

LUCIA KOHLHAUF, ULRIKE RUTKE UND BIRGIT NEUHAUS

Entwicklung eines Kompetenzmodells zum biologischen Beobachten ab dem Vorschulalter

Development of a competency model on biological observation suitable from preschool age

ZUSAMMENFASSUNG

In den Naturwissenschaftsdidaktiken existieren bereits Kompetenzmodelle zur Erkenntnisgewinnung durch Experimentieren (z. B. Grube et al., 2007; Hammann, 2007). Obwohl die Biologie ihre Erkenntnisse keineswegs nur durch das Experiment gewinnt, steht das Beobachten als wissenschaftliche Erkenntnismethode meist im Schatten des Experiments. Jedoch ist auch das Beobachten als eigenständige biologische Arbeitsmethode zu sehen. Als im Jahr 2005 deutschlandweit Bildungs- und Erziehungspläne für Kindertageseinrichtungen verabschiedet wurden, hielt die Diskussion um Kompetenzen und Kompetenzförderung, die seit der Einführung der Bildungsstandards für den Mittleren Schulabschluss die Bildungsgespräche im Schulbereich dominiert, auch Einzug in den Vorschulbereich. Eine entwicklungsangemessene individuelle Förderung der Beobachtungskompetenz scheint bereits im Vorschulalter möglich, da Kinder ab ca. vier Jahren dazu in der Lage sind, selbstständig Fragen und Hypothesen zu generieren (Sodian & Thoermer, 2006). Da hierzu die Feststellung der aktuellen Kompetenz des Kindes anhand eines Kompetenzmodells, das schon ab dem Vorschulalter eingesetzt werden kann, nötig ist, wurde ein Modell zur Erfassung der Beobachtungskompetenz entwickelt und empirisch evaluiert. Anhand des Modells sollen in Zukunft Module zur individuellen Förderung der Beobachtungskompetenz im Vorschulalter entwickelt werden, auf die die Grundschule aufbauen kann.

Schlüsselwörter: Vorschule, Beobachtungskompetenz, Kompetenzmodell

ABSTRACT

It has been decided in German preschool curricula in 2005, that children should be exposed to natural science from preschool. Due to these developments, the topic of competence and competence training, which have dominated educational debates in schools since 2005, were introduced into the preschool level. Though the scientific education competency models in experimentation already exist (e.g., Grube et al., 2007; Hammann, 2007), biology does not always use experimentation to uncover new information, using instead, observation as a scientific research method, which is often shadowed by experimentation.

However, observing must be seen as an independent biological research method. An age-appropriate individual promotion of the observation competency seems possible as early as pre-school age, because children from approximately four years of age are

able to generate questions and hypotheses independently (Sodian & Thoermer, 2006). To foster this competency individually, it is necessary to develop and to evaluate a competency model, which can be used at the preschool level.

Based on this model, training units will then be developed which promote pupils' observational abilities step-wise.

Keywords: Preschool, observation competency, competency modell

1 Einleitung

Seit bereits mehr als fünf Jahren gibt es Bildungsstandards für den Mittleren Schulabschluss im Fach Biologie, die in den vier Bereichen Fachwissen, Erkenntnisgewinnung, Kommunikation und Bewertung diejenigen Kompetenzen beschreiben, die Lernende am Ende der Mittelstufe im jeweiligen Fach erlangt haben sollten (KMK, 2005). Was Kompetenz meint, ist am ehesten über die Definition von Weinert (2001) verständlich, der Kompetenzen als diejenigen verfügbaren oder erlernbaren kognitiven Fähigkeiten und Fertigkeiten bezeichnet, die nötig sind, um bestimmte Probleme zu lösen. Hinzu kommen die damit verbundenen motivationalen, volitionalen und sozialen Fähigkeiten, das Problemlösen in verschiedenen Situationen erfolgreich nutzen und anwenden zu können (Weinert, 2001). Um diese Kompetenzen erfassen und individuelle Maßnahmen zur Förderung beim Einzelnen ergreifen zu können, sind empirisch überprüfte Kompetenzstruktur- und -entwicklungsmodelle nötig. Da Weinert von einer Domänenspezifität der Kompetenzen

ausgeht, müssen für jedes Fach derartige Modelle entwickelt werden, bevor nach fächerübergreifenden Gemeinsamkeiten gesucht werden kann (Neumann, Kaurertz, Lau, Notarp & Fischer, 2007). Im Bereich der Biologie existieren bereits erste Kompetenzmodelle für die Erkenntnisgewinnung durch Experimentieren (z. B. Grube et al., 2007; Hammann, 2007). Da die Wissenschaft Biologie ihre Erkenntnisse jedoch nicht nur durch das Experiment gewinnt – Aristoteles, Hippokrates oder Darwin sind nur einige große Namen aus der Wissenschaftshistorie des Faches, deren Erkenntnisse auf wissenschaftlichen Beobachtungen beruhen – muss das wissenschaftliche Beobachten als eigenständige wissenschaftliche Erkenntnis-methode angesehen werden, sodass auch Kompetenzmodelle in diesem Bereich erforderlich sind.

Ein systematisches Einüben dieser wissenschaftlichen Arbeitsmethode sowie deren Reflexion ist deshalb unabdingbar im Schulfach Biologie.

Da mit der Verabschiedung der Bildungs- und Erziehungspläne für Kindertageseinrichtungen ab dem Jahr 2005 in

Deutschland erneut die Diskussion auflebte, ob und inwieweit bereits Kinder unter sechs Jahren in Kindertageseinrichtungen entwicklungsangemessen gefördert werden sollten, sollte auch über eine Förderung der Beobachtungskompetenz ab dem Vorschulalter nachgedacht werden. Gemäß neuer Erkenntnisse aus der Theory-of-Mind-Forschung (Sodian & Thormer, 2006) sind Kinder ab ca. 4 Jahren in der Lage, zwischen der Meinung eines anderen und der eigenen Meinung zu unterscheiden. Wenn Kinder diese Perspektivenübernahme durchführen können, ist ihre Theory-of-Mind entwickelt. Da es für die Fähigkeit zur Hypothesenbildung wichtig ist, zu verstehen, dass viele mögliche Vermutungen zu einer Fragestellung existieren können, ist ca. ab diesem Alter mit einer eigenständigen Hypothesenbildung zu rechnen. Weil Kinder zudem eine natürliche Neugier besitzen, die sie zum Erkenntnisgewinn motiviert (Sacher, 2006), scheint eine individuelle Förderung der Beobachtungskompetenz ab dem Vorschulalter möglich und angemessen. Damit eine solch entwicklungsangemessene Förderung gewährleistet werden kann, ist es nötig, die aktuelle Kompetenz des Einzelnen anhand eines Kompetenzmodells eruieren zu können. Um aber der sicher in einigen Fällen gerechtfertigten Kritik an der Erstellung von Kompetenzmodellen und besonders deren Einsatz im Vorschulbereich (Grochla, 2009) zu begegnen, sei betont, dass hierbei keinesfalls eine Förderung angestrebt wird, die sich am Bedarf unserer Wirtschaftsgesellschaft orientiert, sondern vielmehr eine bestmögliche individuelle Förderung,

die ihren Ausgangspunkt im Individuum selbst nimmt, Selbstzweck ist und sich einzig am Bedarf ebendieses Menschen orientiert. Nur wenn man um die Fähigkeiten des Einzelnen weiß und ihn wahrnimmt, ist es möglich, der Heterogenität der Bevölkerung zu begegnen, Schwächen zu erkennen, sowie Stärken und Interessen zu fördern.

2 Beobachten

Um ein Kompetenzentwicklungsmodell im Bereich des wissenschaftlichen Beobachtens erstellen zu können, bedarf es zuerst einer genauen Begriffsklärung und einer detaillierten theoretischen Annäherung an die Arbeitsweise des wissenschaftlichen Beobachtens aus psychologischer, biologischer und fachdidaktischer Sicht.

2.1 Theoretische Klärung eines Alltagsbegriffes

Oft werden im alltäglichen Sprachgebrauch die Wörter Beobachten, Wahrnehmen, aktives oder passives Aufnehmen von Sinneseindrücken nahezu als Synonym für ein und dieselbe Tätigkeit verwendet, wobei im Begriff Beobachten noch am ehesten eine aufmerksame, konzentrierte, zielgerichtete Komponente mitschwingt. Im Alltag ist eine genaue Begriffsklärung nicht nötig; ganz im Gegensatz zur Wissenschaft, in der eine klare Definition dessen, was eine wissenschaftliche Beobachtung von einer Wahrnehmung, von einer Alltagsbeobachtung,

aber auch vom Experiment unterscheidet, unerlässlich, jedoch bisher kaum erfolgt ist.

2.2 Abgrenzung zur Erkenntnis- methode des Experimentierens

Auch wenn die Übergänge zwischen den hypothetisch-deduktiven Erkenntnis-
methoden oft fließend erscheinen, ist es für die Erstellung eines Kompetenzmodells wichtig, so genau wie möglich zwischen ihnen zu differenzieren. Während für das Experiment die Produktion von Daten wesentlich ist, soll Beobachten eher eine Datenerhebung sein (Martin & Wawrinsowski, 2006). Beim biologischen Beobachten werden Merkmale, räumliche Beziehungen und zeitliche Abfolgen einer Erscheinung ermittelt, ohne dabei grundlegend verändernde Eingriffe an Objekten oder Prozessen vorzunehmen (Dietrich et al., 1979), wohingegen das Experimentieren das Eingreifen in Objekte oder Zusammenhänge erfordert. Es wird eine Messgröße unter künstlich hergestellten Umständen beobachtet – hier ist die Beobachtung eine Phase des Experiments –, aber zusätzlich auch eine Einflussgröße isoliert und systematisch variiert (Gropengießer & Kattmann, 2006). Für eine Beobachtung, z. B. eine biologische Verhaltensbeobachtung, ist dagegen keine künstliche Situation erforderlich. Sie ist auch unter natürlichen Bedingungen möglich. Das Nicht-Eingreifen und Nicht-Manipulieren grenzt das wissenschaftliche Beobachten damit von den anderen Erkundungsformen ab.

Manchmal kann es notwendig sein, Situationen zu schaffen, um gewisse Details genauer beobachten zu können, doch je mehr ein Beobachter versucht, auf die gegebenen Bedingungen Einfluss zu nehmen, desto mehr nähert er sich einer experimentellen Situation (Martin & Wawrinsowski, 2006).

2.3 Wahrnehmen, unsystematisches und systematisches Beobachten

Da Wahrnehmen alltäglich für jeden von uns ist, scheint auch Beobachten trivial. Doch muss Beobachten etwas anderes sein als passives Wahrnehmen; da sich prinzipiell alle Dinge und Erscheinungen als komplexe Gebilde darstellen und eine Beobachtung ohne eine Zerlegung in einzelne Merkmale nicht möglich ist, stellt die Beobachtung eine kognitive Leistung dar. Dies betont bereits Georg Kerschensteiner (1914), der damit denkendes Beobachten als Tautologie bezeichnet.

Auch in der Psychologie wird der Beobachtungsprozess von der Wahrnehmung abgegrenzt und in unsystematische Alltagsbeobachtung und systematische Beobachtung unterteilt (vgl. hierzu Bortz & Döring, 1995; Oguz & Yurumezoglu, 2007; Martin & Wawrinsowski, 2006). Obwohl Wahrnehmen und Beobachten meist nur mit visuellen Vorgängen assoziiert werden, sind beides Prozesse, die mit vielen Sinnesorganen ablaufen können und sollen (vgl. z. B. Becker-Textor, 1992; Österreicher, 1999; Martin & Wawrinsowski, 2006). Wahrnehmen meint dabei in einem passiven, beiläufigen Prozess mit den Sin-

nesorganen all das aufzunehmen, das ins eigene Umfeld gerät. Es kann zwar nicht als wissenschaftlich eingestuft werden, ist aber nichtsdestotrotz bereits von großer Bedeutung, da es die nötigen Voraussetzungen für das Beobachten schafft.

(Vor-)wissenschaftliches Beobachten aber ist ein aktives Sich-Aneignen von Wirklichkeit (Mollenhauer und Rittelmeyer, 1977), das absichtlich und aufmerksamselektiv den Fokus auf bestimmte Details richtet und dabei ganz bewusst andere Dinge außer Acht lässt. Beobachten ist von einer Suchhaltung, einem Erkenntnisinteresse, geprägt und zielgerichtet, da es von Anfang an darauf ausgelegt ist, dieses Erkenntnisinteresse zu befriedigen und so zu einem auswertbaren Ergebnis zu gelangen (Graumann, 1966; Martin & Wawrinowski, 2006).

Unsystematisches Beobachten läuft noch sehr subjektiv und relativ beliebig, oft spontan und damit ohne vorher festgelegte Regeln ab; allerdings darf es im Gegensatz zur Wahrnehmung nicht mehr als unwissenschaftlich eingestuft werden (Martin & Wawrinowski, 2006, Bortz & Döring, 1995, Sumaski, 1977). Am Beginn verhaltensbiologischer Beobachtungen steht z. B. meist die Neugier, mehr über konkrete Zusammenhänge zu lernen. Oft führen erst diese vorwissenschaftlichen Beobachtungen zu Fragen, die der Beobachter gerne beantwortet wissen möchte.

Beim Systematischen Beobachten muss vor Beginn feststehen, welche Frage konkret geklärt werden soll. Auch wird von Anfang an eine mögliche Antwort auf diese Frage, eine wissenschaftliche Hypothese, aufgestellt und die Absicht verfolgt,

diese zu prüfen; wissenschaftliches Beobachten muss dabei in Bezug auf Planung und Bewertung den Kriterien der Selektion, Abstraktion, Klassifikation, Systematisierung und Relativierung genügen (Feger, 1983; Bortz & Döring, 1995). Das zu Beobachtende muss aus einer Vielzahl wahrnehmbarer Reize selektiert und auf seine wesentliche Bedeutung reduziert werden (vgl. auch Sjøberg, 2007). Zwar ist auch Wahrnehmen immer selektiv, da nicht alle Sinneseindrücke vom Gehirn gleichzeitig aufgenommen und verarbeitet werden können (Sumaski, 1977), doch während beim Wahrnehmen ein unbewusster Selektionsvorgang abläuft, zeichnet sich wissenschaftliches Beobachten durch eine bewusste Auswahl der für die Beantwortung der Fragestellung und Testung der Hypothese essentiellen Details aus (Sumaski, 1977).

Die selektierten Merkmale müssen anschließend klassifiziert und relativiert, d. h. in einen übergeordneten theoretischen Rahmen eingebettet werden (Bortz & Döring, 1995). Eine systematische Beobachtung rückt spezifische Details in den Fokus der Aufmerksamkeit, läuft unter standardisierten Bedingungen ab, soll intersubjektiv überprüfbar sein und schließlich dokumentiert werden (Sumaski, 1977).

Zusammenfassend wird deshalb in der hier beschriebenen Studie von den folgenden drei Ausprägungen der Beobachtungskompetenz ausgegangen: Wahrnehmung, Unsystematisches Beobachten und Systematisches Beobachten. Da das Wort ‚wahrnehmen‘ jedoch sehr viele alltagssprachliche Konnota-

tionen trägt, wurde für diese Ausprägung im Rahmen der vorliegenden Studie die Bezeichnung des Inzidentellen Beobachtens gewählt, um ungenaue Assoziationen zu vermeiden. Der Begriff ist in Anlehnung an Rolf Oerters Inzidentelles Lernen (Oerter, 1996) entstanden, der beiläufiges Lernen meint, das halb oder nicht bewusst abläuft. Nachdem auch das Wahrnehmen im Kontext der Beobachtungskompetenz einen passiven, beiläufigen und eher zufälligen Prozess bezeichnet, lässt sich das Inzidentelle Beobachten daraus ableiten.

2.4 Wissenschaftliches Beobachten als komplexe Tätigkeit

Beobachten ist als komplexe, vielschichtige Tätigkeit zu sehen (Oguz & Yurumezoglu, 2007). Zur Erarbeitung eines Kompetenzmodells gilt es, diese Komplexität in Dimensionen zum Ausdruck zu bringen. Zur Entwicklung eines Kerncurriculums für den Biologieunterricht in der gymnasialen Oberstufe wurden von Harms, Mayer, Hammann, Bayrhuber und Kattmann (2004) Teilkompetenzen erarbeitet, deren Ausbildung die Entwicklung einer biologischen Kompetenz ermöglichen soll. Diese wissenschaftspropädeutischen Teilkompetenzen reichen von Konvergentem und Divergentem Denken, Denken in Systemen, Denken in mentalen Modellen und Denken über Fächergrenzen hinweg (Ebene Naturwissenschaftliches Denken) über naturwissenschaftliches Argumentieren und Verbalisieren eines Sachverhaltes (Ebene Natur der Naturwissenschaft) bis hin zum Bewerten (Ebene Wissenschaft

und Gesellschaft). Diese generellen Dimensionen wurden mit den angesprochenen Merkmalen einer Beobachtung in Beziehung gesetzt, um Teilkompetenzen spezifisch für biologisches Beobachten zu entwickeln. Da beim Beobachten Fragestellung und Hypothesenbildung eine wichtige Rolle spielen, ist sowohl divergentes als auch konvergentes Denken nötig. Um Ergebnisse zu interpretieren, muss naturwissenschaftlich argumentiert werden und da die Beobachtung dokumentiert werden soll, ist Verbalisieren unerlässlich. Hiervon ausgehend wurden die Dimensionen Beschreiben, Fragen, Vermuten, Testen und Interpretieren als essentiell für biologische Beobachtungen angenommen und sollen empirisch überprüft werden.

Beschreiben

Während Fragen, Vermuten, Testen und Interpretieren Teilkompetenzen sind, die allen wissenschaftlichen Erkenntnismethoden eigen sind, ist das objektive Beschreiben von Details ohne selbst eingreifen bzw. handelnd aktiv sein zu dürfen eine wichtige Komponente allein des Beobachtens. Den beobachteten Gegenstand oder Prozess verbalisieren und letztendlich dokumentieren zu können, ist eine wesentliche Kompetenz des wissenschaftlichen Beobachtens, da Sprache bzw. jegliche Form der Äußerung zum einen den Grad des Verständnisses für den jeweiligen Sachverhalt reflektiert (vgl. Sjøberg, 2007) und außerdem eine Dokumentation des Beobachteten ermöglicht. Weil Sprache und Literalität wichtige Werkzeuge der

Wissenschaft darstellen, sind gute Sprachfertigkeiten bei wissenschaftlichen Tätigkeiten unabdingbar; Sprache fungiert als grundlegendes Vermittlungselement des Beobachteten und übernimmt eine aufmerksamkeitsleitende Funktion, indem der Fokus eines Anderen auf die wesentlichen Details gelenkt wird (Lück, 2003). Wesentlich beim Beschreiben ist es demnach nicht nur, Details wahrzunehmen, sondern auch den Fokus auf spezifische Details zu richten und die Beobachtungen in geeignete Worte zu fassen.

Obwohl der Nachteil der Verwendung von Alltagssprache beim Beschreiben von Beobachtungen ist, dass sie theoriefrei erscheint und unscharfe Begriffe enthält, ist Alltagssprache diejenige Sprachebene, die letztlich für alle Menschen verständlich ist (Greve & Wentura, 1997). Je nach Alter und Vorwissensstand des Beobachters sollte deshalb beim Beschreiben von Beobachtungen zwar auf eine Verwendung von präzisen Fachtermini Wert gelegt werden, eine gute und richtige Verwendung von Alltagssprache jedoch auch ihre Berechtigung behalten (Greve & Wentura, 1997).

Wissenschaftliches Denken: Fragen, Vermuten und Testen

Wer wissenschaftlich beobachtet, sucht aktiv nach Lösungen für vorher aufgeworfene Fragen, das heißt, er muss Forschungsfragen und zugehörige Hypothesen aufstellen und diese überprüfen können.

Grundlage hierfür ist zwischen Hypothese und Evidenz unterscheiden und aus Überzeugungen Handlungsvorhaben ableiten

zu können (Sodian & Thoermer, 2006). Essentiell dafür ist, zu verstehen, dass verschiedene Menschen evtl. unterschiedliche Vermutungen zur selben Frage anstellen werden. Wie eingangs erwähnt, beginnt dies in der kindlichen Entwicklung mit ca. vier Jahren (Sodian & Thoermer, 2006), sodass ab diesem Alter mit Kompetenzen im Bereich des wissenschaftlichen Denkens gerechnet werden kann.

Die Deutung einer Beobachtung

Beim naturwissenschaftlichen Argumentieren, also in diesem Fall beim Auswerten und Deuten einer Beobachtung, müssen hypothesengeleitete Bewertungen von Untersuchungsergebnissen vorgenommen werden. Denn Ziel einer Interpretation ist es, die erhobenen Daten in Bezug auf die eingangs gestellte Hypothese zu bewerten und zu entscheiden, ob diese verifiziert oder falsifiziert werden muss (Naguib, 2006); hierbei ist darauf zu achten, dass Beobachtung und Deutung klar voneinander getrennt werden. Auch in dieser Dimension ist ein gewisses Maß an Kreativität erforderlich, muss divergent gedacht werden, um eine Beobachtung deuten zu können. Zudem muss der Beobachter sich dessen bewusst sein, dass eine Deutung immer subjektiv geprägt ist, während eine Beobachtung rein neutral und objektiv abläuft (Schuster & Leland, 2008). Um diese Unterscheidung treffen zu können, ist also auch hier die Fähigkeit essentiell, zwischen Evidenz und Überzeugung zu unterscheiden, was voraussetzt, dass bereits eine Theory-of-Mind ausgebildet ist.

3 Zielsetzung

Zielsetzung des hier dargestellten Projektes ist es, auf Basis der beschriebenen Theorie ein Modell zu erarbeiten und empirisch zu überprüfen, das die unterschiedlichen Kompetenzausprägungen und Teilkompetenzen im Bereich des wissenschaftlichen Beobachtens beschreibt, sodass im Vorschulbereich Maßnahmen zu deren Förderung beim Einzelnen ergriffen werden können. Aus der in der Theorie vorgestellten Literatur wurden die nachstehenden Hypothesen entwickelt:

H1: Als Strukturen der Beobachtungskompetenz lassen sich die Dimensionen Beschreiben, Fragen, Vermuten, Testen und Interpretieren unterscheiden (Harms et al., 2004; Naguib, 2006; Oguz & Yurumezoglu, 2007).

H2: Es lassen sich in jeder Dimension die drei Ausprägungen Inzidentelles Beobachten, Unsystematisches Beobachten und Systematisches Beobachten unterscheiden (Graumann, 1966, Sumaski, 1977; Bortz & Döring, 1995; Martin & Wawrinowski, 2006; Oguz & Yurumezoglu, 2007).

Hieraus ergibt sich zusammenfassend nachstehendes Modell, das im Folgenden empirisch überprüft werden soll (Abb. 1).

4 Methodisches Vorgehen

Basierend auf den theoretischen Vorüberlegungen zu Ausprägungen und Dimensionen der Beobachtungskompetenz wurde

ein Testinstrument zu deren Erhebung entwickelt. Der Test erfolgte als mündlich-praktischer Test, der die Beobachtung von drei ausgewählten Tieren umfasste und mittels Videoanalyse ausgewertet wurde. Er wurde als Interview im offenen Antwortformat entworfen, anhand eines Kodiermanuals von zwei unabhängigen Ratern bewertet und mittels Rasch-Analyse ausgewertet, um die Ausprägungen und Dimensionen der Beobachtungskompetenz ermitteln zu können.

Zur Ausschärfung der Items und Standardisierung der Interviewsituation wurden mehrere Vorpilotierungen durchgeführt, bevor der Test in der eigentlichen Stichprobe zur Anwendung kam.

4.1 Stichprobe

Um eine möglichst hohe Varianz der Stichprobe zu erreichen, wurden Teilnehmer von 4 – 29 Jahren aus verschiedenen Bildungseinrichtungen getestet. Von den 110 Probanden besuchten 27 verschiedene Kindergärten in einer ländlichen Umgebung im Landkreis Miesbach, 27 die örtliche Grundschule der Stadt Miesbach und 28 das Gymnasium der Stadt Tegernsee. Die restlichen 28 Teilnehmer waren Biologie-Lehramtsstudenten verschiedener Semester der Ludwig-Maximilians-Universität in München.

Die Durchführung wurde in den Kindergarten- bzw. Schulalltag eingebunden, verlief aber auf freiwilliger Basis. Die jüngsten Kinder wurden mit Hilfe einer Handpuppe mit der Situation vertraut gemacht und immer wieder durch die Puppe

Dimension	Beschreiben	Wissenschaftliches Denken			Interpretieren
		Fragen	Vermuten	Testen	
Ausprägung					
Systematisches Beobachten	Beschreiben vieler Details; starker Fokus auf spezifische Details	selbstständig	selbstständig	selbstständig	Beobachtung und Deutung werden getrennt
Unsystematisches Beobachten	Beschreiben von wenigen Details; wenig Fokus	nach Aufforderung	nach Aufforderung	nach Aufforderung	Beobachtung und Deutung werden kombiniert
Inzidentelles Beobachten	Wahrnehmen ohne Details	keine Fragestellung	keine Hypothesenbildung	kein Testen der Hypothese	keine Deutung

Abb. 1: Theoretisch-hypothetisches Modell zur Beobachtungskompetenz.

ermutigt – eine Herangehensweise die aus einigen Gründen sinnvoll erscheint: Der Einsatz von sogenannten Living-Puppets, die oft in psychologischen Interviews mit Kindern (Krüger, 2006; Krüger & Grunert, 2001) und im schulischen Bereich vor allem im Sprachunterricht Verwendung finden, wird nun auch im naturwissenschaftlichen Unterricht immer populärer. Neuesten Forschungen zufolge setzt der Lehrer naturwissenschaftliche Problemstellungen mit Hilfe der Puppe mehr in einen narrativen Kontext, was zur verstärkten Mitarbeit anregt. Generell fördert die Puppe das Gespräch, da Kommunikation mehr auf Augenhöhe ablaufen kann (Simon, Naylor, Keogh, Maloney & Downing, 2008). Auch positive Effekte auf die Gesamtmotivation für naturwissenschaftlichen Unterricht konnten festgestellt werden (Simon et al., 2008).

4.2 Messinstrument

Die Teilnehmer wurden darum gebeten, sich in die Rolle eines Biologen hineinzuversetzen, der ein Tier beobachtet. Die drei Tiere (Fisch, Weinbergschnecke, Rennmaus) wurden der Testperson daraufhin nacheinander präsentiert. Diese sollte nun in der Rolle eines Biologen zu jedem Tier erklären, wie sie als Beobachter vorgehen würde und dementsprechend handeln. Papier, Stifte, Lineal, Waagen, Thermometer, Stoppuhr und Lupe lagen zum fakultativen Einsatz bereit. Die Kindergartenkinder wurden bei diesem Rollenspiel von der Puppe begleitet, die die Kinder inspirierte Fragen zu stellen und Neues über das jeweilige Tier in Erfahrung zu bringen. Wollten die Kinder mit Messinstrumenten arbeiten, wurde Ihnen der Gebrauch wenn nötig erklärt. Da

eine grobe Vorstellung von Zahlen und Abständen bereits bei allen Kindern vorhanden war, fiel die einfache Verwendung den Meisten nicht schwer. Zudem waren bei jedem Tier vielfältige Beobachtungen auch immer ohne Messinstrumente möglich. Ähnliches gilt für den Umgang mit dem unterschiedlichen Verhalten der Tiere: In einer relativ authentischen Beobachtungssituation ergeben sich unterschiedlich inspirierende Konstellationen. Da die Möglichkeiten für Beobachtungen jedoch derartig mannigfaltig sind, dürfte das Testergebnis dadurch nicht stark verzerrt worden sein.

Konnte der Proband mit der Situation nicht umgehen, wurde schrittweise Hilfe durch den Interviewer angeboten. Der Teilnehmer wurde vorerst gebeten, das jeweilige Tier zu beschreiben, um ihm, durch die intensivere Auseinandersetzung mit dem Objekt, Zeit zu lassen, eine Forschungsfrage aufzuwerfen. Wurde selbstständig keine Frage aufgestellt, wurde er explizit darauf hingewiesen, sich eine Frage zu überlegen, die er anhand einer eigenen Beobachtung klären zu können meinte. Konnte auch daraufhin keine

Frage generiert werden, wurde sie vom Interviewer vorgegeben. Ebenso wurde beim Aufstellen und Prüfen der Hypothese vorgefahren. Außerdem wurde der Teilnehmer gebeten, seine Beobachtungen beim Prüfen der Hypothese zu beschreiben, wobei darauf geachtet wurde, ob er das Ergebnis auswertete und wenn ja, ob dabei zwischen Beobachtung und Deutung differenziert werden konnte. Auf diese Weise war es möglich den Test in 15 Items, 5 pro Tier, zu unterteilen (Tab. 1).

Der gesamte Test dauerte je nach Alter, Fähigkeit und Motivation des Teilnehmers ca. 30 min und wurde auf Video aufgezeichnet, um eine Störung des Gesprächs durch Notieren der Antworten zu vermeiden und zudem eine spätere Auswertung anhand eines Kodiermanuals durch einen zweiten unabhängigen Rater zu ermöglichen. Um jedoch auch qualitative Unterschiede innerhalb der einzelnen Kategorien (0, 1, 2; siehe Einteilung der Ausprägungen in Abb. 1) unterscheidbar zu machen, wurden sie teilweise in mehrere Subkategorien, sogenannte Rubriken (01 – 23) unterteilt. Die Rubriken spiegeln in der Dimension Beschreiben die

Tab. 1: Items des Beobachtungstests hier am Beispiel der Schnecke; Fisch und Rennmaus analog

Items des Beobachtungstests
1. Was möchtest du gern über die Schnecke wissen? Überleg dir eine Frage, die du durch Beobachten erforschen kannst!
2. Was vermutest du?
3. Teste deine Vermutung!
4. Beschreibe deine Beobachtung und die Schnecke so genau wie möglich!
5. Was konntest du beobachten?

genauere Anzahl der genannten Details, bzw. deren Qualität (spezifisch oder unspezifisch) wider, während in der Teilkompetenzen Vermuten anhand der Rubriken eine Aussage darüber möglich ist, ob die Hypothesenbildung auf eine vorgegebene, die eigene oder vollkommen ohne Fragestellung erfolgte. Die Rubriken der Dimension Interpretieren lassen zusätzlich erkennen, ob die Deutung nach Aufforderung oder selbstständig erfolgte. Um zusätzlich den Einfluss von Vorwissen, Interesse oder Sprachfertigkeit auf die so erhobene Beobachtungskompetenz feststellen zu können, wurden auch hierzu Daten erhoben. Eine Analyse mittels Strukturgleichungsmodellen ist bereits in Planung.

4.3 Datenauswertung

Die Daten wurden mit WINSTEPS (Mc Linacre, 2009), einer Software zur Berechnung probabilistischer Testmodelle ausgewertet. Dabei wird die Wahrscheinlichkeit einer korrekten Antwort als logistische Funktion der Differenz von Personen- und Itemparametern modelliert. Dies liefert Informationen darüber, wie gut das jeweilige Item die Fähigkeit der einzelnen Person beschreibt. Da ein plausibler Zusammenhang zwischen Personenfähigkeit und Itemlösungswahrscheinlichkeit angenommen wird, kann zudem für jede Person eine konkrete Verhaltensvorhersage getroffen werden, indem ermittelt wird, mit welcher Wahrscheinlichkeit eine Person bei Kenntnis der Itemschwierigkeit und der Personenfähigkeit ein Item löst

(Bühner, 2008). Pro Item wurden 3 Hauptkategorien festgelegt (0, 1, 2), sodass bei einer vollkommen richtigen Lösung die volle Punktzahl, bei einer teilweise richtigen Lösung 1 Punkt und kein Punkt für eine unzureichende Antwort vergeben wurde. Da es diese Software erlaubt, auch bei kleinen Stichprobengrößen ($N \geq 30$), sinnvolle Aussagen über die Daten zu treffen (Mc Linacre, 1994), erschien WINSTEPS zur Analyse der vorliegenden Daten der relativ kleinen Stichprobe von 110 Probanden geeigneter als beispielsweise das Programm ConQuest (Wu, Adams, Wilson & Haldane, 2007), das zur Interpretation der Daten von PISA verwendet wurde.

WINSTEPS wendet zur Parameterschätzung die JML-Methode (Joint Maximum-Likelihoodmethode) an, die zwar Personen- und Itemparameter simultan schätzt, jedoch für kleine Stichproben empfohlen wird (Bühner, 2008, S. 338).

Die Reliabilitätsbestimmung eines Tests erfolgt bei WINSTEPS anhand der Personen-Reliabilität, deren Wert dem Cronbachs- α der klassischen Testtheorie entspricht und genauso interpretiert werden kann. Über eine fit-Statistik kann festgestellt werden, ob und wie geeignet das jeweilige Item ist, bzw. für welche Personen der Test passt. Outfit-Werte, die auf den konventionellen Chi-Quadrat-Statistiken basieren, sind hierbei eher sensitiv für Outlier, das heißt deuten darauf hin, dass ein Item für eine Person entgegen der Personenfähigkeit besonders leicht oder schwer zu lösen war. Hohe Outfit-MNSQ-Werte resultieren somit aus einzelnen (zufällig) richtigen Antworten von Personen niedriger Personenfähigkeit. Infit-Werte sind dage-

gen Indikatoren für unerwartete Antwortmuster und deuten darauf hin, dass Items für Zielpersonen nicht passen, was für die Gültigkeit der Daten ein größeres Problem darstellt (Mc Linacre, 2010). Items oder Personen mit Infit- bzw. Outfit-MNSQ-Werten ≤ -1.5 oder ≥ 1.5 sollten deshalb einer genaueren Analyse unterzogen werden. Hierzu sollten die z -standardisierten (ZSTD-)Werte, die die Wahrscheinlichkeit der Hypothese ‚die Daten passen perfekt auf das theoretische Modell‘ widerspiegeln, betrachtet werden. ZSTD-Werte sollten ≤ 2 sein, um sicherzugehen, dass Items oder Personen das Messergebnis nicht negativ beeinflussen (Mc Linacre, 2010). Solch schlecht geeignete Items sollten verworfen werden. Eine derartige Misfit-Diagnose wurde auf alle Items wie Personen des vorliegenden Datensatzes angewandt. Zudem wurden die Schwellenparameter der einzelnen Items zur Analyse der Daten herangezogen. Über diese Parameter (Item-measure-Werte, d. h. Itemschwierigkeiten), ist es möglich zu erkennen, ob die Kategorien bei jedem Item in der richtigen Reihenfolge vorliegen, d. h. ob Kategorie 1 tatsächlich schwieriger als 0 und 2 schwieriger als 1 ist.

Zum Nachweis verschiedener Dimensionen im Datensatz, wurden die von WINSTEPS ausgegebenen Inter-Item-Korrelationen einer klassischen Faktorenanalyse mit SPSS unterzogen.

5 Ergebnisse

Zur Überprüfung der Interraterreliabilität wurden wie üblich 10 % der Daten, d. h. in

diesem Fall 12 der 110 Videos, von zwei unabhängigen Ratern kodiert. Die Schulung des zweiten Raters erfolgte anhand eines Beispielvideos und eines Auswertleitfadens. Da die Daten auf Ebene der Antwortkategorien (0,1,2) mindestens ordinal skaliert sind, wurde die Beurteilerübereinstimmung mittels SPSS (SPSS Inc., 2009) über die Rangkorrelationskoeffizienten Spearman's rho (alle Skalen: $p = .77$, $p \leq 0.001$; Skala Beschreiben: $p = 0.87$, $p \leq 0.001$; Skala Wissenschaftliches Denken: $p = 0.72$, $p \leq 0.001$; Skala Interpretieren: $p = .80$, $p \leq 0.001$) und Kendalls tau (alle Skalen: $\tau = 0.74$, $p \leq 0.001$; Skala Beschreiben: $\tau = 0.84$, $p \leq 0.001$; Skala Wissenschaftliches Denken: $\tau = 0.69$, $p \leq 0.001$; Skala Interpretieren: $\tau = 0.75$, $p \leq 0.001$) bestimmt (Wirtz & Caspar, 2002). Alle Werte lassen auf eine ausreichende Interrater-Reliabilität schließen. Dass die Übereinstimmung trotz kurzer Schulung ausreichend ausfällt, spricht für die gute Handhabbarkeit des Tests samt Kodiermanual.

Die mit WINSTEPS errechnete Personenreliabilität für den Gesamttest (15 Items) beträgt .85 ($p \leq 0.001$; $N = 110$) und ist damit nach den Kriterien für die Höhe von Cronbachs- α für einen reliablen Test als ausreichend hoch einzustufen (Bühner, 2006, 139f.). Für die einzelnen Skalen ergeben sich zusätzlich folgende Testgütewerte, die alle bis auf den Reliabilitätswert der Skala Interpretieren, der einzig durch die sehr niedrige Itemzahl ($n = 3$) bei geringer Stichprobe ($N = 110$) erklärt werden kann, in annehmbaren Bereichen liegen (Tab. 2). Bei einer Misfit-Analyse der einzelnen Testpersonen wurden zu hohe Infit/Outfit-MNSQ-Werte für 12 Personen festgestellt;

Tab. 2: Werte zur Testgüte der einzelnen Skalen, bzw. Items des Beobachtungstest

Skala	Anzahl Items	Reliabilität (Person reliability)	Items	Item-Schwierigkeit (Item-measure)	Trennschärfe (pt-measure correlation)
Beschreiben	3	.66	Beschreiben Fisch	-.47	.89
			Beschreiben Schnecke	1.26	.92
			Beschreiben Maus	-.79	.86
Wissenschaftliches Denken	9	.68	Fragen Fisch	-.66	.81
			Fragen Schnecke	-.21	.69
			Fragen Maus	-.69	.74
			Vermuten Fisch	.36	.28
			Vermuten Schnecke	-.19	.36
			Vermuten Maus	-.22	.34
			Testen Fisch	.41	.78
			Testen Schnecke	.71	.74
			Testen Maus	.48	.76
Interpretieren	3	.18	Interpretieren Fisch	.06	.84
			Interpretieren Schnecke	.34	.88
			Interpretieren Maus	-.39	.86

Tab. 3: Misfit-Diagnose zur Skala ‚Vermuten‘

Item	Itembezeichnung	Infit		Outfit	
		MNSQ	ZSTD	MNSQ	ZSTD
7	Vermuten Fisch	1.59	3.9	1.86	4.6
8	Vermuten Schnecke	1.55	3.5	1.91	4.1
9	Vermuten Maus	1.47	3.1	2.10	4.7

in diesen 12 Fällen waren auch die ZSTD-Werte > 2 , so dass davon ausgegangen werden muss, dass das Modell für ca. 11 % der Teilnehmer nicht optimal passt. Eine Abhängigkeit vom Bildungsgrad konnte jedoch nicht ausgemacht werden, da sich diese 11 % aus fast allen Ausbildungsbe-reichen (Vorschule, Grundschule, Gymnasium, Studium) zusammensetzen. Die

12 Teilnehmer wurden daher in der Stichprobe belassen.

Bei der anschließenden Misfit-Analyse der 15 Items stellten sich nur die 3 Items der Skala Vermuten als unpassend heraus (Infit bzw./und Outfit > 1.5 ; ZSTD > 2 ; siehe Tab. 3).

Um zu klären, worauf diese Werte zurückzuführen sind, wurde eine Analyse der

Schwellenparameter durchgeführt. Hier zeigt sich deutlich, dass jeweils die Rubrik 21 (Hypothese wird selbstständig auf vorgegebene Frage formuliert), bei Item 7 auch die Rubrik 23 (Hypothese wird ohne Fragestellung formuliert) niedrigere measure-Werte als Rubriken der Antwortkategorie 1 aufweist, was bedeutet, dass diese Rubrik(en) der Kategorie 2 leichter zu erreichen sind als Rubriken der Kategorie 1 (siehe Tab. 4).

Da es demzufolge einfacher ist, selbstständig Hypothesen auf vorgegebene Fragen zu generieren als auf eigene Fragen, kann in der Dimension Vermuten eine Einteilung der Ausprägungen nach dem gewünschten Schema nicht angenommen werden. Selbstständig Hypothesen auf vorgegebene Fragen zu generieren muss ein niedrigeres Niveau darstellen als selbstständig Hypothesen auf eigene Fragen zu produzieren, sodass die Bezeichnungen Niveau 1 = nach

Tab. 4: Ungeordnete Schwellenparameter der Items zur Skala Vermuten; zum Vergleich dazu geordnete Schwellenparameter eines Items aus der Skala Fragen

Item	Itembezeichnung	Kategorie	Rubrik	Item-measure
7	Vermuten Fisch	0	01	-1.89
		1	11	-1.55
			12	.89
		2	21	-1.07*
			22	1.62
			23	-.22*
missing	missing	1.62		
8	Vermuten Schnecke	0	01	-2.16
		1	11	-1.39
			12	.84
		2	21	-.47*
			22	1.37
			23	1.38
missing	missing	/		
9	Vermuten Maus	0	01	-2.96
		1	11	-1.07
			12	.62
		2	21	-1.00*
			22	1.15
			23	1.35
missing	missing	/		
6	Fragen Maus	0	01	-1.85
		1	11	-.35
		2	21	1.12
		missing	missing	1.26

* Der Measure-Wert steigt nicht mit der Kategorie

Tab. 5: Mustermatrix der Faktorenanalyse (ProMax Rotation mit Kaiser-Normalisierung, Anzahl der Faktoren vorgegeben, Werte < .40 unterdrückt)

Item	Komponente		
	1	2	3
Beschreiben Fisch			.85
Beschreiben Schnecke			.87
Beschreiben Maus			.71
Fragen Fisch	.88		
Fragen Schnecke	.80	-.54	
Fragen Maus	.61		
Vermuten Fisch	-.74		
Vermuten Schnecke	-.78		
Vermuten Maus	-.71		
Testen Fisch	.62		
Testen Schnecke	.84		
Testen Maus	.65		
Interpretieren Fisch		.66	
Interpretieren Schnecke		.82	
Interpretieren Maus		.80	

Aufforderung und Niveau 2 = selbstständig in dieser Dimension nicht mehr haltbar sind. Zudem ist diese Teilkompetenz damit abhängig von der Teilkompetenz Fragen und kann deshalb nicht mehr als eigenständige Dimension gewertet werden. Die Items aus dem Bereich Vermuten müssen deshalb verworfen bzw. modifiziert werden. Zur weiteren Klärung der Dimensionenanzahl wurde eine Faktorenanalyse der Inter-Item-Korrelationen mit SPSS durchgeführt (Tab. 5).

Deutlich lässt sich für die Skala Beschreiben ein eigener Faktor differenzieren, die Items der Skala Interpretieren bilden ebenfalls eine eigene Komponente. Dagegen laden die Items der anderen Dimensionen Fragen, Vermuten und Testen auf einen gemeinsamen Faktor, bilden also gemeinsam eine Teilkompetenz.

6 Diskussion

Die genannten ungeordneten Schwellenparameter der Items der Skala Vermuten führten eindeutig dazu, diese Skala als eigenständige Dimension zu verwerfen. Die Ergebnisse der Faktorenanalyse lieferten zudem eine relativ eindeutige Aufteilung der Skalen in drei unterschiedliche Komponenten.

Einzig das Item *Fragen Schnecke* liefert einen nicht ganz eindeutigen Wert, da es zudem mit einem geringeren Wert auf die Skala Interpretieren lädt. Dies kann darauf zurückzuführen sein, dass ein Teil der Teilnehmer im Test lediglich das Schneckenhaus beobachten konnte, da die Schnecke sich zurückgezogen hatte, was das Generieren der Frage erschwert haben kann.

Aus diesem Grund und da dieser Wert der niedrigste Wert der gesamten Faktorenanalyse ist, scheint es gerechtfertigt, diese Ladung zu vernachlässigen.

Die Faktorenanalyse wirft drei Komponenten der Beobachtungskompetenz auf, sodass von einer eindimensionalen Lösung abgesehen, aber auch Hypothese 1 (*H1: Als Strukturen der Beobachtungskompetenz lassen sich die Dimensionen Beschreiben, Fragen, Vermuten, Testen und Interpretieren unterscheiden.*) nur teilweise bestätigt werden kann. Zwar lassen sich unterschiedliche Dimensionen der Beobachtungskompetenz unterscheiden, doch müssen die drei Teilkompetenzen Vermuten, Fragen und Testen für das endgültige Testinstrument der Beobachtungskompetenz zu neuen Items einer gemeinsamen Teilkompetenz kombiniert werden. Sinnvoll erscheint es, diese drei Komponenten in Anlehnung an die Ebene des Naturwissenschaftlichen Denkens der wissenschaftspropädeutischen Teilkompetenzen nach Harms et al. (2004) unter der gemeinsamen Dimension wissenschaftliches Denken zusammenzufassen. Anhand dieser Ergebnisse kann davon ausgegangen werden, dass sich als Strukturen der Beobachtungskompetenz die drei Dimensionen Beschreiben, wissenschaftliches Denken (mit Fragen/Vermuten/Testen) und Interpretieren unterscheiden lassen.

Dass in allen Teilkompetenzen die drei Ausprägungen Inzidentelles Beobachten, Unsystematisches Beobachten und Systematisches Beobachten vorliegen, wie es Hypothese 2 postuliert, lässt sich anhand der Ergebnisse für vier der fünf vorher an-

genommenen Dimensionen bestätigen, da alle Items der Skalen Fragen, Testen, Beschreiben und Interpretieren geordnete Item-measure-Werte und damit klare Schwellenparameter aufweisen. Für die Skala Vermuten ist dies nicht der Fall, weshalb diese drei Items in ihrer alten Form verworfen werden.

Das anfangs vorausgesetzte theoretische Modell zur Beobachtungskompetenz muss aufgrund der hier beschriebenen Ergebnisse modifiziert werden, kann jedoch in seinen Grundzügen beibehalten werden. Das empirisch gestützte Modell weist drei Teilkompetenzen in je drei Ausprägungen auf (vgl. Abb. 2).

Auch einer Betrachtung vor dem theoretischen Hintergrund hält das empirisch entwickelte Modell stand. Wie in Kapitel 2.4 erläutert, ist das objektive Beschreiben ohne selbst eingreifen zu dürfen auch von einem theoretischen Standpunkt aus eine eigenständige Komponente, die den anderen Teilkompetenzen gegenübersteht. Dass sich auch die Deutung einer Beobachtung als Komponente des wissenschaftlichen Argumentierens, in der wesentlich ist, zwischen Deutung und Beobachtung unterscheiden zu können, von den Teilkompetenzen des wissenschaftlichen Denkens abgrenzen lässt, erscheint genauso wenig verwunderlich. Dass hingegen die drei Komponenten Fragen, Vermuten und Testen so stark miteinander verknüpft und nicht unabhängig voneinander zu sehen sind, war nicht offensichtlich; da sie jedoch auch theoretisch unter der Komponente Wissenschaftliches Denken zusammenfassbar sind, stützen sich auch hier Theorie und Empirie gegenseitig.

Dimension Ausprägung	Beschreiben	Wissenschaftliches Denken Fragen/Vermuten/Testen	Interpretieren
Systematisches Beobachten	Beschreiben von mehreren Details; starker Fokus auf spezifische Details	durchweg selbstständig	Beobachtung und Deutung werden voneinander getrennt
Unsystematisches Beobachten	Beschreiben von wenigen Details; schwacher Fokus auf spezifische Details	teilweise mit Aufforderung	Beobachtung und Deutung werden vermischt
Inzidentelles Beobachten	Wahrnehmung ohne Details	komplett oder meist mit Aufforderung	keine Deutung

Abb. 2: Empirisches Modell zur Beobachtungskompetenz.

Zudem wird die theoretisch angedachte Einteilung wissenschaftlichen Beobachtens in drei Ausprägungen (vgl. z. B. Martin & Wawrinowski, 2006, Bortz & Döring, 1995, Sumaski, 1977), hier als inzidentell, unsystematisch, systematisch bezeichnet, durch das Modell einmal mehr untermauert.

7 Ausblick und Bedeutung für die Praxis

Der Test zur Erfassung der Beobachtungskompetenz wurde in einem gemischten Sample geprüft, um ein allgemeingültiges Kompetenzmodell für das wissenschaftliche Beobachten zu entwickeln. Einsatz finden soll es vorerst aber vor allem bei der Diagnose und Förderung der Beobachtungskompetenz im Vor- und Elementarschulbereich. Der Test soll deshalb an einer größeren Stichprobe dieser Altersklasse erneut überprüft werden. Ausgehend von dem empirisch erarbeiteten Modell werden schließlich Module zur

Förderung der Beobachtungskompetenz im Vorschulbereich entwickelt werden. Der Heterogenität und Individualität der Kinder gerecht zu werden ist hierbei zentral. Dies bedeutet, die Unterschiedlichkeit der einzelnen Kinder zu berücksichtigen und sie in ihrer jeweils nächsten Entwicklungszone (Vygotsky nach Tudge, 1990) agieren zu lassen.

Dies wird anhand des nun vorliegenden Modells einfacher möglich sein, da der Stand jedes einzelnen Kindes in jeder Teilkompetenz erfasst werden kann. Für jede Ausprägung wird in jeder Teilkompetenz ein Modul entwickelt werden, sodass jedes Kind mit genau dem Modul trainieren kann, das in der jeweiligen Dimension in der nächsthöheren Ausprägung seines bisherigen Kenntnisstandes, sprich in seiner Zone der nächsten Entwicklung, liegt.

Wichtig ist hierbei, dass die erarbeiteten Module keine Lerneinheiten darstellen, sondern vielmehr als in den Kindergartenalltag integrierte guided play-activities (Singer, Golinkoff & Hirsh-Pasek, 2006)

erfolgen, sodass spielerisches interesseförderndes Lernen ohne Leistungsdruck möglich ist. Dies scheint am besten möglich, wenn die Erzieher und Erzieherinnen der jeweiligen Kindertageseinrichtung die Vermittlung der Inhalte übernehmen, was eine intensive Zusammenarbeit mit dem Kindergartenpersonal voraussetzt. Zu diesem Zweck sind Fortbildungen für Erzieherinnen und Erzieher geplant und eine enge Zusammenarbeit mit Einrichtungen der Erzieher/innenausbildung angestrebt. Die Entwicklung von Lernspielen für ein unabhängiges Training der einzelnen Dimensionen stellt sicherlich eine Herausforderung dar, doch erste Ideen hierzu – vom Brettspiel über Bewegungs- und Outdoor-Spiele mit Naturmaterialien – liegen jedoch bereits vor und lassen erahnen, dass eine Ausarbeitung durchaus sinnvoll ist. Eine Evaluation der entwickelten Spiele im Rahmen einer Langzeitintervention ist für das Jahr 2011 geplant.

Literatur

- Becker-Textor, I. (1992). Sehen lehren. Sehen lernen. In I. Becker-Textor, *Mit Kinderaugen sehen. Wahrnehmungserziehung im Kindergarten*. Freiburg: Herder, 17–42. Leicht bearbeitete Fassung [online]. Verfügbar unter: <http://www.kindergartenpaedagogik.de/1028.html> [17.12.2009].
- Bortz, J. & Döring, N. (1995). *Forschungsmethoden und Evaluation*. Berlin: Springer.
- Bühner, M. (2006). *Einführung in die Test- und Fragebogenkonstruktion*. München: Pearson Studium.
- Dietrich, G. et al. (Hrsg.) (1979). *Methodik Biologieunterricht*. Berlin: Volk und Wissen.
- Garson, G. D. (2009). *Reliability Analysis*. [online]. Verfügbar unter <http://faculty.chass.ncsu.edu/garson/PA765/reliab.htm> [28.06.2010].
- Singer, D.G., Golinkoff, R.M., Hirsh-Pasek, K. (2006). *Play equals learning*. Oxford: Oxford University Press.
- Graumann, C. F. (1966). Grundzüge der Verhaltensbeobachtung. In E. Meyer (Hrsg.), *Wahrnehmung und Bewußtsein. Handbuch der Psychologie, Bd.11* Göttingen: Verlag für Psychologie.
- Greve, W. & Wentura, D. (1997). *Wissenschaftliche Beobachtung. Eine Einführung*. München: Psychologie Verlags Union.
- Gropengießer, H. & Kattmann, U. (2006). *Fachdidaktik Biologie. Begründet von D. Eschenhagen, U. Kattmann und D. Rodi*. Köln: Aulis.
- Grochla, N. (2008). Qualität und Bildung. Eine Analyse des wissenschaftlichen Diskurses in der Frühpädagogik. In A. Stroß, R. Rehn & E. Spiegel (Hrsg.), *Vechtaer Beiträge zur Frühpädagogik. Bd.2*. Berlin: LIT Verlag.
- Grube, C., Möller, A., Mayer, J. (2007). Dimensionen eines Kompetenzstrukturmodells zum Experimentieren. In Bayrhuber et al. (Hrsg.), *Ausbildung und Professionalisierung von Lehrkräften. Abstracts zur Internationalen Tagung der Fachgruppe Biologiedidaktik im VBIO* (S.31–34). Kassel: Universität Kassel.
- Hammann, M. (2007). Das Scientific Discovery as Dual Search-Modell. In D. Krüger & H. Vogt (Hrsg.), *Theorien in der biologiedidaktischen Forschung (187–196)*. Berlin: Springer
- Harms, U., Mayer, J., Hammann, M., Bayrhuber, H. & Kattmann, U. (2004). Kerncurriculum und Standards für den Biologieunterricht in der gymnasialen Oberstufe. In H.-E. Tenorth (Hrsg.), *Kerncurriculum Oberstufe II. Biologie, Chemie, Physik, Geschichte, Politik* (22–84). Weinheim: Beltz.
- Sachser, N. (2006). Neugier, Spiel und Lernen: Verhaltensbiologische Anmerkungen zur Kindheit. In U. Herrmann (Hrsg.), *Neurodidaktik. Grundlagen und Vorschläge für gehirngerechtes Lehren und Lernen* (19–30). Weinheim: Beltz.

- Kerschensteiner, G. (1914). Wesen und Wert des naturwissenschaftlichen Unterrichts. In K.T. Fischer (Hrsg.), *Die Schule der Naturwissenschaften*. Leipzig: Teubner.
- Krüger, H.-H. (2006): Forschungsmethoden in der Kindheitsforschung. *Diskurs Kindheits- und Jugendforschung* 1, 93–116.
- Krüger, H.-H. , Grunert, C. (2001): Biographische Interviews mit Kindern. In I. Behnken & J. Zinnecker (Hrsg.), *Kinder. Kindheit. Lebensgeschichte. Ein Handbuch* (S. 129–143). Seelze: Kallmeyersche Verlagsbuchhandlung.
- Lück, G. (2003). *Handbuch der naturwissenschaftlichen Bildung. Theorie und Praxis für die Arbeit in Kindertageseinrichtungen*. Freiburg: Herder.
- Martin, E. & Wawrinowski, U. (2006). *Beobachtungslehre. Theorie und Praxis reflektierter Beobachtung und Beurteilung*. Weinheim: Juventa.
- Mc Linacre, M. (1994): Sample Size and Item Calibration Stability. In J.M. Linacre, *Rasch Measurement Transactions*, 7:4, 328.
- Mc Linacre, J. M. (2009). Winsteps (Version 3.68) [Computer Software]. Beaverton, Oregon: Winsteps.com.
- Mc Linacre, J. M. (2010). *WINSTEPS Multiple-Choice, Rating Scale and Partial Credit Rasch Analysis*. [online]. Verfügbar unter: <http://www.winsteps.com/winsteps.htm> [28.06.2010].
- Mollenhauer, K. & Rittelmeyer, C. (1977). *Methoden der Erziehungswissenschaft*. München: Juventa.
- Naguib, M. (2006). *Methoden der Verhaltensbiologie*. Berlin: Springer.
- Neumann, K., Kauertz, A., Lau, A., Notarp, H. & Fischer, H. E. (2007). Die Modellierung physikalischer Kompetenz und ihrer Entwicklung. *Zeitschrift für Didaktik der Naturwissenschaften*. 13, 103 – 123.
- Oerter, R. (1996). Spielendes Lernen, gibt es das? Vom Widerspruch zwischen Spiel und Lernen und wie man ihn überwinden kann. *Praxis Schule* 5–10, 4, 6–9.
- Österreicher, H. (1999). Wahrnehmung mit allen Sinnen. In H. Rieder-Aigner (Hrsg.). *Handbuch Kindertageseinrichtungen. Ergänzbare Sammlung*. Regensburg: Walhalla [online]. Verfügbar unter: <http://www.kindergartenpaedagogik.de/684.html> [17.12.2009].
- Oguz, A. & Yurumezoglu, K. (2007). *The primacy of observation in inquiry-based science teaching*. Paper presented at the International Workshop: Science Education in School (Bucharest, Romania, Oct 11–14, 2007).
- Simon, S., Naylor, S., Keogh, B., Maloney, J. & Downing, B. (2008). Puppets promoting engagement and talk in science. *International Journal of science education*, 30 (9), 1229–1248.
- Schuster, D. & Leland, C. (2008). Considering Context. *Science*, 45, 22–24.
- Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der Bundesrepublik Deutschland (2005). *Bildungsstandards der Kultusministerkonferenz*. München, Neuwied: Wolters Kluwer Deutschland GmbH.
- Sjøberg, S. (2007). *Naturfag som allmenndannelse. En kritisk fagdidaktikk*. Oslo: Gyldendal.
- Sodian, B. & Thoermer, C. (2006). Theory of Mind. In W. Schneider & B. Sodian (Hrsg.), *Kognitive Entwicklung. Enzyklopädie der Psychologie. Entwicklungspsychologie Bd. 2* (S.495–579). Göttingen: Hogrefe.
- SPSS Inc. (2009). SPSS for Windows (Version 18.0) [Computer Program]. Chicago: SPSS.
- Sumaski, W.(1977). Systematische Beobachtung – Grundlagen einer empirischen Methode. In G. Leder et al., (Hrsg.), *Hildesheimer Beiträge zu den Erziehungs- und Sozialwissenschaften. Bd. 5*. Hildesheim: Georg Olms Verlag.
- Tudge, J. (1990). Vygotsky, the zone of proximal development, and peer collaboration: Implications for classroom practice. In L.C. Moll (Hrsg.), *Vygotsky and education: Instructional implications and applications of sociohistorical psychology* (S. 155–172). Cambridge: Cambridge University Press.

- Weinert, F.E. (2001). Vergleichende Leistungsmessung in Schulen – eine umstrittene Selbstverständlichkeit. In F. E. Weinert (Hrsg.), *Leistungsmessung in Schulen* (S. 17–31). Weinheim: Beltz.
- Wirtz, M. & Caspar, F. (2002). *Beurteilerübereinstimmung und Beurteilerreliabilität. Methoden zur Bestimmung und Verbesserung der Zuverlässigkeit von Einschätzungen mittels Kategoriensystemen und Ratingskalen*. Göttingen: Hofgrefe.
- Wu, M. L., Adams, R. J., Wilson, M. R., & Haldane, S. (2007). ConQuest (Version 2.0) [Computer Software]. Camberwell, Australia: ACER.

KONTAKT

Lucia Kohlhauf
 LMU München
 Institut für Didaktik der Biologie
 Winzererstrasse 45/II
 80797 München
lucia.reichart@lrz.uni-muenchen.de

AUTORENINFORMATION

Lucia Kohlhauf
 arbeitet seit 2009 als wissenschaftliche Mitarbeiterin am Institut für Didaktik der Biologie an der Ludwig-Maximilians-Universität München an ihrer Doktorarbeit. Ihr Forschungsschwerpunkt liegt dabei auf der Förderung von Beobachtungskompetenz als wissenschaftlicher Erkenntnismethode im Vorschulbereich.

Dr. Ulrike Rutke

ist Grundschulpädagogin und arbeitet seit 2003 als Wissenschaftliche Mitarbeiterin am Institut für Didaktik der Biologie an der Ludwig-Maximilians-Universität München. Ihr Forschungsschwerpunkt liegt auf Naturwissenschaftlicher Bildung im Grundschulbereich.

Prof. Dr. Birgit Neuhaus

ist Professorin für Didaktik der Biologie an der Ludwig-Maximilians-Universität München. Ihre Forschungsschwerpunkte liegen im Bereich der fachspezifischen Unterrichtsqualitätsforschung, der fachspezifischen Analysen zur Professionalisierung von Lehrenden und der video-basierten Lehr-Lernforschung vom Vorschulalter bis zur universitären Lehrerausbildung.