

**Förderung der Koordination von Theorie und Evidenz  
bei Kindern im Elementarbereich**

Inauguraldissertation zur Erlangung des Grades eines Doktors  
der Philosophie  
am Fachbereich Erziehungswissenschaften  
an der Johann Wolfgang-Goethe-Universität zu Frankfurt am Main

vorgelegt von  
Simone Stephan-Gramberg

November 2015

-Korrigierte Version-

Erstgutachterin: Prof. Dr. Ilonca Hardy  
Zweitgutachterin: Prof. Dr. Diemut Kucharz

## Danksagung

Einige Menschen haben mir während meiner Promotionszeit mit Rat und Tat zur Seite gestanden – ihnen allen gilt mein Dank. Besonders hervorheben möchte ich die Unterstützung meiner beiden Betreuerinnen, Professor Dr. Ilonca Hardy sowie Professor Dr. Diemut Kucharz, die meine Arbeit durch ihre fachliche Begleitung erst ermöglicht haben. Durch Hinweise und Ideen haben sie mich immer wieder zur kritischen Hinterfragung angeregt. Außerdem geht mein Dank an meine beiden Projektkollegen in Münster, Christin Robisch und Steffen Tröbst. Sie hatten stets ein offenes Ohr für fachliche Fragen; die Zusammenarbeit mit beiden war immer überaus konstruktiv. Zudem möchte ich mich bei den studentischen Mitarbeitern an der Johann Wolfgang-Goethe-Universität, vor allem bei Kai Löhner und Theresa Walther, für ihre Unterstützung und ihr Engagement bedanken. Nicht zuletzt bin ich meiner Familie, vor allem meinem Mann und meinen beiden Töchtern, für ihre Geduld und Unterstützung zu Dank verpflichtet. Während der ganzen Zeit haben sie mich immer wieder ermutigt und in vielfältiger Weise unterstützt.

Bad Homburg im November 2015

## Zusammenfassung

Die vorliegende Arbeit, die im Rahmen des zwischen 2011 und 2013 durchgeführten Forschungsprojektes „Förderung von Modellbildungs- und Falsifikationsprozessen im Elementar- und Primarbereich“ entstanden ist, untersuchte auf Grundlage neuerer entwicklungspsychologischer Forschungsbefunde die Möglichkeiten der Förderung im naturwissenschaftlichen Denken bei Kindern im Elementarbereich. Die vorliegende Fassung enthält Korrekturen in den Tabellen 33, 34, 39 und 40 in der Trainingsstudie.

Nach der theoretischen Einordnung des Themas und der Darstellung der Forschungslage wurden im empirischen Teil in einem ersten Schritt die Kompetenzen beim Schlussfolgern im Themengebiet *Elastizität und Plastizität* und beim Wissenschaftsverständnis von Kindern im Alter von vier bis zehn Jahren, eingeteilt in vier Altersstufen, ermittelt; weiterhin wurden die Verknüpfungen beider Kompetenzbereiche untersucht. Als Instrumente dienten ein bereits erprobter Schlussfolgerungstest sowie ein neu entwickelter Test zur Kompetenzmessung von Wissenschaftsverständnis. In der Grundschule wurden die Tests jeweils als Gruppentests und im Kindergarten als Einzeltests durchgeführt. Die Stichprobe umfasste 142 Kinder, 82 Kinder aus dem Primarbereich und 60 Kinder aus dem Elementarbereich. Beim Schlussfolgern zeigte sich, dass es für Kinder aller einbezogenen Altersgruppen deutlich leichter war, mit Ereignissen umzugehen, die eine Vermutung bestätigen, als mit solchen, die eine Vermutung widerlegen. Zudem stellte sich heraus, dass der Umgang mit Ereignissen, die im Hinblick auf eine Vermutung irrelevant sind, noch schwieriger war. Mit zunehmendem Alter war eine Kompetenzsteigerung erkennbar. Die Analyse der Tests ergab außerdem einen Zusammenhang zwischen Wissenschaftsverständnis und Schlussfolgern sowie einen deutlichen Einfluss exekutiver Funktionen. Im zweiten Schritt wurden zwei ausgewählte Trainingsmaßnahmen zur Förderung der Koordination von Theorie und Evidenz bei Kindern im Alter von fünf bis sechs Jahren auf ihre Wirksamkeit hin geprüft, und zwar einerseits durch die Unterstützung mit adaptivem Nachfragen bei fehlerhaften Antworten sowie andererseits durch eine intensive Förderung mit Modellierung. Die an einer Stichprobe von 63 Kindern durchgeführte Trainingsstudie war als Prä-Post-Studie angelegt und umfasste die Überprüfung des erworbenen Wissens. Die Studie ergab, dass die intensiv geförderten Kinder deutlich höhere Kompetenzen erworben hatten als die durch adaptives Nachfragen unterstützten Probanden. Außerdem wurde ein Transfer-Test im Inhaltsgebiet *Schwimmen und Sinken* durchgeführt; hierbei wurden beide Trainingsgruppen in gleicher Weise mit adaptivem Nachfragen unterstützt. Dabei zeigten Kinder beider Trainingsgruppen deutlich höhere Kompetenzen beim Schlussfolgern als im Post-Test, dennoch zeigten Kinder

mit vorheriger intensiver Förderung durch Modellierung wiederum höhere Kompetenzen im Transfer-Test als Kinder aus der Gruppe mit adaptiver Unterstützung. Abschließend wurde noch ein Argumentationstest durchgeführt, bei dem Kinder aller drei Experimentalgruppen (Trainingsgruppe 1, Trainingsgruppe 2, Kontrollgruppe) übergreifende Kompetenzen beim Schlussfolgern zeigen konnten. In diesem Test zeigte sich zwischen den drei Gruppen kein Unterschied im Hinblick auf angemessene Antworten beim Schlussfolgern.

## Inhaltsverzeichnis

Danksagung .....	2
Zusammenfassung .....	3
1 Einleitung .....	9
2 Frühkindliche Bildung in Deutschland.....	13
2.1 Frühkindliche naturwissenschaftliche Bildung in Deutschland .....	13
2.2 Drei Dimensionen von Qualität im Elementarbereich .....	15
2.3 Studien zur Prozessqualität.....	17
2.4 Zusammenfassung und Fazit .....	21
3 Förderung von Lernprozessen .....	23
3.1 Die Rolle des sozialen Konstruktivismus im Rahmen von Lehr- und Lernprozessen 23	
3.2 Forschungsstand: Scaffolding sowie verwandte Konzepte zur Lerner-Unterstützung 24	
3.3 Scientific Literacy im Elementarbereich in Deutschland .....	34
3.4 Exkurs: US-amerikanische Förderprogramme im letzten Jahrzehnt.....	37
3.5 Zusammenfassung und Fazit .....	40
4 Die Entwicklung der Fähigkeit zur Theorie-Evidenz-Koordination .....	41
4.1 Verschiedene Kompetenzniveaus.....	41
4.2 Koordination von Theorie und Evidenz .....	46
4.3 Bereichsübergreifende Fähigkeiten .....	50
4.4 Empirische Studien zur Koordination von Theorie und Evidenz.....	53
4.5 Zusammenfassung und Fazit .....	64
5 Die Schlussfolgerungsfähigkeit in der kognitiven Entwicklung .....	65

---

5.1	Theorie der Mentalen Modelle beim schlussfolgernden Denken.....	67
5.2	Empirische Studien zur Entwicklung schlussfolgernden Denkens .....	70
5.3	Exekutive Funktionen und Schlussfolgern.....	75
5.4	Schlussfolgern und Argumentieren .....	80
5.5	Zusammenfassung und Fazit .....	81
6	Forschungslücken und Ziele.....	82
7	Querschnittstudie.....	85
7.1	Vorstudie .....	85
7.1.1	Fragestellungen und Hypothesen.....	85
7.1.2	Design.....	86
7.1.3	Analysestrategien bei der Auswertung.....	86
7.1.4	Methoden.....	88
7.1.5	Stichprobe.....	100
7.1.6	Ergebnisse.....	100
7.2	Hauptstudie.....	103
7.2.1	Fragestellungen und Hypothesen.....	104
7.2.2	Methoden.....	104
7.2.3	Stichprobe.....	106
7.2.4	Design.....	106
7.2.5	Ergebnisse zur ersten Forschungsfrage: Fähigkeiten beim Schlussfolgern.....	107
7.2.6	Ergebnisse zur zweiten Forschungsfrage: Kompetenzen beim Wissenschaftsverständnis .....	112
7.2.7	Ergebnisse zur dritten Forschungsfrage: Zusammenhang von Schlussfolgern und Wissenschaftsverständnis .....	118
7.3	Diskussion .....	119

---

8	Trainingsstudie .....	122
8.1	Vorstudie .....	125
8.1.1	Fragestellung und Hypothese .....	125
8.1.2	Methoden .....	126
8.1.3	Stichprobe .....	133
8.1.4	Design .....	133
8.1.5	Vorgehen bei der Durchführung .....	134
8.1.6	Analysestrategien bei der Auswertung .....	135
8.1.7	Ergebnisse .....	135
8.2	Hauptstudie .....	138
8.2.1	Fragestellungen und Hypothesen .....	138
8.2.2	Methoden .....	141
8.2.3	Stichprobe .....	141
8.2.4	Design .....	141
8.2.5	Analysestrategien bei der Auswertung .....	142
8.3	Ergebnisse der Hauptstudie .....	143
8.3.1	Ergebnisse zur ersten Forschungsfrage: Förderung der Schlussfolgerungsfähigkeit .....	145
8.3.2	Ergebnisse zur zweiten Forschungsfrage: Unterschiede zwischen den Trainingsmaßnahmen .....	153
8.3.3	Ergebnisse zur dritten Forschungsfrage: Langfristige Trainingseffekte .....	157
8.3.4	Ergebnisse zur vierten Forschungsfrage: Vergleich der Effekte von Trainingsdomäne und Transfer-Test .....	164
8.3.5	Ergebnisse zur fünften Forschungsfrage: Argumentationstest .....	168
8.3.6	Ergebnisse zur sechsten Forschungsfrage: Schlussfolgern, Argumentieren .....	170

8.3.7	Zusammenfassung und Fazit .....	171
8.4	Diskussion .....	173
9	Gesamtdiskussion .....	176
10	Literaturverzeichnis .....	185
11	Tabellen- und Abbildungsverzeichnis .....	207
12	Anhang .....	212



## 1 Einleitung

Die vorliegende Arbeit ist im Rahmen des Forschungsprojektes „Förderung von Modellbildungs- und Falsifikationsprozessen im Elementar- und Primarbereich“<sup>1</sup> entstanden. Die zentrale Fragestellung der Untersuchung zielt auf die Möglichkeiten der Förderung von Kindern im Elementarbereich im naturwissenschaftlichen Denken.

Aus unterschiedlichen Gründen besteht aktuell in Deutschland ein besonderer Bedarf an umfassender früher Förderung von Kindern im naturwissenschaftlichen Bereich. In einer mehr und mehr von Naturwissenschaft und Technik geprägten Umwelt ist eine aktive Teilnahme an wichtigen gesellschaftlichen Diskursen nur möglich, wenn eine naturwissenschaftliche Grundbildung vorhanden ist (Prenzel & Artelt, 2007). Die Förderung naturwissenschaftlicher Kompetenzen bei der nachwachsenden Generation stellt zudem einen wichtigen wirtschaftlichen Faktor dar, denn in Deutschland besteht bereits heute ein Mangel an Fachkräften dieser Ausbildungsrichtung (Bundesbildungsministerium, 2012). Verschiedene Schulleistungsstudien der Organisation für wirtschaftliche Zusammenarbeit und Entwicklung (OECD) zeigen demgegenüber, dass deutsche Schüler in den letzten Jahren im internationalen Vergleich mittelmäßige Leistungen in den Naturwissenschaften erbrachten (z.B. PISA 2000). Die Ergebnisse dieser Schulleistungsuntersuchungen führten zu einer breiten gesellschaftlichen Debatte und dazu, dass Wege gesucht wurden, um die Qualität der Schulbildung in Deutschland zu verbessern (BMBF, 2004). Auch Ergebnisse der TIMSS Studie (Trends in International Mathematics and Science Study), einer international vergleichenden Leistungsstudie, zeigten große Wissensdefizite bei deutschen Schülern im Alter von etwa 15 Jahren im naturwissenschaftlich-technischen Bereich (Bos et al., 2008).

Auf diesen Bedarf reagierte die deutsche Bildungspolitik und regte dazu an, naturwissenschaftliche Grundbildung bereits bei Kindern ab etwa vier Jahren zu fördern. Die wachsende Bedeutung dieses Aspekts in der Elementarbildung zeigt sich unter anderem an den Bildungsplänen der einzelnen Bundesländer (Kultusministerkonferenz, 2004; 2009) sowie an der steigenden Zahl von Projekten zur Förderung junger Kinder im Bereich der Naturwissenschaften.

Studien aus der entwicklungspsychologischen Forschung (Koerber, Sodian,

---

<sup>1</sup> Das Forschungsprojekt mit den Nummern MO 942/5-1 und HA 3205/5-1 wurde von der Deutschen Forschungsgemeinschaft gefördert und in der Zeit von 2011 bis 2013 an den Standorten in Frankfurt am Main und Münster durchgeführt.

Thoermer & Nett, 2005) zeigten, dass bereits junge Kinder über Basisfähigkeiten für wissenschaftliches Denken verfügen (z.B. Zimmermann, 2007) und naive Vorstellungen in naturwissenschaftlichen Bereichen aufbauen können. So können zum Beispiel schon Kindergartenkinder gut beobachten, Details erkennen und benennen. Die vorhandenen Fähigkeiten gilt es auszubauen und durch Anleitung für naturwissenschaftliche Fragestellungen zu nutzen. In der Schule kann an die bereits vorhandenen Kompetenzen angeknüpft werden und diese weiter ausdifferenzieren (Steffensky, 2008).

Das Hauptanliegen der naturwissenschaftlichen Bildung im Elementarbereich besteht darin, Kinder am Beispiel verschiedener Themen anzuleiten, eine Verbindung der Komponenten wissenschaftlichen Denkens (Fragestellung-Beobachtung-Evidenz-Schlussfolgern) herzustellen und anzuwenden. Die Verknüpfung der genannten Komponenten hat einen zentralen Stellenwert im Rahmen der naturwissenschaftlichen Bildung. Insbesondere geht es hier um Kompetenzen im Hinblick auf das Stellen und Beantworten von Fragen (Möller, 2009; Steffensky, 2008), denn auf diese Weise können Kinder ihre Vorstellungen zu einem spezifischen Thema äußern. Dies gelingt, indem pädagogische Fachkräfte Lernumgebungen vorbereiten und mit passender Gesprächsführung Kindern den Zusammenhang der einzelnen Komponenten vermitteln. Besonders eignen sich dabei Themen aus dem Lebensumfeld der Kinder. Die beiden hier vorgestellten Studien beschäftigen sich exemplarisch mit dem Bereich *Elastizität und Plastizität* sowie *Schwimmen und Sinken*, da diese Inhaltsgebiete im kindlichen Alltag beim Spiel mit Bällen bzw. beim Spiel im Wasser bedeutsam und deswegen besonders geeignet für die Förderung frühkindlicher Lernprozesse sind.

### *Ziele*

Empirische Studien zeigen, dass die Fähigkeit zur Koordination von Theorie und Evidenz, auch als *Wissenschaftsverständnis* bezeichnet (Sodian, Jonen, Thoermer & Kircher, 2006), ein zentraler Bestandteil naturwissenschaftlicher Bildung ist (z.B. Klahr & Dunbar, 1988; Klahr, 2000). Hier knüpft die vorliegende Untersuchung an und zielt darauf ab, die Fähigkeit zur Koordination von Theorie und Evidenz bei Kindern im Elementarbereich besser verstehen und geeignete Fördermaßnahmen identifizieren zu können. Dazu werden zunächst in einem theoretischen Teil der Hintergrund sowie die Forschungslage zum Themenfeld skizziert und anschließend zwei im Rahmen des eingangs genannten Projektes durchgeführte empirische Studien (Querschnittstudie/Trainingsstudie) vorgestellt und im Zusammenhang mit den Fragestellungen ausgewertet.

Die Querschnittstudie untersuchte, inwieweit Kinder zwischen vier und zehn Jahren im Zusammenhang mit einer Fragestellung selbständig beobachten und den Beleg ihrer Beobachtungen als Grundlage für eine eigene Schlussfolgerung bzw. zum Aufzeigen einer Wenn-Dann-Beziehung nutzen können. Ziel war es, empirisch zu belegen, inwieweit die adäquate Verknüpfung der vier Komponenten Fragestellung-Beobachtung-Evidenz-Schlussfolgern in den unterschiedlichen Altersgruppen (vierjährig, fünf- bis sechsjährig, sieben- bis achtjährig, neun- bis zehnjährig) gelingt. Neben dem Wissenschaftsverständnis haben exekutive Funktionen, wie vor allem Arbeitsgedächtnis und Inhibitionsfähigkeit, Einfluss auf die Fähigkeit zum Schlussfolgern. Durch unterschiedliche Testverfahren werden Wissenschaftsverständnis und exekutive Funktionen in der Querschnittstudie einzeln erhoben, um den Einfluss dieser Faktoren auf das Schlussfolgern genauer zu fassen. Dabei wird der Frage nachgegangen, ob sich die Einflussnahme von Wissenschaftsverständnis bzw. exekutiven Funktionen mit wachsenden kognitiven Fähigkeiten der Kinder steigert bzw. mindert sowie welches Maß der exekutiven Funktionen, Arbeitsgedächtnis, phonologische Bewusstheit oder Inhibition, den stärksten Einfluss auf einen Schlussfolgerungsprozess nimmt.

Zwar beschäftigen sich einige wissenschaftliche Studien mit Kompetenzen bei der Verknüpfung von Theorie und Evidenz (Koerber et al., 2005; Sodian, Jonen, Toermer & Kircher, 2006), aber welche Maßnahmen sich zur Förderung dieser Fähigkeiten im Elementarbereich eignen, ist bislang nicht umfassend untersucht. Dieser Fragestellung widmete sich die Trainingsstudie. Sie prüfte zwei unterschiedliche Trainingsmaßnahmen zur Förderung der Koordination von Theorie und Evidenz bei Kindern im Alter von fünf bis sechs Jahren auf ihre Wirksamkeit hin. Die erste Maßnahme (Trainingsgruppe 1, Scaffolding durch adaptive Unterstützung) sollte ein Kind durch direktes Nachfragen zum Denken anregen. Bei der zweiten Maßnahme (Trainingsgruppe 2, Unterstützung durch Modellierung und adaptives Nachfragen) nahm ein Kind eine weniger aktive Rolle ein. Hier erklärte der Erwachsene einem Kind seinen eigenen Denkprozess in Einzelschritten. Die Studie folgte einem Prä-Post-Design, um die Entwicklung des Antwortverhaltens beider Trainingsgruppen vor und nach dem Training vergleichen zu können. Aus diesem Grund umfasste die Studie zwei Trainingsgruppen sowie eine Kontrollgruppe. Die Studie war als Laborstudie mit möglichst optimalen Bedingungen angelegt, um die Wirksamkeit der beiden Trainingsmaßnahmen im Unterschied zur Kontrollgruppe zu prüfen. Insgesamt zielte diese Studie darauf ab, empirische Daten für die Wirksamkeit unterschiedlicher Unterstützungsmaßnahmen in der fokussierten Altersgruppe zu erheben.

### *Gliederung*

Der theoretische Teil, Kapitel 2, beschäftigt sich zunächst mit grundlegenden und bildungspolitisch aktuellen Ansätzen für den Elementarbereich. Im Zusammenhang damit werden Ziele frühkindlicher Bildung sowie Möglichkeiten der Qualitätsmessung im Elementarbereich besprochen. Kapitel 3 stellt für die Studie relevante Theorien und dazugehörige Formen der Förderung von Lernprozessen vor und zeigt Möglichkeiten der Gestaltung naturwissenschaftlicher Grundbildung in frühkindlichen Bildungsprozessen auf. In diesem Zusammenhang wird die Einbettung des naturwissenschaftlichen Lernens in Bildungsplänen beschrieben sowie unterschiedliche Bestrebungen in Deutschland, Großbritannien und den USA, naturwissenschaftliche Kompetenzen bei Kindern zu fördern. Kapitel 4 erläutert Modelle der Verknüpfung von Theorie und Evidenz; zudem werden dort Theorien über die Entwicklung der Fähigkeit zur Koordination beider Komponenten (Wissenschaftsverständnis) beschrieben. Das Vermögen, systematisch Strategien zur Prüfung von Hypothesen einzusetzen, gilt dabei als übergreifende Kompetenz (Sodian, Joneke, Thoermer & Kircher, 2006). Zur Erläuterung des Vorgehens beim wissenschaftlichen Arbeiten werden zwei Modelle dargestellt. Das erste Modell (Klahr & Dunbar, 1988; Klahr, 2000) konstruiert drei Schritte der Hypothesenprüfung (Suche im Hypothesenraum, Hypothesen testen, Evaluation); das zweite Modell stellt das Wissenschaftsverständnis als übergeordnete Kompetenz wissenschaftlichen Denkens ins Zentrum. Beide Modelle veranschaulichen wissenschaftliche Arbeitsweisen mit dem Ziel der Hypothesenprüfung. Theorien zum schlussfolgernden Denken werden in Kapitel 5 näher betrachtet, zunächst die Theorie der mentalen Modelle. Im Anschluss daran wird die Entwicklung schlussfolgernden Denkens bei Kindern bis hin zu Jugendlichen und Erwachsenen erläutert. Dabei stellen die verschiedenen Interpretationsniveaus (Barrouillet, Gauffroy & Lecas, 2008) die Ausprägung der menschlichen Schlussfolgerungsfähigkeit auf unterschiedlichen Kompetenzniveaus dar. Empirischen Forschungen zufolge nehmen exekutive Funktionen, Arbeitsgedächtnis und Inhibitionsfähigkeit, bedeutsamen Einfluss auf das Schlussfolgern (z.B. Markovits, 2006, Röthlisberger, Neuenchwander, Michel & Roebbers, 2010).

Kapitel 6 zeigt Forschungslücken auf, die die Zielsetzung der vorliegenden Arbeit bestimmen und die Grundlagen für den empirischen Teil bilden. In Kapitel 7 und 8 werden die beiden empirischen Studien detailliert vorgestellt.

## 2 Frühkindliche Bildung in Deutschland

Einleitend soll in diesem Kapitel der Hintergrund abgesteckt werden, vor dem sich frühkindliche naturwissenschaftliche Bildung in Deutschland abspielt, wobei sowohl die aktuelle Diskussion in der Forschung als auch der Bezug zum Bildungsplan hergestellt wird. Anschließend werden Qualitätsdimensionen im Elementarbereich sowie Studien zur Prozessqualität erläutert.

### 2.1 *Frühkindliche naturwissenschaftliche Bildung in Deutschland*

Die aktuelle Bildungsdiskussion in Deutschland ist von den Ergebnissen der PISA-Studie (2001) geprägt. Im internationalen Vergleich zeigten 15-jährige Schüler aus Deutschland in zentralen Kompetenzbereichen wie Lesen, Rechnen oder Naturwissenschaften mittelmäßige Leistungen (Oberhuemer 2004; Artelt & Baumert, 2001; Rost, 2004), während beispielsweise Schüler aus Kanada und Finnland deutlich besser abschnitten. In diesen beiden Ländern wird viel Wert auf frühkindliche Bildung (Stamm 2010) gelegt und insbesondere auf die Förderung von Kindern aus sozial benachteiligten Familien. Dank dieser Förderung scheinen Kinder in die Lage versetzt zu werden, sozial bedingte Defizite auszugleichen (BMSFJ, 2003). In den letzten Jahren wird in Deutschland ebenso verstärkt Wert auf bereichsspezifische Förderung von Kindern gelegt. Neben dem sprachlichen und mathematischen zählt auch der naturwissenschaftliche Bereich dazu. Naturwissenschaftliches Denken und Interesse prägen das Denken von Menschen in unserer Gesellschaft, da es zur Weiterentwicklung und Lösung verschiedener Probleme Grundvoraussetzung ist (OECD 2000; Bybee, 1997).

Die Relevanz von frühkindlichen Bildungsprozessen wird in den Bildungsplänen aller Länder beschrieben (Kucharz, Mackowiak, Ziroli, Kauertz & Mathgreb-Schnierer, 2014). Dazu gehören der mathematische, sprachliche und der naturwissenschaftliche Bildungsbereich. Da es in der vorliegenden Arbeit um frühkindliche Bildungsprozesse naturwissenschaftlichen Bereich geht, wird der Bezug im Bildungsplan dazu hergestellt. Im hessischen Bildungsplan wird erläutert, dass Kindergartenkinder in der Lage sind, „Wenn-Dann Beziehungen in einfachen Kontexten zu verstehen und einfache Schlussfolgerungen zu ziehen“ (Hessisches Sozialministerium & Hessisches Kultusministerium, 2007, S.77). Daher sollte insofern sowohl das Interesse als auch die kindliche Neugier genutzt werden können, um Kindern „spannende Erfahrungen im naturwissenschaftlichen Bereich“ zu eröffnen (ebenda, S.77). Auf diese Weise lernen Kinder einfache naturwissenschaftliche Arbeitsweisen kennen. Zu ihnen gehören das systematische Beobachten, Vergleichen, Beschreiben ebenso wie das Aufstellen und Überprüfen

von Hypothesen. Aktuelle Forschungen zeigen, dass zielgerichtetes Beobachten im Zusammenhang mit einer Fragestellung eine zentrale Grundlage für die Koordination von Theorie und Evidenz darstellen (Steffensky, Lankes, Carstensen & Nölke 2012). Möller & Steffensky (2010) weisen darauf hin, dass zielgerichtetes Beobachten möglichst früh, bereits im Elementarbereich, gestärkt werden soll. So wird es auch im Hessischen Bildungsplan gefordert (Hessisches Sozialministerium & Hessisches Kultusministerium, 2007, S.78). Als geeignete Möglichkeit der Umsetzung wird das Experimentieren bereits für Vorschulkinder angesehen (Steffensky et al., 2012). Daher wurde es auch in die Bildungspläne aufgenommen. Kinder stellen dabei zunächst Vermutungen (Theorien) auf und überprüfen diese im Anschluss. Kinder können so früh mit naturwissenschaftlichen Themen vertraut werden und Interesse aufbauen (Hessisches Sozialministerium & Hessisches Kultusministerium, 2007, S.78). Sie lernen, bewusst naturwissenschaftliche Vorgänge im Zusammenhang mit einer Vermutung zu betrachten, die Evidenz im Hinblick darauf zu bewerten und eine adäquate Schlussfolgerung zu ziehen. (Wissenschaftliches Denken als Koordination von Theorie und Evidenz wird ausführlich in Kapitel 4 behandelt).

Frühkindliche naturwissenschaftliche Bildung im Elementarbereich bezieht sich auf die Beschäftigung mit Fragen aus der Lebenswelt der Kinder. Kauertz unterscheidet dabei zwei Arten von Fragen. Die erste Frage bezieht sich meist auf Abläufe von Prozessen (z.B. „Wie funktioniert das?“), während die zweite Frage häufig auf die Klärung von wahrgenommenen Phänomenen aus der Natur ausgerichtet ist (z.B. „Wieso gibt es Frost?“, „Warum wird es Nacht?“) (Kauertz, 2012, S. 99). Naturwissenschaftliche Bildung im Elementarbereich wird zum Beispiel erreicht, indem ein Experiment durchgeführt und Beobachtungen sowie Erkenntnisse formuliert werden. Dazu werden meist qualitative Aussagen zur Beschreibung (z.B. Vergleich von zwei Größen) genutzt. Messinstrumente sowie andere technische Geräte spielen in der Regel keine bedeutende Rolle. Nach Auffassung von Kauertz (2012) sollten Angebote eine altersentsprechende Offenheit beinhalten, so dass ein Kind sich selbsttätig mit dem jeweiligen Inhalt auseinandersetzen kann. Bei einem Kind mit geringem Vorwissen ist, so Kauertz, zunächst eine Phase eines strukturierten Angebots nötig. Anschließend kann ein Kind im Rahmen eines offenen Angebotes eigene Erfahrungen machen und diese mithilfe von Erklärungen eines Erwachsenen in die eigene Struktur von vorhandenen Wissens-elementen einbauen. Kauertz kritisiert stark strukturierte Angebote, da bei ihnen häufig Phasen des gemeinsamen Austausches von Erzieher und Kind nicht zustande kommen (Kauertz, 2012). Ziel naturwissenschaftlicher Bildung im Elementarbereich (2012) sind „qualitative Beschreibungen von Prozessen und Zuständen“ (Kauertz, 2012, S. 99).

Die Förderung naturwissenschaftlichen Verständnisses von Kindergartenkindern wird von weiteren Forschern (z.B. Leuchter; Saalbach & Hardy, 2011, Steffensky & Lankes 2011) als wichtig angesehen, um sukzessive Interesse und Wissen bei Kinder zu wecken. Steffensky & Lankes (2011) sprechen in diesem Zusammenhang davon, dass Wissen anschlussfähig sein muss an themenspezifische Inhalte, die im Anschluss in der Grundschule vermittelt werden. Das erfahrungsbasierte, möglichst alltagsnahe Wissen kann Kinder befähigen, naturwissenschaftliche Phänomene zu beschreiben und zu begründen. In ihrer Studie untersuchten Steffensky, Lankes, Carstensen & Nölke (2012), welche Formen von Lernangeboten (Experimente versus Alltagsangebote) sich eignen, um naturwissenschaftliche Lernprozesse zu unterstützen.

Während Steffensky & Lankes (2011) in ihren Ausführungen Wert auf die Anschlussfähigkeit bei naturwissenschaftlicher Förderung im Bildungskontext (z.B. zwischen Kindergarten und Grundschule) legten, siedelte Kauertz die Förderung der Fähigkeiten von Kindern ausschließlich im jeweiligen Bildungsbereich (z.B. im Kindergarten) an. Seiner Auffassung nach steht die Verknüpfung der einzelnen Bildungsbereiche nicht im Vordergrund bei der Vermittlung naturwissenschaftlichen Lernens. Während Steffensky et al., (2012) strukturierte Lernangebote bevorzugen, erachtet Kauertz (2012) offene Lernangebote als lernwirksam. Trotz der genannten Unterschiede zwischen Kauertz (2012) und Steffensky et al. (2012), gibt es auch eine Gemeinsamkeit. Sowohl Kauertz (2012) als auch Steffensky et al., (2012) erachten das gemeinsame Suchen nach Begründungen für naturwissenschaftliche Phänomene von Erwachsenen und Kind als relevant für die Entwicklung naturwissenschaftlichen Verständnisses. Eine sinnvolle und lernwirksame Vermittlung von Inhalten ist abhängig von unterschiedlichen Aspekten im Hinblick auf verschiedene Qualitätsdimensionen.

## 2.2 *Drei Dimensionen von Qualität im Elementarbereich*

Im folgenden Absatz werden Qualitätsaspekte von Kindertageseinrichtungen näher betrachtet. In Längsschnittstudien (z.B. Burger, 2010; Roßbach, 2008; Weiland & Yoshikawa, 2013) wurde die Wirkungsweise des Besuchs eines Kindergartens auf Kind-bezogene Merkmale betrachtet. Dabei wurde sowohl die sozio-emotionale als auch die kognitive-leistungsbezogene Entwicklung von Kindern im Zusammenhang mit Qualitätsaspekten eines Kindergartens untersucht. Differenziert wird in der Regel zwischen strukturellen und prozessorientierten Bedingungen (Anders, 2013; Kuger & Kluczniok, 2008). In einer Studie (Tietze, 1998), in einem Überblick (Hardy & Steffensky, 2014) sowie einem Modell zur Qualität von Kindertageseinrichtungen (Stamm, 2010) wurde zusätzlich die *Orientierungsqualität*

in den Blick genommen. Unter dem Begriff werden sowohl die professionelle Einstellung, die Überzeugung als auch die Motivation von Fachkräften gefasst.

Bei der *Strukturqualität* handelt es sich um äußere Faktoren, die sich auf die Organisation innerhalb der Kindertagesstätte beziehen (beispielsweise die Gruppengröße oder der Erzieher-Kind Schlüssel). Die *Strukturqualität* mit ihren äußeren Bedingungen, die *Orientierungsqualität* mit dem Professionswissen sowie der Einstellung der Fachkräfte bilden den Rahmen, in dem Prozesse im Alltag der Fachkräfte und Kinder stattfinden.

Bei der *Prozessqualität* stehen dynamische Aspekte, wie beispielsweise Interaktionen zwischen Fachkraft und Kind bzw. der entwicklungsangemessene Umgang der Fachkraft mit dem Kind im Zentrum eines Kindergartenalltags. In einer Definition von Tietze (2008, S. 18) wird Prozessqualität als „[...] ein prozesshaftes Geschehen [... definiert, bei dem] Anregungen, die die Kinder in den einzelnen Bildungs- und Entwicklungsbereichen erhalten, die Art der Interaktionen zwischen dem pädagogischen Personal und den Kindern, wie auch [zwischen] den Kindern untereinander [bestimmen].“ Der Begriff *Prozessqualität* wird in zwei Dimensionen, in globale und bereichsspezifische Formen eingeteilt. Globale Prozessqualität bezieht sich auf den pädagogischen Umgang der Fachkraft mit dem Kind. Dabei werden sowohl Interaktionen zwischen Fachkraft und Kind im Hinblick auf Lerner-Unterstützung, Qualität der Betreuung als auch die Gestaltung der Umgebung mit validen Messinstrumenten beleuchtet (Kuger & Kluczniok, 2008).

Nach dem Modell von Fröhlich-Gildhoff, Nentwig-Geseman & Pietsch (2011) (siehe Abbildung 1) wird globale und bereichsspezifische Prozessqualität maßgeblich von der Disposition und Performanz der Fachkräfte bestimmt. Fachkräfte zeigen mit ihrer Wahrnehmung und Analyse der Situation, ihrer Motivation sowie ihrem Wissen, ihren Fähigkeiten und Fertigkeiten eine individuelle Performanz in Form von Handlungsplanung und Handlungsbereitschaft, die in pädagogischem Handeln sichtbar wird. Fachkräfte können mittels bestimmter Kriterien (z.B. im Hinblick auf Motivation der Kinder, Ablauf einer pädagogischen Situation) ihr pädagogisches Handeln einschätzen und zudem ihr Handeln selbst reflektieren. Die genannten Aspekte bilden die Grundlage, damit Erzieherinnen ihre zukünftigen pädagogischen Handlungen verbessern können.



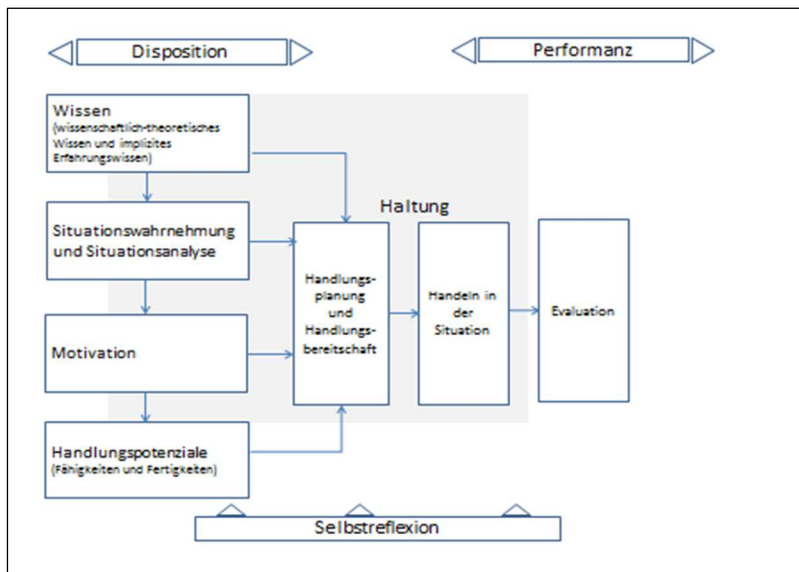


Abbildung 1. Modell Prozessqualität nach Fröhlich-Gildhoff et al (2011, S. 23).

Im Vergleich dazu wird mit der bereichsspezifischen Prozessqualität die Qualität von Trainingsmaßnahmen in spezifischen Inhaltsgebieten untersucht. Beziehungen zwischen globalen und bereichsspezifischen Aspekten der Förderung sind häufig nicht eindeutig voneinander abzugrenzen (Hardy & Steffensky, 2014; Kuger & Kluczniok, 2008). Empirische Untersuchungen (Welzel-Breuer & Zimmermann, 2007) haben gezeigt, dass pädagogische Fachkräfte in ihrer Schulzeit wenige gute Erfahrungen in naturwissenschaftlichen Fächern machen konnten; daher ist ihr Interesse vielfach kaum ausgeprägt (Welzel-Breuer & Zimmermann, 2007). Der Aufbau von Wissen, Interesse sowie Motivation im naturwissenschaftlichen Bereich ist eine wichtige Voraussetzung, damit pädagogische Fachkräfte Lerngelegenheiten mit naturwissenschaftlichen Themen inhaltlich gut und altersgerecht aufbauen können. Nur so kann bereichsspezifische Prozessqualität bei naturwissenschaftlichen Themen erreicht werden. Prozessqualität wirkt im Gegensatz zur Strukturqualität direkt auf die Lernentwicklung eines Kindes (Kuger & Kluczniok, 2008) und spielt damit eine zentrale Rolle bei der Elementarbildung und insbesondere bei der Gestaltung von anregenden Lernumgebungen.

### 2.3 Studien zur Prozessqualität

In den folgenden Absätzen werden vier Studien vorgestellt, die die positiven Auswirkungen von Einrichtungen mit hoher Prozessqualität aufzeigen. Im Zusammenhang mit einer qualitativ guten Interaktion zwischen pädagogischer Fachkraft und Kind stehen positive Effekte in Bezug auf die Lernentwicklung eines Kindes (Fröhlich-Gildhoff et al., 2011). Daher spielt die

Prozessqualität, global bzw. bereichsspezifisch, eine große Rolle für die Lernentwicklung eines Kindes. Bei den Studien handelt es sich um die englische Studie *Effective Pre-School Education* (EPPE-Studie) (Sylva et al., 2004), eine deutsche Studie *Bildungsprozesse, Kompetenzentwicklung, Selektionsentscheidungen* (BiKS-Studie), (von Maurice et al., 2007), um eine aktuelle US-amerikanische Studie (Weiland & Yoshikawa, 2013), sowie um die aktuelle deutsche Studie zum Professionellen Handeln im Elementarbereich (Primel) (Kucharz et al., 2014).

Empirische Befunde der englischen EPPE-Studie zeigten, dass Kindergärten mit einem Konzept, orientiert am Curriculum besonders wirkungsvolle Effekte auf die Kompetenzentwicklung von Kindern haben. In der *EPPE*-Studie (Effective Pre-School Education) einer Längsschnittstudie wurden verschiedene Einflüsse auf die kindliche Entwicklung untersucht (Sylva, Melhuish, Sammons, Siraj-Blatchford und Taggart, 2004). Die Studie wurde in der Zeit von 1997 bis 2003 durchgeführt; in dieser Zeit wurden circa 3000 Kinder aus 141 Kindertageseinrichtungen<sup>2</sup> im Alter von drei bis elf Jahren mehrfach im Hinblick auf ihre kognitive, soziale und emotionale Entwicklung mit validen Testinstrumenten untersucht. Beispielsweise wurde die domänenspezifisch ausgerichtete Ratingskala ECERS-E eingesetzt (Sylva et al., 2006). In der Studie wurden Kinder, die eine Kindertageseinrichtung besuchten, mit solchen verglichen, die von ihren Eltern betreut wurden. Die Forschergruppe sammelte umfangreiche Informationen in Bezug auf den familiären Hintergrund der Kinder sowie im Hinblick auf die Einrichtungen, die die Kinder besuchten. Positive Effekte werden bei Kindern berichtet, die eine qualitativ hochwertige vorschulische Einrichtung besucht hatten. Diese Vorteile von Kindern waren bis zum Ende des zweiten (Melhuish, 2013; Sylva, 2004) bzw. des dritten Schuljahrs (NICHD-ECCRN, 2005) nachweisbar. Kinder aus Einrichtungen mit einem Konzept, das am Curriculum orientiert war, und die gleichermaßen kognitive und soziale Förderziele verfolgten, zeigten die besten Fortschritte in der Längsschnittuntersuchung. Die einflussreichsten Effekte auf die kognitive Entwicklung zeigten sich bei Kindern, die eine maximale Verweildauer bis zum Schuleintritt in einer qualitativ hochwertigen Einrichtung verbracht hatten (Sammons et al., 2004; Melhuish, 2013).

Effekte guter bereichsspezifischer Prozessqualität auf die kognitive Entwicklung von Kindern zeigten sich ebenso in der groß angelegten deutschen *BiKS*-Studie, eine Studie der Universität Bamberg (von Maurice et al., 2007). Dort wurde der Einfluss von globaler und bereichsspezifischer Prozessqualität des Kindergartens im Zusammenhang mit dem familiären

---

<sup>2</sup> Der Begriff Kindertageseinrichtung wird synonym mit Kindergarten verwendet.

Hintergrund auf die Entwicklung kognitiver Leistungen im Rahmen einer Längsschnittstudie erhoben. Insgesamt gab es acht Teilprojekte, an denen insgesamt 550 Kinder aus 97 Kindergärten aus Bayern und Hessen teilnahmen. Untersucht wurden Kinder in „BiKS 3-8“ im Alter von drei bis acht Jahren auf institutioneller Ebene. Das bedeutet, dass jeweils Kindergartengruppen bzw. ganze Schulkassen das Design der Stichprobe ausmachten (von Maurice et al., 2007). Dabei berücksichtigten die Forscher das Konzept der globalen Prozessqualität, das Aspekte der Qualität von Betreuung und Pflege sowie die Qualität der Gruppenführung (nach Klieme et al., 2006) umfasst. Dieses Konzept wurde durch eine bereichsspezifische Prozessqualität erweitert (Kuger & Kluczniok, 2008), dabei wurden spezifische Bereiche wie Sprache sowie mathematische Kompetenzen einzeln erfasst. Kognitive Merkmale wie Sprache, Arbeitsgedächtnis sowie Informationsverarbeitungsgeschwindigkeit wurden durch Einzeltestungen bei jedem Kind erhoben (von Maurice et al., 2007). Die Befunde zeigten, dass eine gute bereichsspezifische Prozessqualität verbunden ist mit der Entwicklung früher mathematischer Kompetenzen. Im Rahmen des Projekts „Families, Children and Childcare“ (Sylva et al., 2012) wurden ebenso positive Effekte guter Qualität einer frühkindlichen Betreuung und der sprachlich-kognitiven Entwicklung bei Kindern im Alter von drei Jahren berichtet.

In einer US-amerikanischen Studie (Weiland & Yoshikawa, 2013) wurde der Einfluss von Effekten unterschiedlicher Treatments auf die bereichsspezifische Entwicklung von Kindern geprüft. Dabei wurde die Wirkungsweise von fachspezifischen Curricula (Mathematik, Sprache) in Kombination mit Fortbildungen von Fachkräften im Hinblick auf die Entwicklung der bereichsspezifischen Kompetenzen (frühe mathematische Kenntnisse, Sprache, exekutive Funktionen und emotionale Unterstützung) untersucht. Weiterhin wurde geprüft, ob spezifische Subgruppen in Bezug auf Familieneinkommen, ethnische Herkunft oder Gender signifikant mehr von den Programmen profitieren als andere. Insgesamt haben an der Studie 2000 Kinder im Alter von vier bis fünf Jahren teilgenommen. In ihren empirischen Befunden konnten Weiland & Yoshikawa (2013) belegen, dass Kinder, die an der Intervention über ein Jahr partizipierten, signifikant bessere Leistungen aufwiesen als Kinder aus der Kontrollgruppe. Dabei zeigten sich mittlere bis starke Effekt im sprachlichen sowie im mathematischen Bereich. Ebenso konnten schwache Effekte bei Maßen der exekutiven Funktionen gefunden werden. Im Hinblick auf die Subgruppen zeigten sich statistisch signifikante Effekte. Sowohl Kinder aus sozial benachteiligten Familien als auch Kinder unterschiedlicher ethnischer Herkunft (lateinamerikanisch, asiatisch) profitierten signifikant mehr von der Teilnahme an den Interventionen als Kinder aus gut situierten Familien ohne Migrationshintergrund.

Als Beispiel zur Ausweitung der Kompetenzen pädagogischer Fachkräfte wird im folgenden Absatz das Projekt *Professionalisierung von Fachkräften im Elementarbereich* (Primel) vorgestellt, das ein Verbundprojekt zwischen verschiedenen Universitäten (z.B. Goethe Universität Frankfurt, Pädagogische Hochschule Weingarten, Leibniz Universität Hannover, Universität Koblenz-Landau, Universität St. Gallen, Universität Schaffhausen) darstellt (Kucharz et al., 2014). Im Fokus stand die Frage, inwiefern sich die unterschiedlichen Formen der Ausbildung (Fachschulausbildung versus Hochschulausbildung) in der Qualität unterschiedlicher Settings im Kindergartenalltag zeigten. Als Erhebungsinstrumente wurden sowohl Fragebögen für die Fachkräfte als auch Videoanalysen verwendet. In den Fragebögen wurden Informationen über die Art der Ausbildung der Fachkräfte sowie über deren Fähigkeiten einer domänenbezogenen Anregung von Lernprozessen eingeholt. Zusätzlich wurden für die Messung der Qualität sowohl Situationen im Freispiel als auch konkret gestaltete domänenspezifisch geprägte Bildungsangebote in den Blick genommen. Diese wurden mit Videokameras aufgenommen und in Intervallen von jeweils 10 Sekunden im Hinblick auf drei Kategorien (Lernprozessgestaltung, Beziehungsgestaltung, Klassenführung) mikroanalytisch beleuchtet (Kucharz et al., 2014). Die gebildeten Skalen waren nicht reliabel, was vermutlich mit dem geringen Stichprobenumfang sowie der

Item-Anzahl zusammenhing (ebenda, S.175); in einem der nächsten Schritte wird eine Validierung erforderlich sein (ebenda, S.175). Einige wichtige Aspekte konnten dennoch herausgestellt werden. In Bezug auf die Qualität der Umsetzung naturwissenschaftlicher Angebote nahmen folgende Merkmale Einfluss: Beziehung der Fachkraft zu den Naturwissenschaften, das eigene Fähigkeitsselbstkonzept, ein grundlegendes Verständnis eines Beispielthemas (ebenda, S. 174). Insgesamt zeigte sich nur ansatzweise ein Einfluss durch die Art der Ausbildung auf die Qualität eines naturwissenschaftlichen Angebotes. Statistisch konnte der Einfluss in Regressionsanalysen nicht nachgewiesen werden (ebenda, S.177). Weiterhin stellte sich heraus, dass die Konzeption eines Kindergartens nur einen geringen Einfluss auf die Qualität von Lernprozessen ausübt. Vorhandenes bzw. nicht vorhandenes Material übt keinen messbaren Einfluss darauf aus, ob ein Lernangebot gut ist oder nicht. Zusammenfassend lässt festhalten, dass 40% aller naturwissenschaftlichen Interaktionen zwischen Kind und Fachkraft als kognitiv aktivierend bezeichnet werden können.

Insgesamt wurde im Primel-Projekt (Kucharz et al., 2014) deutlich, dass Fachkräfte häufig zwei unterschiedliche Vorgehensweisen bei der Gestaltung didaktischer Angebote verwenden. Einige Fachkräfte bereiteten Lerngelegenheiten genauso vor, wie dies in einer didaktischen Handreichung zu einem speziellen Thema beschrieben wird; andere Fachkräfte konzentrierten

sich in ihrer Arbeit verstärkt darauf, gemeinsam mit Kindern über Ursachen bzw. Zusammenhänge von Phänomenen nachzudenken. Nach Steffensky (2012) ist eine Mischung aus beiden Aspekten sinnvoll und lernförderlich.

Die dargestellten empirischen Befunde belegen, dass Einrichtungen mit einer besseren Struktur- und Prozessqualität die Entwicklung eines Kindes positiv beeinflussen können (Tietze, 1998; Colberg & Schrader, 1998; Anders et al., 2012), wohingegen umgekehrt Einrichtungen mit schlechter pädagogischer Qualität negative Auswirkungen auf die Entwicklung eines Kindes haben können (Dohmen, 2010). Demzufolge spielen Struktur- und Prozessqualität eines Kindergartens eine große Rolle im Hinblick darauf, ob ein Kind vom Besuch eines Kindergartens profitieren kann (Roßbach & Anders, 2013). Besonders deutlich nachweisbare Effekte übt die Qualität eines Kindergartens, insbesondere die globale sowie die bereichsspezifische Prozessqualität auf sozial benachteiligte Kinder aus (Bassok et al., 2008; Sylva et al., 2004). Ebenso konnte die BiKS-Studie zeigen, dass eine gute Prozessqualität eines Kindergartens mit positiven Effekten auf die kognitive Entwicklung von Kindern verbunden ist und noch am Ende der ersten Klasse (mit sieben Jahren) nachweisbar ist (Anders et al., 2012). Das Konstrukt der globalen sowie der bereichsspezifischen Prozessqualität im Elementarbereich ist geprägt durch Interaktionsprozesse, die zwischen Fachkraft und Kind ablaufen (z.B. Weiland & Yoshikawa, 2013). Wie gerade aufgezeigt worden ist, spielt das pädagogische Handeln von Fachkräften eine maßgebliche Rolle für die bereichsspezifische Prozessqualität.

#### *2.4 Zusammenfassung und Fazit*

In diesem Kapitel zur frühkindlichen naturwissenschaftlichen Bildung wurde als erstes die Relevanz der Förderung in diesem Bereich erläutert. Naturwissenschaftliches Verständnis im Zusammenhang mit frühkindlicher Bildung wurden bei Steffensky & Lankes (2011), Steffensky et al., (2012) und Kauertz (2012) in den Blick genommen (Kapitel 2.1). Weiterhin wurden Bezüge zu deutschen Bildungsplänen, insbesondere zum hessischen Bildungsplan vorgenommen, um die staatliche Verankerung sowie den Bildungsanspruch im Elementarbereich zu verdeutlichen. Kinder sollen bereits im Elementarbereich am Beispiel von lebensweltbezogenen Themen Wenn-Dann Beziehungen herstellen können. Dadurch können sie zunehmend Zusammenhänge von Ursache-Wirkungsprozessen begreifen (Kapitel 2.2). In Kapitel 2.3 wurden drei Dimensionen von Qualität (Struktur-, Prozess-, sowie Orientierungsqualität) erläutert, da sie zentral sind, um Alltagsprozesse im Elementarbereich näher beschreiben zu können. Schließlich wurden vier Studien zur Prozessqualität (EPPE-Studie, BiKS-Studie, US-amerikanische Studie

sowie die Primel-Studie) dargestellt, um aufzuzeigen, dass besonders die bereichsspezifische Prozessqualität großen Einfluss auf den Lernzuwachs von Kindern ausübt (Kapitel 2.4). Insgesamt zeigte dieses Kapitel Fundierung und Umsetzung von Bildungsbestrebungen im frühkindlichen Bereich zur Förderung naturwissenschaftlichen Lernens.

### 3 Förderung von Lernprozessen

In diesem Kapitel sollen Möglichkeiten und Theorien der Lerner-Unterstützung näher betrachtet werden. Zunächst wird ein kurzer Überblick über den sozialen Konstruktivismus gegeben, da dieser entscheidende Einfluss auf die Förderung von Lernprozessen genommen hat. Im Anschluss daran werden der Begriff des *Scaffolding* (*scaffold* (engl.) = Gerüst) sowie verwandte in der Forschung entwickelte Konzepte erläutert. Diese beziehen sich auf Hilfestellungen in Lehr-Lernsettings, die in einem engen Abstimmungsprozess mit dem Lernenden je nach dessen individuellen Fähigkeiten jeweils zurückgenommen bzw. verstärkt werden. Abschließend wird das Konzept der *Scientific Literacy* vorgestellt sowie einige Beispiele für die Umsetzung dieses Konzeptes in Deutschland und (als Exkurs) den USA.

#### 3.1 Die Rolle des sozialen Konstruktivismus im Rahmen von Lehr- und Lernprozessen

Bei der Betrachtung von Lernprozessen sowie Maßnahmen zur Lerner-Unterstützung spielt der soziale Konstruktivismus eine große Rolle. Der Sozialpsychologe Dewey hat bereits um 1900 die Diagnose des Lernniveaus eines Lernenden als Ausgangspunkt für Maßnahmen zur Lerner-Unterstützung angesehen. Nach Auffassung von Vertretern des sozialen Konstruktivismus stellt ein Lernprozess einen aktiven Prozess dar, an dem die Umwelt, insbesondere die Bezugspersonen, maßgeblich beteiligt sind (Möller, 2009). Ein zentraler Vertreter des sozialen Konstruktivismus ist Vygotsky, da er mit seinen theoretischen Überlegungen entscheidenden Einfluss auf dessen Ausprägung genommen hat. Vygotsky (1978) ging davon aus, dass Kinder eigenes Wissen, geprägt von ihrem sozialen Umfeld, konstruieren. Dabei ist sowohl ihr Lernen als auch ihre Entwicklung eingebunden in ihren individuellen sozialen Kontext (Bodrova & Leong, 2007). Nach Vygotsky spielt der Austausch zwischen Erwachsenem und Kind eine zentrale Rolle. Ein Erwachsener kann durch passgenaue Unterstützung den Lernprozess eines Kindes fördern. Dabei beschrieb er eine Aufgabe, die auf der nächsthöheren Verständnisstufe eines Kindes liegt, als „Zone der nächsten Entwicklung“ (Vygotsky, 1978, S. 84). Seiner Auffassung nach sind Lern- und Entwicklungsprozesse zwei unterschiedliche Abläufe, wobei Lernprozesse eine zentrale Rolle in der Entwicklung eines Kindes spielen und diese prägen (Bodrova & Leong, 2007). Nach der Theorie Vygotskys zeigt ein aktueller Entwicklungsstand die bereits erfolgten mentalen Fortschritte eines Kindes, woran sich der nächste Entwicklungsschritt anschließt (ebenda). Eine Fachkraft reduziert ihre Unterstützung innerhalb eines Lernprozesses, wenn ein Kind eine zuvor unterstützte Tätigkeit alleine ausübt, denn: „What a child is able to

do in collaboration today, he will be able to do independently tomorrow. “ (Vygotsky, 1978, S. 211).

Dieser Gedanke in Vygotskys Theorie (Vygotsky, 1978) hat innerhalb des sozialen Konstruktivismus das Konzept der *Ko-Konstruktion* geprägt. Der Begriff der *Ko-Konstruktion* bezeichnet den Prozess des gemeinsamen Austausches, beispielsweise zwischen Fachkraft und Kind, mit dem Ziel, gemeinsam die Bedeutung eines Sachverhaltes herauszuarbeiten (Carle & Wenzel, 2007; Wannack, 2010). Voraussetzung für eine erfolgreiche Ko-Konstruktion ist, dass beide Kommunikationspartner Intersubjektivität herstellen (Rogoff, 1990; Völkel, 2002). Das bedeutet, dass beide die Perspektive des jeweils anderen verstehen und sich entsprechend aufeinander abstimmen (Völkel, 2002). Nur auf diese Weise kann sich ein Gespräch entwickeln, in dem gemeinsames Wissen konstruiert wird. Bei einem erfolgreichen Ko-Konstruktionsprozess greift der Gesprächspartner die Ideen oder Vorstellungen des Gegenübers auf und regt diesen mit weiterführenden Fragen zum Denken an (Krammer, 2010). Ein Ko-Konstruktionsprozess kann zwischen Erwachsenem und Kind stattfinden, aber auch zwischen Kindern. Er wird dann als symmetrischer (unter Peers) oder komplementärer (zwischen Erwachsenem und Kind) Prozess bezeichnet (Carle & Wenzel, 2007). Frühkindliche Bildungsprozesse sind in der Praxis durch Prozesse der Ko-Konstruktion geprägt, da Erwachsener und Kind Inhalte im gemeinsamen Gespräch aus Situationen heraus bedeutungsvoll werden lassen. Eine Ko-Konstruktion von Wissen bildet die Grundlage für einen positiven Lernverlauf (Krammer, 2010), in dem ein Kind Wissenszuwachs erlangt. Im Prozess sind die Aussagen des Kindes genauso wichtig wie die des Erwachsenen, da der Austausch wechselseitig und im gemeinsamen inhaltlichen Verstehen zu sehen ist (Hogan & Pressley, 1997). Allmählich verändert sich der von äußeren Kommunikationshandlungen mit einem Partner bestimmte Prozess hin zu einer Denkweise, die sich ausschließlich im Menschen abspielt. In diesem Zusammenhang wird häufig von *Internalization* (Vygotsky, 1978) gesprochen. Insgesamt hat der soziale Konstruktivismus die zentrale Position des kooperativen Prozesses zwischen Lehrendem und Lernendem sowie zwischen Gleichaltrigen bei der Konstruktion von Wissensinhalten (Gerstenmaier & Mandl, 1995; Möller, 2001) im Rahmen von Lehr-Lernprozessen unterstrichen.

### 3.2 Forschungsstand: *Scaffolding* sowie verwandte Konzepte zur Lerner-Unterstützung

In diesem Kapitel werden der Begriff *Scaffolding* sowie relevante Konzepte zur Lerner-Unterstützung in Lehr-Lernprozessen dargestellt. Neben Vygotsky (siehe Kapitel 3.1) haben die englischen Kognitionspsychologen Wood, Bruner und Ross (1976) den Begriff



*Scaffolding* geprägt (Pea, 2004; Van de Pol et al., 2010). *Scaffolding* kann als Lerner-Unterstützung übersetzt werden und leitet sich von der in diesem Zusammenhang häufig verwendeten Metapher eines Baugerüsts (Englisch: *scaffold*) her. Die drei Psychologen haben die Interaktion zwischen Kind und Erwachsenen im Rahmen eines tutoriellen Settings beschrieben: Zu Beginn braucht ein Kind am Anfang seines Lernprozesses umfangreiche Unterstützung, um eine Aufgabe zu erledigen. Dabei wird es vom Erwachsenen mit verschiedenen Maßnahmen unterstützt. Nach Vygotsky sollten diese Unterstützungsmaßnahmen im Sinne der „Zone der nächsten Entwicklung“ (Vygotsky, 1978, S. 84) eine Stufe über dem Lernniveau des Lerners angesiedelt sein, sodass der Lernende unmittelbar davon profitieren kann. Auf diese Weise kann die Motivation beim Lernen aufrechterhalten werden und ein effektiver Lernprozess stattfinden (Brophy, 1999). Die Lerner-Unterstützung bzw. *Scaffolding* wird im Rahmen eines dynamischen Prozesses von einem Erwachsenen gestaltet und ist zeitlich begrenzt, da die Maßnahmen entsprechend des Lernprozesses reduziert werden (Kleickmann et al., 2016).

Der Begriff *Scaffolding* ist nicht klar definiert, sondern umfasst ein breites Spektrum unterschiedlicher theoretischer Ansätze (Van de Pol, Volman & Beishuizen, 2010; Pea, 2004). In der vorliegenden Arbeit wird *Scaffolding* im Sinne von Vygotsky (1978) verwendet. Ebenso beinhaltet der Begriff praktische Maßnahmen zur Lerner-Unterstützung, von denen die wichtigsten am Beispiel verschiedener Studien in Tabelle 1 dargestellt werden. Dort ist eine Übersicht über wichtige theoretische und empirische Studien zur Lerner-Unterstützung durch *Scaffolding* sowie verwandte Konzepte zusammengestellt, um die große Bandbreite an Formen von Lerner-Unterstützung aufzuzeigen. Daher finden sich an erster Stelle der Liste drei Studien (De Jong, 2006; Puntambekar & Hübscher, 2005; Van de Pol et al., 2010), die die breite Verwendung des Begriffes *Scaffolding* verdeutlichen und ihn theoretisch fassen. Anschließend werden einzelne Studien aufgeführt, die relevant für die Maßnahmen zur Lerner-Unterstützung der vorliegenden Arbeit sind. Dazu gehören Studien, die Formen zur Unterstützung von Lernprozessen beschreiben wie z.B. die *Modellierung* (Pea, 2004), *Kognitive Strukturierung* (Hardy, 2012), adaptive Lerner-Unterstützung (Decristan et al., 2015; Hardy et al., 2006) oder *Sustained Shared Thinking* (Siraj-Blatchford & Sylva, 2012) untersuchen. Die Studien sind nach ihren Maßnahmen der Unterstützung geordnet, wobei die für die vorliegende Trainingsstudie relevanten Formen von *Scaffolding* als erstes aufgeführt sind.

Tabelle 1

*Studien zur Lerner-Unterstützung*

Maßnahmen	Autoren	Methodik	Probanden
<b>Scaffolding</b>			
	Puntambekar & Hübscher, 2005	theoretisch	Keine
	De Jong, 2006	theoretisch	Keine
	Van de Pol, 2010	theoretisch	Keine
	Simons & Klein, 2007	empirisch	111 Schüler im Alter von 13-14 Jahren
	Decristan et al., 2015	empirisch	1071 Probanden im Alter von 8-9 Jahren
<b>Modellierung</b>			
	Wood et al., 1976	empirisch	30 Kinder im Alter von 3- 5 Jahren
	Rogoff, 1990	empirisch	121 Schüler im Alter von 11-13 Jahren
	Hogan & Pressley, 1997	empirisch	50 Schüler im Alter von 8-10 Jahren
	Murphy & Messer, 2000	empirisch	122 Kinder im Alter von 5-7 Jahren
<b>Fokussieren/Modellierung</b>			
	Pea, 2004	theoretisch	Keine Probanden
<b>Kognitive Strukturierung</b>			
	Einsiedler & Hardy, 2010	theoretisch	Keine Probanden
	Hardy & Stern, 2011	empirisch	49 Schüler im Alter von 8-9 Jahren
	Hardy, 2012	theoretisch	Keine Probanden
<b>Tutorielleres Setting mit Feedback</b>			
	Chi et al., 2001	theoretisch	Keine Probanden
	Krammer, 2010	theoretisch	Keine Probanden

Maßnahmen	Autoren	Methodik
Sustained-Shared Thinking		
Sylva et al., 2004	empirisch	3000 Probanden im Alter von 3-7 Jahren
Siraj-Blatchford & Sylva, 2012	empirisch	2860 Kinder im Alter von 3-6 Jahren

Van de Pol, Volman & Beishuizen (2010) definierten *Scaffolding* in ihrer Überblickstudie als eine Methode der Lerner-Unterstützung, die die Entwicklung eines Kindes in den Mittelpunkt rückt. Ähnlich wie bei Vygotsky wird ihrer Auffassung nach einem Lernendem unterstützt bei der Bewältigung einer Aufgabe, die er alleine nicht bearbeiten kann. Dabei steht bei Van de Pol et al. (2010) der Begriff in engem Zusammenhang mit drei charakteristischen Merkmalen. Der Erwachsene nimmt mit Hilfe diagnostischer Strategien eine fortwährende Erfassung der Lernentwicklung vor. Passend dazu stimmt der Erwachsene die Unterstützungsmaßnahmen auf das Kind hin ab, was auch als Kontingenzprinzip (Brophy, 1999; Van de Pol et al., 2010) bezeichnet wird. Durch einen engen Abstimmungsprozess von Diagnose der Lernentwicklung und Maßnahmen zur Unterstützung erfolgt eine optimale Passung, wodurch ein Kind eine effektive Förderung erhalten kann (siehe Abbildung 2 in Anlehnung an Van de Pol et al., 2010). Parallel zum Zuwachs an Fähigkeiten eines Kindes und der Übernahme von Verantwortung für den eigenen Lernprozess, verringert ein Erwachsener sukzessive das Maß an Unterstützung; diese Rücknahme an Hilfestellung wird auch als *Fading* (Van de Pol et al., 2010) bezeichnet.

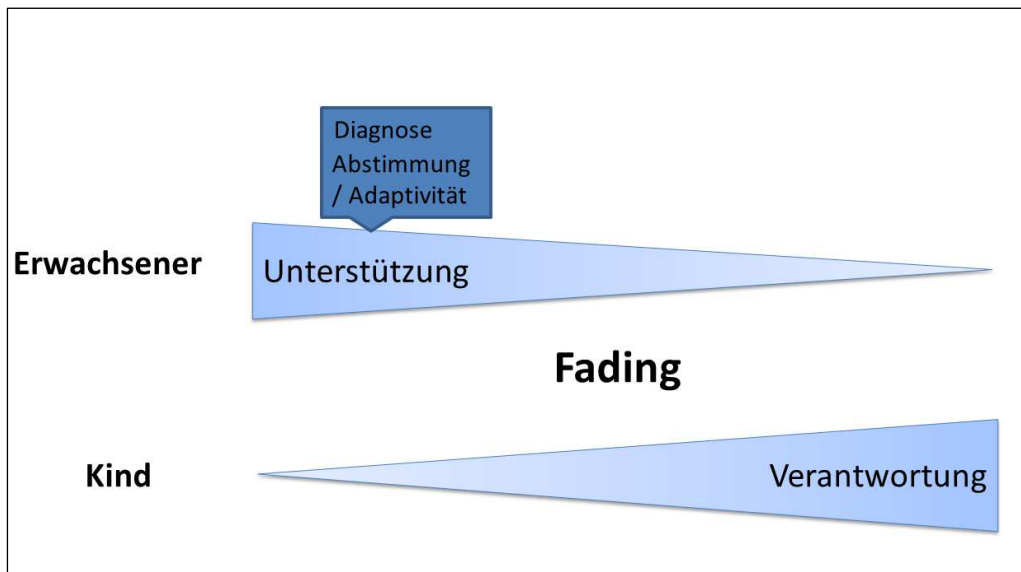


Abbildung 2. Schema zu *Scaffolding* nach Van de Pol (2010, S.15).

*Scaffolding* bezeichnet ein Vorgehen, das individuell auf den Lerner bzw. die Lerngruppe sowie die Situation bezogen wird (ebenda). Nach Reiser (2004) zielt eine Scaffolding-Maßnahme auf zwei zentrale Ziele: Zum einen kann durch die Unterstützung eine aktuelle Aufgabe gelöst werden, zum anderen wird ein Kind durch den Prozess in die Lage versetzt, seine eigene Anstrengung zu steigern und dadurch die Performanz zu optimieren.

Puntambekar & Hübscher (2005) bezeichnen die Herstellung eines gemeinsamen Verständnisses zu Beginn des Unterstützungsprozesses als eigenständiges Merkmal von *Scaffolding*, während dies bei Van de Pol et al. (2010) mit in die diagnostische Phase einfließt. Entsprechend kennzeichnen die beiden zuerst genannten Autoren *Scaffolding* durch vier Merkmale: das Herstellen eines gemeinsamen Verständnisses, die ständige Diagnose des aktuellen Verständnisniveaus des Lernenden, darauf abgestimmte Strukturierungsmaßnahmen sowie die Verringerung der Unterstützung des Erwachsenen im Laufe des Lernprozesses (Puntambekar & Hübscher, 2005).

Pea (2004) spezifiziert *Scaffolding* in seiner Analyse, indem er die Begriffe „channeling“, „focusing“ und „modeling“ einführt (Pea 2004, S.432). Beim *modeling* (Deutsch: Modellierung) spricht der Erwachsene die einzelnen Schritte seines Denkvorgangs laut aus (Collins, Brown & Newman 1989; Rogoff, 1990) und präsentiert abschließend eine richtige Lösung für die gestellte Aufgabe. Die Modellierung bietet eine intensive Form der Lerner-Unterstützung (Pea, 2004, S.432), da sie dem Lerner einen optimalen Weg zur Lösung einer Aufgabe aufzeigt. Die empirische Untersuchung von Rogoff (1990) beschäftigte sich ausführlich mit der Maßnahme der Modellierung, indem sie eine modellierende Unterstützungsmaßnahme in

tutoriellen Situationen empirisch untersuchte. Rogoff verglich zwei Gruppen im Rahmen eines experimentellen Designs miteinander und kam zu dem Ergebnis, dass fünfjährige Kinder deutlich besser lernen, wenn Erwachsene ihnen während des Prozesses die einzelnen Denkschritte aufzeigen. Ebenso lernten Kinder in einer Gruppe mit älteren Kindern besser, wenn sie von diesen durch Modellierung sowie andere prozessbegleitende Maßnahmen unterstützt wurden. Die Ergebnisse dieser Studie waren ähnlich den Ergebnissen der Trainingsstudie von Murphy & Messer (2000), in der ebenfalls Kinder zwischen fünf und sieben Jahren durch Modellierung einen effektiveren Lernzuwachs erzielten. Modellierung ist eine intensive Form der Unterstützung und daher gerade bei Kindern im Elementarbereich sinnvoll. Junge Kinder haben in der Regel wenig Vorwissen und benötigen eine umfassende Förderung zum Aufbau von eigenen Konzeptvorstellungen, insbesondere bei komplexen Denkprozessen. Aus diesen Gründen scheint Modellierung auch eine geeignete Maßnahme zur Förderung des Theorie-Evidenz Verständnisses bei jungen Kindern. Inwiefern hier Erfolge erzielt werden können, wurde in der vorliegenden Untersuchung empirisch überprüft (siehe Trainingsstudie, Kapitel 8).

Neben der Modellierung beschreibt Pea (2004) auch die Kanalisierung sowie die Fokussierung als Scaffolding-Maßnahme, die der Aufmerksamkeitslenkung dienen. Damit lassen diese Maßnahmen zur Lerner-Unterstützung sich an das Konzept der *Kognitiven Strukturierung* annähern, die ebenfalls als Teilaspekt von *Scaffolding* angesehen werden kann. Die *Kognitive Strukturierung* richtet sich aber stärker auf die prozessorientierte Unterstützung während des Lernprozesses (Einsiedler & Hardy, 2010). *Kognitive Strukturierung* dient der Verarbeitung komplexer Inhalte sowie der motivationalen Unterstützung des Lerners (Hardy, 2012). Hier werden Inhalte zum Wissensaufbau sowie zur Verfeinerung konzeptueller Vorstellungen von Lernenden aufgeschlüsselt. Bei der *Kognitiven Strukturierung* erreicht der Erwachsene nach Hardy (2012, S. 240), dass „[...] durch spezifische Aufgabenstellungen, Materialangebote, Impulse sowie eine strukturierende, fachbezogene Gesprächsführung [...] der Lernende sich kognitiv aktiv und konstruktiv mit den jeweiligen Inhalten auseinandersetzen [kann].“

Im naturwissenschaftlichen Lernen können insbesondere Visualisierungen zur kognitiven Strukturierung eingesetzt werden, da sich so Beziehungen zwischen relevanten Größen veranschaulichen lassen. Sie können als Mittel zur kognitiven Entlastung eingesetzt werden, um Inhalte besser greifbar zu machen. In der Studie von Hardy & Stern (2011) wurden zwei unterschiedliche Formen der Visualisierung bei acht- bis neunjährigen Kindern eingesetzt, um den konzeptuellen Aufbau des Dichtekonzeptes im Inhaltsgebiet *Schwimmen und Sinken* zu überprüfen. Einerseits arbeitete die Studie mit Quadraten in unterschiedlichen Größen und Farben, andererseits mit einer Balkenwaage als Repräsentationshilfe zum Aufbau eines

Dichtekonzeptes. Im Rahmen eines Prä-Post Testdesigns wurden die beiden Trainingsgruppen mit der Kontrollgruppe verglichen. Die empirischen Befunde zeigten, dass die Visualisierung des Dichtekonzeptes mittels einer Balkenwaage zum Abbau von Fehlkonzepten und gleichzeitig zum Aufbau von fortgeschrittenen Erklärungen beitrug. Im Gegensatz zur Balkenwaage trug die Visualisierung von relevanten Größen (Volumen und Masse) durch unterschiedlich große Quadrate nicht zu einem deutlichen Lernzuwachs im Hinblick auf das Dichtekonzept in dieser Gruppe bei. Eine Visualisierung bei der Lerner-Unterstützung wirkte sich positiv auf den Denkprozess eines Kindes aus, wenn die Veranschaulichung dem individuellen Verständnisniveau entsprach. Auch die empirische Studie von Simons & Klein (2007) verwendete physische Unterstützungsmaßnahmen, Artefakte oder Visualisierungen, die dort als „Hard Scaffolds“ (ebenda, S. 44) bezeichnet wurden. Durch diese Maßnahmen konnte eine Verbesserung der Performanz bei Schülern im Alter von 13 und 14 Jahren beim forschenden Lernen erreicht werden. Insgesamt hat sich die Visualisierung in empirischen Studien (z.B. Simons & Klein, 2007; Hardy & Stern, 2011) als effektive Form der Förderung sowohl bei Kindern im Grundschulbereich als auch bei Jugendlichen herausgestellt und wurde insofern auch in der vorliegenden Studie zur Verbesserung der Ergebnisse angewendet.

Eine weitere Form von Scaffolding stellt die adaptive Lerner-Unterstützung dar. Sie beruht auf Vygotskys Konzept der Zone der nächsten Entwicklung und bemüht sich um eine individuelle Abstimmung der Maßnahme auf das Lernniveau des Kindes (Decristan et al., 2015, Krammer, 2010; Van de Pol et al., 2010). Eine an das Kind angepasste Unterstützung hat den Vorteil, dass der Erwachsene ein Kind individuell unterstützen, es zu eigenen Denkschritten anregen und daher optimal fördern kann (Krammer, 2010). Die adaptive Lerner-Unterstützung zielte darauf ab, durch ein Angebot mit individueller Unterstützung, angepasst an das Lernniveau eines Kindes, eine optimale Nutzung einer Lerngelegenheit herbeizuführen (Decristan et al., 2015). Adaptive Lerngelegenheiten und ihre Wirkung in Bezug auf Lernzuwachs wurden in der groß angelegten Studie von Decristan und Kollegen (2015) untersucht. Hier wurden 1071 Kinder im Alter von acht und neun Jahren im naturwissenschaftlichen Sachunterricht über ein Schuljahr zum Thema *Schwimmen und Sinken* von 54 Lehrkräften begleitet. Im Vorfeld der Untersuchungen erhielten die Lehrer naturwissenschaftlich ausgerichtete Fortbildungen im Inhaltsgebiet *Schwimmen und Sinken*. Es wurden fünf Sitzungen über jeweils viereinhalb Stunden abgehalten, in denen Lehrern relevante Fachinhalte sowie fachdidaktisches Wissen nahegebracht wurden. So konnten Lehrer fachspezifisches Professionswissen in beiden Bereichen aufbauen (Decristan et al., 2015). Die zu sechs Messzeitpunkten erhobenen Daten gaben sowohl über die Entwicklung der Schüler wie der Lehrer Aufschluss. Die Studie war als quasi-

experimentelles Design angelegt. Das bedeutete, dass Lehrkräfte und Unterrichtsmethoden zufällig zusammengestellt wurden. In vier Experimentalgruppen wurden unterschiedliche Lerner-Unterstützungen im Rahmen eines Prä-Post-Designs untersucht: Kognitive Strukturierung (*Scaffolding instructional discourse*), das Lernen mit einem gleichaltrigen Partner (*Peer-assisted learning*) sowie Formatives Assessment (*Formative Assessment*) sowie eine Kontrollgruppe (*control group*). Dabei spielte der Umgang mit dem Vorwissen in allen vier Experimentalgruppen eine zentrale Rolle. Schülerinnen und Schüler der Gruppe mit Formativem Assessment zeigten den meisten Lernzuwachs in der Studie (Decristan et al., 2015). Dabei hatte sich in der zuvor genannten Gruppe individuelles Feedback während des Trainings nachweislich positiv auf die Entwicklung konzeptionellen Verständnisses bei Schülern ausgewirkt; sie hatten im Post-Test deutlich höhere Werte als Schüler der anderen drei Gruppen (*Peer-assisted learning, Scaffolding instructional discourse, control group*) (Decristan et al., 2015).

Als weitere Form der Lerner-Unterstützung gilt *Sustained Shared Thinking* (Siraj-Blatchford & Sylva, 2012; Hopf, 2011). Der Begriff *Sustained Shared Thinking* (Siraj-Blatchford, Sylva, Muttock, Gilden & Bell, 2002, S. 8; Hopf, 2011) beschreibt eine Unterstützungsmaßnahme innerhalb tutorieller Settings. Die Unterstützung durch *Sustained Shared Thinking* umfasst einen umfassenden Prozess des Austausches zwischen Fachkraft und Kind bzw. Kindern und findet in der Regel in 1:1 Situationen statt. Definiert wird *Sustained Shared Thinking* als Interaktion zwischen zwei oder drei Menschen, die in intellektueller Weise zusammenarbeiten, um ein Problem zu lösen, ein Konzept zu erläutern oder Aktivitäten zu evaluieren (Siraj-Blatchford et al, 2002). Die Form der Unterstützung ist durch Interaktion bzw. gemeinsames Nachdenken geprägt: „[...] an effective pedagogic interaction, where two or more individuals ,work together in an intellectual way to solve a problem, clarify a concept, evaluate activities, or extend a narrative.“ (Siraj-Blatchford et al., 2002, S.8)

Dabei können Gedanken, die von einem Kind ausgehen, vom Erziehenden ausgeführt und erweitert werden. Ebenso können Gesprächsanlässe, die vom Erziehenden initiiert sind, als Grundlage für einen Gedankenaustausch genutzt werden (Siraj-Blatchford, 2012). Diese Form des Austausches ist im Zusammenhang mit Ko-Konstruktion (siehe auch Kapitel 3.1) von Wissen zu betrachten (Siraj-Blatchford et al., 2002). Frei gewählte Aktivitäten eines Kindes bieten die beste Möglichkeit, um die zuvor erläuterte Situation zu schaffen. *Sustained Shared Thinking* legt anders als andere Scaffolding-Maßnahmen den Schwerpunkt auf einen ausgewogenen Austausch von Ideen zwischen den Gesprächspartnern, um gemeinsam Bedeutungen zu Sachverhalten bzw. Dingen herauszuarbeiten.

Die groß angelegte Längsschnittstudie „Effective Provision for Pre-School Education“ (EPPE) (Sylva, Melhuish, Sammons, Siraj-Blatchford, Taggart, 2004) analysierte kognitiv herausfordernde Gesprächsepisoden bei Kindern im Alter von vier bis fünf Jahren (Siraj-Blatchford & Sylva, 2012) und brachte somit Erkenntnisse für die Wirkung von Sustained Shared Thinking. Bei der Qualitätserfassung in dieser Studie spielte immer die Angemessenheit der kognitiven Herausforderung, abgestimmt auf den zu erwartenden Entwicklungsstand des Kindes, eine Rolle (Sylva, Roy & Painter 1980). Die Befunde (siehe Abbildung 3) zeigten, dass in exzellenten Erwachsenen-Kind-Interaktionen/Erzieherinnen/Erzieher-Kind-Dialogen, gemessen an der kognitiven Herausforderung, die initiativen Gesprächsanteile des Kindes etwas geringer waren als in qualitativ guten Interaktionen (Siraj-Blatchford & Sylva, 2012). Entsprechend sind Gesprächsanteile, die vom Erwachsenen ausgingen, in exzellenten Dialogen höher als in guten Erzieherinnen/Erzieher-Kind-Dialogen.

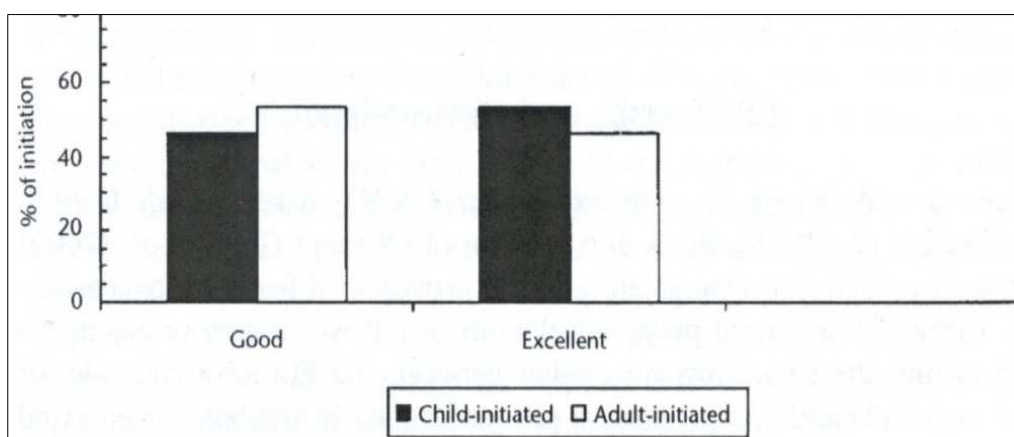


Abbildung 3. Erzieher-Kind-Dialoge I: Verhältnis der Lernprozesse, die von Erwachsenen bzw. Kindern initiiert wurden (Abbildung aus Siraj-Blatchford & Sylva, 2012, S.155).

Bei der Analyse der Gesprächsepisoden stellte sich heraus, dass Aussagen besonders kognitiv aktivierend waren und folglich den Lernenden auf die nächste Ebene in seinem Denken führten, wenn der Erwachsene die Gedanken eines Kindes im Rahmen des Gesprächs themenbezogen erweiterte und mit altersangemessenen Inhalten anreicherte (siehe Abbildung 4). Ein Erwachsener formulierte in exzellenten Erwachsenen-Kind-Situationen deutlich ausführlicher die kindlichen Gedanken aus als in guten pädagogischen Situationen. Dabei griff ein Erwachsener in sehr guten Interaktionen die Gedanken des Kindes auf und elaborierte sie, indem er sie in einen größeren fachlichen Zusammenhang stellte.



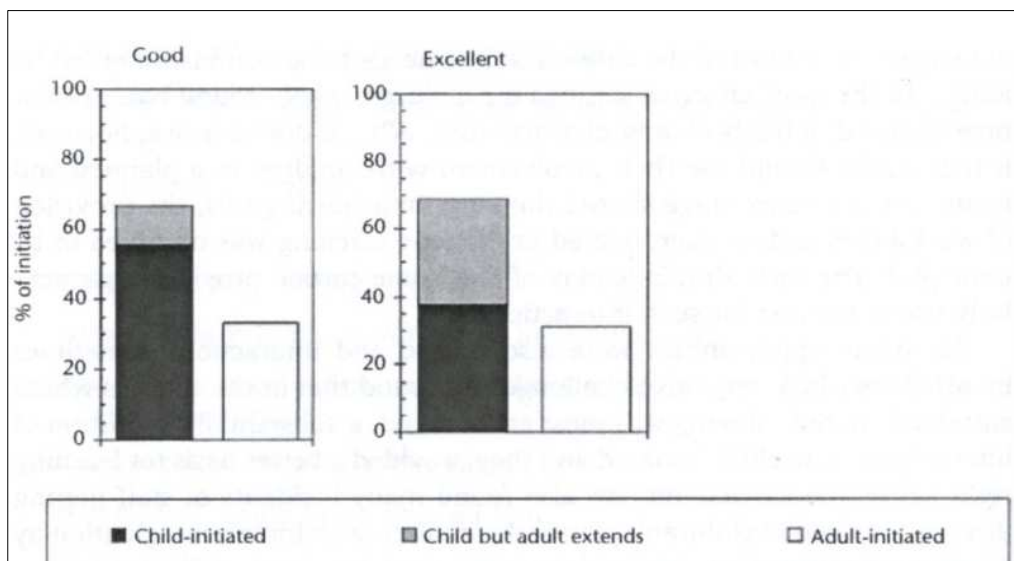


Abbildung 4. Erzieher-Kind-Dialoge II: Initiation zu hoher kognitiver Leistung (Abbildung aus Siraj-Blatchford & Sylva, 2012, S.155).

In der Erweiterung der Überlegungen eines Kindes durch einen Erwachsenen wurde das kindliche Vorwissen auf besondere Weise berücksichtigt (Brophy, 1999; Siraj-Blatchford & Sylva, 2012), wodurch die Anschlussfähigkeit an vorhandenes Wissen ermöglicht wurde. So konnte das Denken eines Kindes auf die nächste Ebene geführt werden. Dieser entscheidende Moment in der Entwicklung wurde von Siraj-Blatchford & Manni daher auch als „kritische[r] Moment“ bezeichnet (Siraj-Blatchford & Sylva, 2008). Als kognitiv herausfordernd galten ebenso offene Fragen, die ein Kind zum Denken anregen (Siraj-Blatchford & Manni, 2008; Siraj-Blatchford & Sylva, 2012; Clark & Grave, 2005). Meist traten Situationen, in denen gemeinsam Wissen im Sinne von *Sustained Shared Thinking* konstruiert wurde, in 1:1 Erwachsenen-Kind-Situationen auf. Analysen haben offengelegt, dass qualitativ sehr hochwertige pädagogische Interaktionen (Abbildung 4) deutlich mehr kritische Momente (Siraj-Blatchford & Sylva, 2012, S.155) aufgewiesen hatten als solche von guter Qualität (ebenda). Die Befunde (Siraj-Blatchford & Sylva, 2012) zeigten, dass exzellente Interaktionen zwischen pädagogischer Fachkraft und Kind im Vergleich zu guten Gesprächssituationen häufiger durch Phasen gemeinsamer Denkkonstruktionen sowie durch umfangreiche Instruktionsphasen gekennzeichnet waren. Die qualitativen Unterschiede beider Situationen waren auch dadurch offensichtlich, dass weniger regulierende Steuerungsmomente im Verlauf der Interaktion in exzellenten Situationen erfolgten. Das zeigte sich daran, dass die vom Kind initiierten Äußerungen einen höheren Anteil haben als in guten Gesprächssituationen; ebenso sind die Gesprächsanteile des Erwachsenen insgesamt geringer in exzellenten Gesprächssituationen (siehe Abbildung 3).

Insgesamt deuteten die Befunde der EPPE-Studie auf die proaktive Rolle der pädagogischen Fachkraft hin, die immer wieder die inhaltliche Fundierung kindlicher Aussagen vornahm und damit ein Kind zum Denken anregen konnte.

In diesem Unterkapitel wurden verschiedene Studien zu Lerner-Unterstützung vorgestellt: eine Studie zu Scaffolding (Van de Pol, 2010), eine Studie zur Kognitiven Strukturierung (Hardy & Stern, 2011) eine Studie zur Modellierung (Murphy & Messer, 2000; Pea, 2004), eine Studie zur adaptiven Lerner-Unterstützung (z.B. Decristan et al., 2015) sowie eine Studie zu *Sustained Shared Thinking* (Siraj-Blatchford & Sylva, 2012). Die vorliegende Arbeit hat die Befunde der empirischen Studien aufgegriffen, um drei erfolgversprechende Formen von Scaffolding für Kinder im Elementarbereich zu konzipieren. In der ersten experimentellen Variation wird die adaptive Lerner-Unterstützung (z. B. Decristan et al., 2015), in der zweiten experimentellen Variation die Modellierung (Pea, 2004) angewendet. In beiden experimentellen Bedingungen werden Formen der Visualisierung (Hardy & Stern, 2011) verwendet, um Kindern spezifische Eigenschaften von Objekten bewusst zu machen.

### 3.3 *Scientific Literacy im Elementarbereich in Deutschland*

Das Konzept *Scientific Literacy* (Bybee, 1997; Prenzel & Artelt, 2007) stammte aus dem anglo-amerikanischen Raum und wird als naturwissenschaftliche Grundbildung übersetzt. *Scientific Literacy* (Bybee, 1997; Prenzel & Artelt, 2007) meint den Erwerb naturwissenschaftlicher Grundkenntnisse mit inhaltsspezifischer Ausprägung sowie damit verbundener Denk- und Arbeitsweisen (Bybee, 1997; Prenzel & Artelt, 2007). Beim Konzept der *Scientific Literacy* wird zwischen dem Wissen über naturwissenschaftliche Methoden (Theorien, Konzepte, Wissenschaftsverständnis) und dem bereichsspezifischen, konzeptuellen Wissen differenziert. *Scientific Literacy* fand Eingang in die Erläuterung von Kompetenzen in der PISA-Studie (2006) und ebenso in die Beschreibung der Standards für den Sachunterricht für den Grundschulbereich in Deutschland (GDSU, 2002), inzwischen ist *Scientific Literacy* in der deutschen Bildungsdebatte breit akzeptiert.

*Scientific Literacy* war ursprünglich ein schulisches Konzept, wird in den letzten Jahren aber auch auf den Elementarbereich übertragen (Steffensky, Lankes, Carstensen & Nölke, 2012). Seit 2009 fordert die Kultusministerkonferenz, Naturwissenschaften in die Bildungspläne der Länder für den Elementarbereich aufzunehmen. Dabei soll naturwissenschaftliche Grundbildung im Elementarbereich im Sinne der *Scientific Literacy* darauf ausgerichtet sein, Wissen über grundlegende naturwissenschaftliche Konzepte sowie Denk- und Arbeitsweisen

als „erfahrungsbasiertes, anschlussfähiges und alltagsnahes Wissen“ (Steffensky et al., 2012, S. 39). Kindern nahebringen, um naturwissenschaftliche Ereignisse angemessen einschätzen zu lernen (ebenda).

Durch das Bereitstellen von Lernangeboten zu unterschiedlichen Themen kann bereits früh konzeptuelles, inhaltliches Wissen aufgebaut werden (Samarapungavan, 2008; Sodian, 1998; Ziegler, Hardy, 2015). Mit naturwissenschaftlichen Lernangeboten im Kindergarten können Kinder früh anschlussfähiges Wissen entwickeln (Möller, 2004; Spreckelsen, 2007; Steffensky, 2008), das in der Grundschule mit weiteren Inhalten angereichert und so zunehmend differenziert werden kann. Auf diese Weise können Kinder zu einem vertieften Verständnis für Naturwissenschaften gelangen. Diese Lernprozesse werden auch als „kumulative Bildungsprozesse“ (Kauertz, 2012, S. 92) bezeichnet. Damit naturwissenschaftliche Angebote lerneffektiv sind, müssen sie auf Grundlage prozessbezogener Denk- und Arbeitsweisen konzipiert sein (Steffensky, 2008; Möller & Steffensky, 2010). Die genannten Autoren gehen davon aus, dass naturwissenschaftliches Lernen erfolgreich sein kann, wenn eine Verknüpfung von inhaltlichem Lernen mit dem Erlernen von naturwissenschaftlichen Arbeitsweisen angestrebt wird.

Neben kognitiven Aspekten wird im Rahmen der naturwissenschaftlichen Grundbildung die Förderung emotionaler Fähigkeiten auch angestrebt, um Kinder auf der Gefühlsebene anzusprechen, sodass sie neugierig werden und eigenständig Interesse entwickeln (Norris & Philips, 2003). Anregende Fragestellungen und altersgemäße Lernsettings gelten als vielversprechend, um Interesse für wissenschaftliche Inhalte zu wecken (Bybee, 1997; Bybee & McCrae & Laurie, 2009; Norris & Philips, 2003). Dabei werden Methoden zur Lösung von Problemen und Fragestellungen aus der Erfahrungswelt der Kinder eingesetzt (Steffensky, 2008). Im Kindergartenalltag kann naturwissenschaftliche Grundbildung insbesondere mit Hilfe von Experimenten angebahnt werden (Steffensky, 2008, Aufschnaiter & Riemann, 2005).

Verschiedene didaktische Anregungen für Kinder werden in den genannten Bildungsbereichen angeboten, damit Fachkräfte die naturwissenschaftlichen Kompetenzen von jungen Kindern unterstützen können (Möller & Steffensky, 2010). Beispielsweise ist das Thema Magnetismus im Spiralcurriculum als Handreichung für Erzieherinnen/Erzieher und Lehrerinnen/Lehrer zur Unterrichtsvorbereitung im Handel erhältlich (Anders, Hardy, Pauen & Steffensky, 2013). Für die drei Bildungsbereiche werden jeweils zusätzlich Materialkisten angeboten, die Fachkräfte bei der Planung ihrer Lerngelegenheiten einsetzen können.

Zudem gibt es Fortbildungen für Fachkräfte der einzelnen Bildungsbereiche im naturwissenschaftlichen Bereich, die entweder von Universitäten (z.B. Westfälische Wilhelms Universität Münster) oder von Stiftungen (z.B. Klaus Tschira Stiftung, Polytechnische Stiftung)

durchgeführt werden. Dabei steht die Erweiterung fachspezifischer sowie didaktischer Kompetenzen von Fachkräften (Erzieherinnen/Erziehern und Lehrerinnen/Lehrern) im Vordergrund.

Die beiden Stiftungen zur naturwissenschaftlichen, mathematischen und technischen Bildung von Kindern im Elementar- und Primarbereich werden im Folgenden kurz erläutert: einerseits die bundesweite Bildungsinitiative Haus der kleinen Forscher, andererseits die Klaus Tschira Stiftung. Die Wirksamkeit der verschiedenen Maßnahmen beider Stiftungen ist bisher noch nicht wissenschaftlich belegt.

Die Stiftung Haus der kleinen Forscher wurde auf Initiative einer Unternehmensberatung (<http://www.haus-der-kleinen-forscher.de/de/ueberuns/die-stiftung/>) gegründet und wird von dieser unterstützt. Zusätzlich fördert das Bundesministerium für Bildung und Forschung (BMBF) die Initiative. Die Stiftung hat lokale Netzwerke aufgebaut, insbesondere in Berlin und Brandenburg sowie in Baden Württemberg. Die einzelnen Zentren unterstützen Kindertagesstätten sowie Grundschulen dabei, naturwissenschaftliche, mathematische und technische Themengebiete interessant und altersgemäß aufzubereiten und Kindern nahezubringen. In Deutschland sind über 1000 Einrichtungen an ein lokales Netzwerk der Kleinen Forscher angegliedert (Gold & Dubowy, 2013). Ziel der Stiftung ist es, Erzieherinnen/Erziehern und Lehrerinnen/Lehrern über Fortbildungen (<http://www.bmbf.de/press/3179.php/>) die Möglichkeit zu bieten, Expertise in naturwissenschaftlichen Themenbereichen sowie in deren Vermittlung aufzubauen. Mit handlungsorientiertem Wissen können pädagogische Fachkräfte Freude und Neugier für naturwissenschaftliche, mathematische sowie technische Phänomene bei Kindern wecken. Die Stiftung Haus der kleinen Forscher bietet außerdem Materialien, beispielsweise Forschungskarten mit Anregungen zu spezifischen Themen, mit denen sie Fachkräfte und Kinder beim Forschen und Experimentieren unterstützt (Gold & Dubowy, 2013).

Die Klaus-Tschira-Stiftung wurde 1995 von dem gleichnamigen Physiker gegründet (<http://www.klaus-tschira-stiftung.de/>). Wie das Haus der kleinen Forscher unterstützt auch diese Stiftung Bildungsmaßnahmen im naturwissenschaftlichen, mathematischen sowie im technischen Bereich in Kindergärten und Grundschulen. Zusätzlich fördert die Klaus Tschira Stiftung Lernen im informatischen Bereich. Neben einem breit gefächerten Fortbildungsangebot für Erzieherinnen/Erzieher und GrundschullehrerInnen/Grundschullehrer verfügt die Stiftung über eine Forscherstation (<http://www.forscherstation.info/>), die eine Auswahl an themenspezifischen Kisten zum Einsatz in der Praxis verleiht. Auf diese Weise können pädagogische Fachkräfte Expertise und Ideen sammeln, um sie in der Praxis anzuwenden.

Die Angebote beider Stiftungen werden jeweils von Universitäten, Fachhochschulen bzw. Pädagogischen Hochschulen begleitet, um die optimale Ausgestaltung und Wirkungsweise früher Förderprogramme in naher Zukunft evaluieren zu können.

### 3.4 Exkurs: US-amerikanische Förderprogramme im letzten Jahrzehnt

Im Folgenden sollen drei US-amerikanische Programme zum frühen naturwissenschaftlichen Lernen vorgestellt werden, da diese Einfluss auf die Förderung frühkindlicher Bildung in den letzten Jahren in Deutschland genommen haben.

Folgende Programme zur Förderung naturwissenschaftlichen Lernens im Vorschulbereich sind in den USA implementiert und evaluiert worden: „Head Start on Science and communication“ (Klein & Hammrich, 2000), „Science Start!“ (French, 2004) „Science learning pathways for young children“ (PrePS) (Gelman & Brenneman, 2004). Diese Programme zeigten in grundlegenden Aspekten Übereinstimmungen. So gingen alle drei von der Annahme aus, dass der Aufbau von mentalen Strukturen durch altersangemessene Angebote zentral für die kognitive Entwicklung von Kindern ist. Alle drei begriffen Lernen außerdem als einen aktiven Konstruktionsprozess und arbeiteten mit Themenfeldern, die einen Bezug zur Lebenswelt von Kindern haben und somit bedeutsam waren.

Ebenso waren alle Programme darauf ausgerichtet, die naturwissenschaftliche Grundbildung zu fördern (Bybee, 1997; OECD, 2004). Kinder sollten durch altersgemäße, ansprechende Lerngelegenheiten am Beispiel von interessanten Themen aus dem naturwissenschaftlichen Bereich ein Verständnis sowohl für Konzepte, Theorien und Begriffe als auch für naturwissenschaftliche Denk- und Arbeitsweisen aufbauen (Steffensky, 2011). Auf diese Weise können Kinder mit dem naturwissenschaftlichen Zyklus vertraut gemacht werden. In chronologischer Ordnung enthält ein naturwissenschaftlicher Zyklus folgende Elemente: Aufstellen von Fragen, Formulieren von Hypothesen, Vermuten, Vorhersagen, Beobachten, Interpretieren von Ergebnissen sowie Schlussfolgern. Neben der Ausprägung kognitiver Fähigkeiten sollten Kinder auch auf affektiv-emotionaler Ebene angesprochen werden, indem ihr Interesse sowie ihre Neugier geweckt werden (Norris & Philipps, 2003).

Trotz der genannten Gemeinsamkeiten können auch bestimmte Schwerpunkte für die einzelnen Programme genannt werden. „Science Start!“ (French, 2004) wurde in Zusammenarbeit von US-amerikanischen Lehrerinnen/Lehrern, Doktorandinnen/Doktoranden sowie Professorinnen/Professoren als umfassendes Konzept entwickelt. Es wurde in 37 ganztägigen Kindertagesstätten in New York und Umgebung umgesetzt. Vorgesehen waren vier thematisch

gebundene, aufeinander aufbauende Module, die jeweils zehn bis zwölf Wochen zum Einsatz kommen und unterschiedliche naturwissenschaftliche Angebote mit ausgeprägter Strukturierung enthalten. Die Lerneinheiten wurden durch verschiedene sprach- oder literaturbezogene Inhalte angereichert. Bei „Science Start!“ (French, 2004) handelte es sich um einen integrativen Ansatz, der naturwissenschaftliches Lernen, insbesondere Themen aus dem physikalischen oder biologischen Bereich, mit der Förderung sprachlicher Fähigkeiten verknüpft. Dabei baute der Ansatz auf der Annahme auf, dass Kinder im Alter von vier bis fünf Jahren über einfache mentale Repräsentationen verfügen, die es ihnen ermöglichen, alltägliche Situationen einzuschätzen:

„These generalized mental representations of events are the foundations for children’s understanding of and action within the everyday world. They enable children to recognize regularities, to interpret their everyday experiences, to predict how and why people will behave in certain ways, and to behave approximately themselves.” (French, 2004, S. 140)

Beim Programm „Head Start on Science and communication“ (Klein & Hamrich, 2000) stand hingegen die individuelle Förderung eines Kindes im Vordergrund. Zentrales Anliegen des Programmes war die Förderung von prozessorientierten Kompetenzen, indem die Kinder bei unterschiedlichen Themen wiederholt eigene Vermutungen äußern, beobachten sowie überprüfen und damit wissenschaftliche Arbeitsweisen kennenlernen. Das Programm bestand aus zwei Phasen und wurde in 18 Schulen in zwei amerikanischen Städten, Philadelphia und New Jersey, durchgeführt. In der ersten Phase wurden die Programmziele mit Lehrerinnen/Lehrern und Eltern erarbeitet, die in der zweiten Phase im Rahmen von instruktionalen Lernsettings zum naturwissenschaftlichen Lernen umgesetzt wurden.

Das Programm richtete sich speziell an Kinder aus sozioökonomisch benachteiligten Familien und war darauf ausgerichtet, auf deren spezielle Bedürfnisse einzugehen und die Kinder optimal zu fördern. Als Ausgangspunkt für den Aufbau von Wissen galten visuell ansprechende Lernmaterialien bzw. eine strukturierte Lernumgebung. Diese wurden in unterschiedlichen Schwierigkeitsstufen angeboten. Sie reichten von Aufgaben, in denen Kinder Objekte miteinander verglichen über Aufgaben zur Diskrimination bis hin zu solchen, in denen Kategorieneildung oder Zuordnungen nach bestimmten Kriterien erforderlich waren. Dabei bestimmten Fragen sowie individuelle Interessen der Kinder die Auswahl der Aufgaben, die der Erwachsene im Lernangebot machte. Dieser nahm eine eher moderierende Funktion ein, indem er auf Fragen der Kinder einging oder gemeinschaftliche Lernsituationen von Kindern organisierte;

von der Seite des Erwachsenen gab es jedoch keine strukturierenden Maßnahmen. Klein und Hammrich (2000) fassten das Ziel des Programms folgendermaßen zusammen:

“The goal of the exploratory teaching model is to create an environment, in which children may explore, learn, and develop when involved with naturally interesting materials and events. In such a setting, there are no structured responses. Rather, activities lend themselves to creativity and exploration.”  
(Klein & Hammrich, 2000, S.84)

Der Fokus bei diesem Programm lag darauf, dass die Kinder im Rahmen ihrer Aktivitäten den naturwissenschaftlichen Zyklus selbst durchlaufen und sukzessive selbstständig Hypothesen aufstellen, Vermutungen äußern sowie Vorhersagen treffen. Nach dem Durchführen der Experimente beobachteten und interpretierten Kinder ihre Ergebnisse, schlussfolgerten und berichteten diese. Hier wurden demzufolge naturwissenschaftliche Fähigkeiten im Sinne einer naturwissenschaftlichen Grundbildung bzw. *Scientific Literacy* sowohl in inhaltsspezifischer als auch prozessualer Hinsicht erreicht. Es ging darum, dass Kinder bessere Fähigkeiten in einer spezifischen Domäne erreichten (Klein & Hammrich, 2000).

Ähnlich wie im zuvor beschriebenen Programm wurde in „Science learning pathways for young children“ (*PrePS*) (Gelman & Brenneman, 2004), einem Programm das von Entwicklungspsychologen zusammengestellt wurde, von einem domänenspezifischen Ansatz ausgegangen. Kinder sollten Strukturen in einzelnen Inhaltsbereichen aufbauen. Die Autoren schlugen vor, Präkonzepte, die vier- bis fünfjährige Kinder in einzelnen Bereichen gebildet hatten, als Basis zum Aufbau von Lernsequenzen zu nutzen (Bowman, 2001; Gelman & Brenneman, 2004). Kinder konnten so ihre eigenen Vorstellungen auf Grundlage von neuen Inhalten abgleichen und umstrukturieren. Ebenso wurde davon ausgegangen, dass Kinder leicht unterschiedliche Themen inhaltlich miteinander verknüpfen und so Verbindungen zwischen diesen Themen herstellen konnten (Bowman, 2001). Wie bei „Science Start!“ wurde auch in diesem Programm Wert auf die Vermittlung eines Verständnisses für den wissenschaftlichen Zyklus gelegt. Daher wurden prozessorientierte Kompetenzen (wie beispielsweise Frage formulieren, Experiment planen und durchführen, Evaluation von Evidenz, Hypothese bestätigen oder widerlegen) im Rahmen unterschiedlicher Themen gefördert.

Zusammenfassend kann festgehalten werden, dass unterschiedliche methodische Ansätze den drei Programmen zugrunde lagen. Während das Programm „Science Start“ strukturierte Lernsettings als Grundlage nutzte, waren die beiden anderen Programme, „Head Start on Science and Communication“ und „Science learning pathways for young children“, stärker auf individuelle Förderung von Kindern ausgerichtet. Sie gingen daher bei der Lerner-

Unterstützung mehr auf die einzelnen Bedürfnisse der Kinder ein und waren bestrebt, eine adaptive Lernumgebung zu bieten. Alle drei Programme hatten die Förderung wissenschaftlicher Denk- und Arbeitsweisen sowie den Aufbau eines konzeptionellen, inhaltlichen Wissens zum Ziel. Die Ausrichtung der Programme sowie deren Zielsetzungen stehen den Trainingsmaßnahmen in der vorliegenden Studie nahe und wurden daher im Exkurs erläutert.

### 3.5 Zusammenfassung und Fazit

In diesem Kapitel wurde zunächst der soziale Konstruktivismus mit Vygotsky als Hauptvertreter aufgeführt, da bei ihm die optimale Passung von Lernniveau eines Kindes und Unterstützungsmaßnahme kennzeichnet für eine optimal gestaltete Maßnahme ist. Dafür charakteristisch ist ebenso ein Ko-Konstruktionsprozess, mit dem ein Austausch zwischen zwei oder mehreren Personen beschrieben werden: zwischen Erwachsenem und Kind oder auch zwischen zwei oder mehreren Kindern (Kapitel 3.1). Weiterhin wurde das Konzept Scaffolding sowie verwandte Formen von Lerner-Unterstützung im Hinblick auf ihre Wirksamkeit am Beispiel von verschiedenen Studien erläutert, um die für die Arbeit zentralen Trainingsmaßnahmen genau zu beschreiben (Kapitel 3.2). Dazu gehören als intensive Form der Lerner-Unterstützung die Modellierung (Pea, 2004), die Visualisierung (Hardy & Stern, 2011; Simons & Klein, 2007) sowie die adaptive Lerner-Unterstützung (Decristan et al., 2015). Ebenso wurden das Konzept *Scientific Literacy* und dessen Verankerung in deutschen Bildungsplänen sowie unterschiedliche Bestrebungen dargestellt, naturwissenschaftliches Lernen bei Kindern im Vorschul- und Grundschulalter umzusetzen (Kapitel 3.3). Das Wirken sowie die Zielsetzungen von zwei privaten, deutschen Stiftungen, Klaus Tschira Stiftung sowie das Haus der kleinen Forscher, wurden kurz erläutert. Im Exkurs (Kapitel 3.4) wurden drei US-amerikanische Programme zur Förderung naturwissenschaftlichen Lernens vorgestellt, da diese Einfluss auf die frühkindliche Förderung im naturwissenschaftlichen Bereich genommen hatten. Insgesamt wurden in diesem Kapitel methodische Vorgehensweisen in ihrer theoretischen Fundierung sowie Möglichkeiten der Umsetzung vorgestellt, da sie für die Konzeption der Trainingsmaßnahmen in der vorliegenden Studie relevant waren.



#### 4 Die Entwicklung der Fähigkeit zur Theorie-Evidenz-Koordination

Die Kompetenz zur Koordination von Theorie und Evidenz ist grundlegend für wissenschaftliches Denken. Kuhn (2002) definierte wissenschaftliches Denken als eine systematische Suche nach Erkenntnis, die sich auf das Zusammenspiel von Theorie und Evidenz konzentriert. Dabei sind Kompetenzen zur Aufstellung, Prüfung und Revision von Theorien auf Basis der Evidenz zentrale Bestandteile des Erkenntnisprozesses bei unterschiedlichen inhaltlichen Phänomenen.

Wissenschaftliches Denken wird in der Literatur definiert als „Fähigkeit, durch systematische Strategien der Hypothesenprüfung zu wissenschaftlicher Erkenntnis zu kommen“ [und beinhaltet ebenso die Reflexion] „der Wissenskonstruktion“ (Sodian, Jonek, Thoermer & Kircher, 2006, S. 149). Demnach umfasst wissenschaftliches Denken zwei Bereiche: Zum einen das prozessbezogene Wissen, das Kenntnisse über wissenschaftliche Denk- und Arbeitsweisen sowie den Wissenschaftlichen Zyklus (Sodian et al., 2006); zum anderen das konzeptuelle, inhaltliche Wissen, das auf spezifische Phänomene ausgerichtet ist. Nach Koerber et al. (2014) bezeichnet wissenschaftliches Denken die Fähigkeit, Theorie und Evidenz zu unterscheiden.

Die Fähigkeit zum wissenschaftlichen Denken, die die zuvor beschriebenen Facetten einschließt, entwickelt sich mit zunehmendem Alter. Daher werden verschiedene Lernniveaus des wissenschaftlichen Wissens beschrieben, Anschließend werden zentrale Fähigkeiten, die sich bereits im frühkindlichen Alter herausbilden, dargestellt. Sie sind relevant für die Entwicklung von Kompetenzen bei der Koordination von Theorie und Evidenz, die danach erläutert wird. Im Anschluss daran wurden empirische Studien zum wissenschaftlichen Denken beschrieben, die verschiedene Facetten wie z.B. die Hypothesenprüfung, Experimentierfähigkeiten sowie Fähigkeiten zum Argumentieren bei Kindern sowie Erwachsenen untersuchten.

##### *4.1 Verschiedene Kompetenzniveaus*

Wissenschaftliches Denken schließt Kenntnisse wissenschaftlicher Methoden über unterschiedliche Inhaltsgebiete und Altersstufen ein. Gemäß der Konzeptwechselforschung entwickelt sich wissenschaftliches Denken gemäß einer sukzessiven Entwicklung (Limón, 2001; Duit & Treagust, 2008; Vosniadou, 2007). Drei Kompetenzniveaus werden dabei unterschieden (z.B. Hardy et al., 2010a). Junge Kinder verfügen häufig über naive Vorstellungen (Niveau 1), die sich auf Fehlannahmen gründen und im Laufe der kognitiven Entwicklung grundlegend umstrukturiert werden (Carey et al., 1989; Hardy et al., 2010a). Dabei lassen sich Unterschiede zum nächst höheren Niveau in naturwissenschaftlichen Inhaltsgebieten feststellen, da Kinder hier bereits belastbare, ausbaufähige Vorstellungen entwickeln (Niveau 2), um

naturwissenschaftliche Phänomene zu erklären. Häufig werden diese auch mit „Zwischenvorstellungen“ (Hardy et al., 2010a, S. 116) bezeichnet. Durch kognitiv aktivierenden Unterricht (Hardy, Jonen, Möller & Stern, 2006) mit geeigneten Strukturierungsmaßnahmen können Kinder im Grundschulalter erste wissenschaftliche Vorstellungen (Niveau 3a) aufbauen, die der Prüfung durch ein Experiment standhalten können. Auf diesem Niveau können Kinder folglich ihre Vorstellung so fundiert formulieren, dass sie diese mit der Evidenz in einem Experiment belegen können. Ein differenzierteres wissenschaftliches Verständnis, bei denen Kinder tragfähige Konzepte simultan in ihre Begründung integrieren, ist auf dem nächst höheren Niveau (Niveau 3b) angesiedelt (Hardy et al. 2006; Hardy et al., 2010a).

Die Entwicklung des inhaltlichen, konzeptuellen Wissens, das in drei Niveaus unterteilt ist (Carey, Evans, Honda & Unger, 1989; Hardy et al., 2006; Hardy et al., 2010a), wird im folgenden Absatz am Beispiel vom Inhaltsgebiet *Schwimmen und Sinken* dargestellt. Für ein inhaltlich ausgereiftes Verständnis ist das Berücksichtigen von zwei relevanten Größen eines Gegenstandes, Volumen und Masse, zentral. Auf dem ersten Niveau verfügen Kinder über naive Vorstellungen, die nur auf einen Aspekt des betrachteten Objekts fokussieren (Hardy et al., 2006; Hardy et al., 2010a).

Eine typische Vermutung von Kindern mit einem naiven Verständnis war beispielsweise „Dinge, die klein sind, schwimmen.“ Die Aussage zeigte, dass Kinder häufig nur auf die Größe eines Gegenstandes achten und in ihrer Aussage noch nicht das Gewicht, Volumen bzw. die umgebende Flüssigkeit berücksichtigen (Hardy et al., 2006). Durch Instruktionen bzw. mit zunehmendem Alter entwickeln Kinder sogenannte „Zwischenvorstellungen“ (Hardy et al., 2010 a, S. 117) und berücksichtigen in ihren Erklärungen die Rolle des Wassers bzw. die Beschaffenheit des Materials. Typische Aussagen auf dem zweiten Niveau waren zum Beispiel: „Das schwimmt, weil es leichter als Wasser ist.“ oder „Der Knopf schwimmt, weil er Löcher hat.“ (ebenda, S. 117). Auf dem dritten Niveau (Niveau 3a) konzeptuellen Verstehens wenden Kinder in ihren Erklärungen eines von verschiedenen physikalischen Gesetzen an, eine typische Aussage war beispielsweise: „Der Knopf schwimmt, weil er vom Wasser nach oben gedrückt wird.“ (Hardy et al., 2006). Integrieren Kinder zwei oder alle drei physikalischen Gesetze (Dichtekonzept, Druck des Wassers sowie Auftriebskraft) in ihrer Erklärung über Schwimmen und Sinken, so zeigen sie damit ein ausgeprägtes konzeptuelles Verständnis über das Inhaltsgebiet *Schwimmen und Sinken*. Empirische Befunde (Tytler & Peterson, 2005) belegten, dass Schüler mit einem besseren konzeptuellen Verständnis auch differenziertere Aussagen im Hinblick auf Theorie und Evidenz zeigen.

Parallel zur stufenweisen Entwicklung vom konzeptuellen, inhaltlichen Wissen wird beim Verständnis über Naturwissenschaften von einer ähnlichen Entwicklung ausgegangen.

Es besteht die Annahme, dass das Verständnis über Wissenschaften sich ebenso wie das konzeptuelle inhaltliche Wissen in drei Stufen entwickelt. Diese wurden im Rahmen der Kompetenzentwicklung in Studien zum Wissenschaftsverständnis beschrieben (Sodian et al., 2006; Sodian & Mayer, 2013). Wissenschaftsverständnis beinhaltet Kenntnisse darüber mit welcher Systematik ein Wissenschaftler arbeitet und vor allem, welche Ziele er damit verfolgt, beispielsweise möchte ein Wissenschaftler mit geeigneten Methoden eine Theorie prüfen und sie entsprechend bestätigen oder widerlegen (Sodian et al., 2006). Zum Wissenschaftsverständnis gehört ein Verständnis über Methoden, Denk- und Arbeitsweisen, der Logik des Experimentierens sowie das Begreifen des wissenschaftlichen Zyklus (Frage formulieren, Experiment planen und durchführen, Evaluation von Evidenz, Hypothese bestätigen oder widerlegen). Im Zusammenhang mit dem Modell von Klahr (*Scientific Dual Discovery Problem Solving Process*, siehe Kapitel 4.3) wird der Vorgang des Experimentierens im Rahmen des wissenschaftlichen Zyklus dargestellt.

Nach dem Modell von Sodian et al. (2006) verläuft die Entwicklung von Wissenschaftsverständnis bei Kindern auf drei Niveaus (siehe Kapitel 4.2). Folglich erkennen Kinder auf dem ersten Niveau noch keinen Zusammenhang zwischen Theorie und Evidenz. Auf dem zweiten Niveau verstehen Kinder die Logik eines Experimentes und können Evidenzen im Zusammenhang mit einer zu prüfenden Hypothese sehen. Sie verstehen jedoch noch nicht, dass Experimente zu einer umfassenden Evaluation einer Theorie beitragen. Ebenso begreifen sie noch nicht, dass eine Hypothese durch Evidenz bestätigt oder widerlegt werden kann (z. B. Mayer, Sodian, Koerber, Schwippert, 2014). Zunächst verstehen Kinder Wissenschaft auf dem ersten, naiven Niveau als reines Sammeln von Fakten, dabei gelten Experimente nicht als Methode, um Hypothesen zu überprüfen; vielmehr werden sie „als Handlungspläne zur Produktion erwünschter Effekte“ (Sodian et al., 2006, S. 150) angesehen. Folglich sehen sie Hypothesen, die durch ein Experiment bestätigt werden, noch isoliert an und können das Ergebnis noch nicht nutzen, um zu einer generalisierenden Aussage zu kommen. Sie können Evidenzen noch nicht im Zusammenhang mit einer übergreifenden Theorie erfassen. Auf dem dritten Niveau erkennen Menschen den zyklischen Charakter wissenschaftlichen Arbeitens, um Theorien zu bilden, zu prüfen und gegebenenfalls zu revidieren (Mayer et al., 2014). Mittelpunkt ist dabei das theoriebasierte Generieren von Hypothesen sowie deren empirische Überprüfung. Die Autoren (Sodian et al. 2006) beschrieben dieses dritte Niveau folgendermaßen: „Personen auf diesem Niveau differenzieren zwischen Hypothesen und Theorien und verstehen Theorien als

forschungsleitend sowohl bei der Bildung von Hypothesen als auch bei der Interpretation von Daten.“ (Sodian et al., 2006, S. 150)

In einer Studie (Carey et al., 1989) antwortete über die Hälfte der Siebtklässler auf dem ersten Niveau. Sie konnten demzufolge keine Beziehung zwischen Theorie und Evidenz herstellen. In weiteren Studien äußerten Teilnehmer bei Interviews mit unterstützenden Bedingungen Antworten zwischen dem ersten und zweiten Niveau (Thoermer & Sodian, 2002). In den genannten Studien (Carey et al., 1989; Sodian et al., 2006) zeigte sich, dass auch Schülerinnen und Schüler der Sekundarstufe noch naive Vorstellungen besitzen; zum anderen wurde deutlich, dass Probanden durch Förderung zu differenzierenden Vorstellungen gelangen können (siehe auch Kapitel 3).

Sodian et al. (2006) untersuchten in ihrer Studie das Verständnis über Wissenschaft bei neun- bis zehnjährigen Schülerinnen und Schülern mit Fragen über Aufgaben, Ziele und Methoden der Arbeit eines Wissenschaftlers. Dazu wurden den Probanden Interviewfragen gestellt, die darauf ausgerichtet waren ihr Verständnis über Wissenschaft zu erfassen. Im Hinblick auf Ziele von Wissenschaftlern wurden beispielsweise folgende Fragen gestellt: Worum geht es in der Wissenschaft? Welche Ziele haben Wissenschaftler? Weitere Rubriken des Wissenschaftstests (ebenda) waren beispielsweise in Bezug auf „Fragen“, „Experimente“, „Hypothesen“ und „Theorien“ ausgerichtet. Zusätzlich wurde eine Experimentieraufgabe gestellt, bei der die Schülerinnen und Schüler überlegen sollten, wie ein Flugzeugingenieur den Verbrauch von Treibstoff verringern konnte (ebenda). Hierbei stand die Erfassung von Experimentierfähigkeiten (Variablenkontrollstrategie, Entwicklung eines kontrollierten Experimentes) im Vordergrund. Im Rahmen eines Prä-Post-Designs wurde eine Trainings- mit einer Kontrollgruppe verglichen.

Nach dem Interview und der Experimentieraufgabe im Prä-Test erhielten die Schülerinnen und Schüler neun Doppelstunden Unterricht zu den Themen *Licht und Schatten* sowie *Warum geht der Brotteig auf?* Zudem gab es eine Unterrichtseinheit zur Frage, „*Wie zuverlässig ist unsere Wahrnehmung?*“ Dabei erfuhren sie, dass es Phänomene gibt, die wir sehen, auch wenn sie nicht wirklich vorhanden sind (z.B. Spiegelungen im Auge). Zudem wurde auch anhand optischer Täuschungen für die Schülerinnen und Schüler deutlich, dass man sich leicht irreführen lassen kann. Anhand des Themas erarbeiteten sie zentrale Variablen, die die Qualität des Schattenbildes beeinflussten (wie z.B. Wahl der Lichtquelle, Abstand von Lichtquelle, Gegenstand und Projektionsfläche). Sie erkannten, dass einzelne Variablen kontrollierbar waren und entscheidenden Einfluss auf die Schärfe des Schattenbildes hatten. Diese Erkenntnis übertrugen die Probanden auf die Arbeit von Wissenschaftlern. Ein Wissenschaftler muss in der

Lage sein, möglichst genau wahrnehmen zu können und passend zu seiner Fragestellung Variablen zu wählen und diese zu kontrollieren (Sodian et al., 2006).

Die Schülerinnen und Schüler der Trainingsgruppe wurden sowohl mit fachlichen Inhalten als auch Reflexionen im Hinblick auf Theorie-Evidenz Bezug vertraut gemacht, während die Kontrollgruppe keine Reflexion auf einer metatheoretischen Ebene erhalten hat (Sodian et al., 2006). Ein Vergleich der beiden Gruppen im Post-Test zeigte, dass die Kinder der Kontrollgruppe Wissenschaft auf basaler Ebene als Sammeln von Fakten begreifen (Niveau 1), während Kinder der Trainingsgruppe Antworten äußern, die zwischen Niveau 1 und 2 liegen (ebenda). Niveau 2 und 3 wurden im Rahmen der kurzen Intervention nicht erreicht (Sodian et al., 2006). Dennoch wurde auch deutlich, dass Kinder der Trainingsgruppe bei der Experimentieraufgabe signifikant bessere Leistungen im Post-Test erreichten als Kinder der Kontrollgruppe (ebenda). Insgesamt verdeutlichten die Befunde, dass sich Verstehen über wissenschaftliche Erkenntnisprozesse und auch Experimentierfähigkeiten durch explizite Förderung steigern ließen.

Verschiedene Bestrebungen zeigten, dass junge Kinder in ihrer Koordination von Theorie und Evidenz unterstützt werden können (z.B. Hardy et al., 2006; Koerber et al., 2005). Dabei gelten Modelle, die Kinder selbst konstruieren und nutzen, als besonders hilfreich (Schwarz, Reiser, Acher, Kenyon & Fortus, 2012). Im Rahmen von „learning progressions“ (Schwarz et al. 2012, S. 113) werden vier Entwicklungsstufen dargestellt, die zeigen, wie Kinder mit angemessener Unterstützung inhaltspezifische Modelle konstruieren und sie diese zu eigenen Erklärungen heranziehen. Zunächst verfügen Kinder über keine oder wenig Erfahrungen im Umgang mit Modellen, beispielsweise stellten sie mit einer Zeichnung ein zuvor betrachtetes Phänomen möglichst realitätsgetreu dar (Niveau 1). Etwas differenzierter können sie Phänomene abbilden und dabei zur Verfügung stehende Informationen (wie beispielsweise durch ihre Lehrer oder durch Literatur) nutzen. Dabei beziehen sie die Evidenz des Phänomens sowie nicht sichtbare Mechanismen und Prozesse in ihr Modell mit ein. Sie sehen dieses als Mittel zur Kommunikation und nicht zur Unterstützung ihres Denkens (Niveau 2). Zunehmend konstruieren Kinder und Erwachsene mehrere Modelle und setzen diese zur Darstellung von Aspekten verknüpfter Phänomene ein (Niveau 3). Auf dem höchsten Niveau entwickeln und nutzen Menschen differenzierte Modelle verschiedener Domänen zur Unterstützung ihres Denkens. Mit Hilfe der Modelle generieren sie innovative Fragen im Hinblick auf Existenz und Verhalten von Phänomenen (Niveau 4). Bei diesem Ansatz steht die angemessene Förderung der Koordination von Theorie und Evidenz mit Repräsentationen, Vergleichen sowie Reflexionen im Vordergrund; dabei entwickeln Menschen mit zunehmend differenzierten Modellen abstrakte Formen naturwissenschaftlichen Denkens (Schwarz et al., 2012).

Bereits im Elementarbereich können einfache Formen zur Theorie-Evidenz-Koordination trainiert werden, dazu gehören beispielsweise zielgerichtetes Beobachten und Vergleichen sowie Vermuten und Prüfen (Steffensky et al., 2012). Die genannten Bestandteile des wissenschaftlichen Zyklus können spielerisch durch das Stellen einer Frage, das Generieren von Hypothesen oder Vermutungen, das Planen und Durchführen eines Experimentes etc. geübt werden. So lernen Kinder die einzelnen Schritte naturwissenschaftlicher Argumentationen kennen (Wilkening & Sodian, 2005). Sie werden mit altersangemessenen Begrifflichkeiten vertraut gemacht, um die entsprechenden Vorgänge zu beschreiben (Steffensky et al., 2012). Dabei scheint eine Differenzierung zwischen Konzepten und Begriffen bei Kindern im Elementarbereich nicht sinnvoll, da Vorschulkinder sich meist aus ihren Beobachtungen eigene Vorstellungen zurechtlegen (Steffensky et al., 2012) und zur Beschreibung noch keine adäquaten Begriffe verwenden. Die Koordination von Theorie und Evidenz kann im Gespräch innerhalb einer Gruppe oder durch einen Austausch zwischen Kind und Erwachsenem gefördert werden, da auf diese Weise Überlegungen ausgetauscht werden (French, 2004, Steffensky et al., 2012). Kinder im Elementarbereich verfügen bereits über grundlegende, bereichsübergreifende Fähigkeiten wie die Koordination von Theorie und Evidenz. (Koerber et al., 2005).

#### *4.2 Koordination von Theorie und Evidenz*

Der Prozess der Koordination von Theorie und Evidenz gilt als metakognitive, übergeordnete Fähigkeit (Kuhn et al., 2004) und stellt die zentrale Voraussetzung für wissenschaftliches Denken dar (Kuhn 1989; Kuhn & Pearsall, 2000; Bullock et al., 2009; Klahr, 2000; Zimmermann, 2007) und wird daher im folgenden Unterkapitel erläutert. Anhand von zwei Modellen (Klahr, 2000; Sodian & Mayer, 2013) wird der Zusammenhang von Theorie und Evidenz erläutert. Im folgenden Absatz wird zunächst das Modell von Klahr & Dunbar (1988) und Klahr (2000) beschrieben (siehe Abbildung 5).

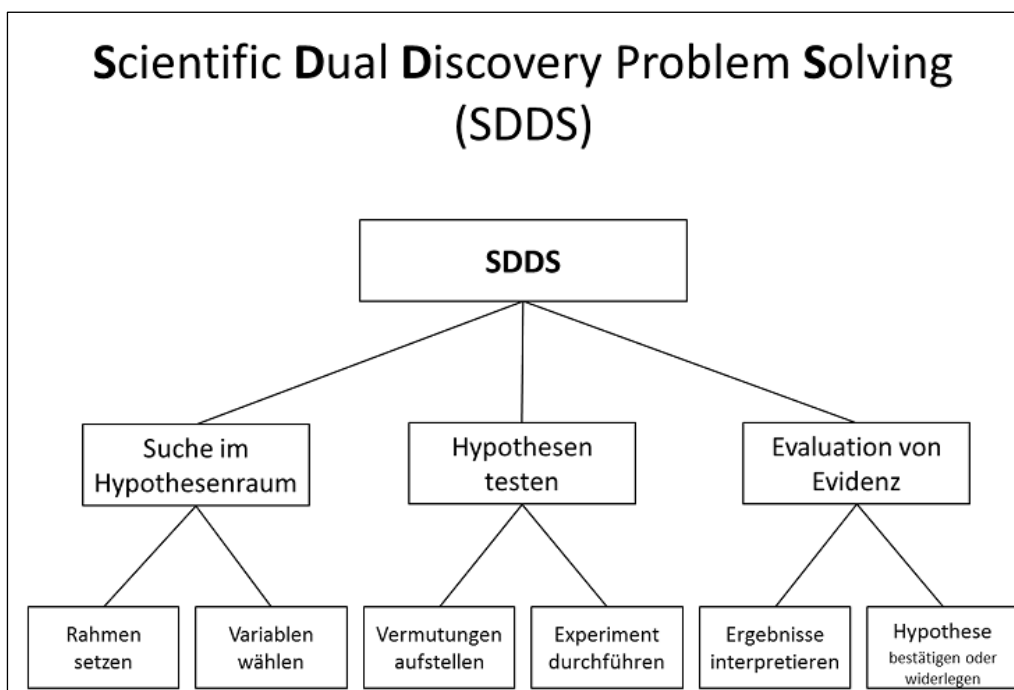


Abbildung 5. Scientific Dual Discovery Problem Solving (SDDS) (Klahr, 2000, S. 82).

Klahr & Dunbar (1988) und Klahr (2000) haben im Modell des *Scientific Dual Discovery Problem Solving* (SDDS) den wissenschaftlichen Prozess als Problemlöseprozess mit zwei Hauptbereichen, dem Hypothesenraum und dem Experimentierraum, dargestellt. Diese beiden Bereiche mit drei Hauptkomponenten beschreiben die Autoren in folgender Weise: Suche im Hypothesenraum, Hypothesen testen und Evaluation von Evidenz (siehe Abbildung 5). Alle drei Elemente, Hypothesengenerierung, Hypothesentestung sowie Evaluation von Evidenz sind im naturwissenschaftlichen Prozess eng miteinander verknüpft und bilden die Grundlage einer kohärenten theoriebasierten wissenschaftlichen Begründung (Klahr & Dunbar, 1988; Klahr, 2000).

Der wissenschaftliche Prozess wird mit der Suche nach Hypothesen eingeleitet und im Anschluss mithilfe eines Experimentes geprüft. Zentral für die Planung ist sowohl domänenübergreifendes als auch domänenspezifisches Wissen (Klahr, 2000), um theoriegeleitet einen plausiblen Rahmen zu setzen und geeignete Variablen zu wählen. Im zweiten Bereich, dem Hypothesen Testen, werden Vermutungen aufgestellt und im Rahmen eines Experimentes überprüft.

Das Aufstellen eines validen Experimentes ist in diesem Zusammenhang relevant, um eine gegebene Vermutung eindeutig bestätigen oder widerlegen zu können. Dafür müssen Kinder die Logik des Experimentierens verstehen und begreifen, dass immer nur eine Variable verändert werden kann und die anderen konstant gehalten werden müssen (Bullock & Ziegler,

1999). Die vermutlich entscheidende Variable sollte dabei besonders genau und im Hinblick auf ihren Effekt isoliert betrachtet werden (Kuhn & Phelps, 1982). Nach Inhelder & Piaget (1958) galt daher die Variablenkontrollstrategie als wichtigste Arbeitsweise beim wissenschaftlichen Experimentieren. Diese Strategie stellte eine domänenübergreifende Fähigkeit dar, um schlüssige Experimente aufzustellen und angemessene Schlussfolgerungen treffen zu können. In neueren Studien wurde diese Auffassung bestätigt (Chen & Klahr, 1999; Klahr et al., 2000; Kuhn, 2002). Nach Bullock & Ziegler (1999) stellt das Verständnis über die Logik eines Experimentes das „Kernstück naturwissenschaftlichen Arbeitens“ (Bullock & Ziegler, 1999, S. 51) dar.

Bei der Evaluation von Evidenz, dem dritten Bereich des Modells von Klahr (2000), gilt es, ein beobachtetes Ereignis als Beleg zu nutzen, um eine Hypothese zu bestätigen oder zu widerlegen (Klahr, 2000; Wilkening & Sodian, 2005). Eine Theorie wird bestätigt, wenn das beobachtete Ereignis, die Evidenz, mit der zuvor aufgestellten Theorie übereinstimmt. Die Theorie wird widerlegt, wenn die Evidenz nicht mit der Theorie übereinstimmt. Dann liegt Gegen-evidenz vor, da ein Beleg gefunden wurde, der der Theorie widerspricht (Wason, 1966). Das dargestellte Modell dient als Grundlage zum Verständnis des Zusammenhangs von Theorie und Evidenz.

Im folgenden Absatz wird ein weiteres Modell zur Koordination von Theorie und Evidenz im Zusammenhang mit einem experimentellen Design dargestellt; es wurde von Kuhn (2002) entwickelt und in einem aktuellen Beitrag aufgegriffen (Sodian & Mayer, 2013, siehe Abbildung 6). Darin wird das Wissenschaftsverständnis, wie zu Beginn des vierten Kapitels erläutert, als systematischer Einsatz von Strategien bezeichnet, um Hypothesen auf Grundlage von Evidenz zu prüfen und adäquate Schlussfolgerungen zu treffen (Sodian et al., 2006). Somit ist dieses grundlegend für ein Verständnis über den Zusammenhang von Theorie und Evidenz. Als bereichsübergreifendes Element spielt das Wissenschaftsverständnis beim Theorieverständnis sowie beim Prozess der Evaluation von Evidenz im Zusammenhang mit der aufgestellten Hypothese eine wichtige Rolle (Grygier, 2008; Sodian et al., 2006; Sodian & Mayer, 2013). In der dargestellten Abbildung 6 nach Sodian & Mayer (2013) steht das Wissenschaftsverständnis daher im Zentrum. Wie zu erkennen, besitzt die Genese wissenschaftlichen Wissens einen zyklischen Charakter (Sodian et al., 2006; Sodian & Mayer, 2013). Weiterhin veranschaulicht die Abbildung, dass die einzelnen Schritte aufeinander folgen und eng miteinander verknüpft sind (Driver et al. 1996; Sodian et al., 2006). Dabei kann eine Theorie auf Grundlage von Evidenz durch systematische Variation durch jeweils eine Variable im Rahmen eines experimentellen Designs bestätigt oder widerlegt werden.



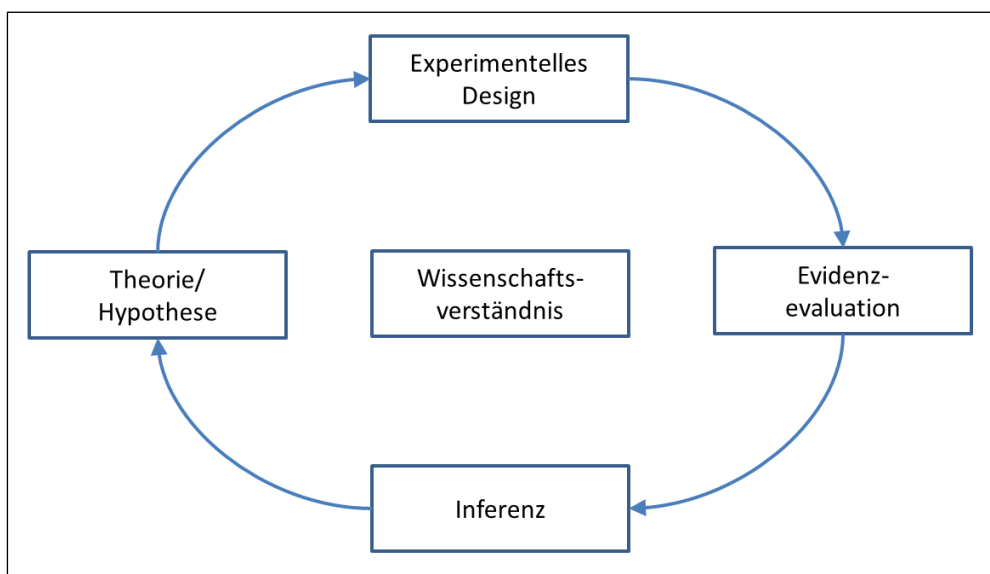


Abbildung 6. Die Relation von Theorie, Hypothese und Evidenz (Sodian & Mayer, 2013, S. 872).

Beide Modelle (Klahr, 2000; Sodian & Mayer, 2013) postulierten, dass eine angemessene Evidenzevaluation im Rahmen von genauen Beobachtungen die Voraussetzung für die Auswertung eines Experimentes in Bezug auf eine Hypothese darstellt. Die angemessene Bewertung von Beobachtungen ist grundlegend für eine folgerichtige Inferenz, da auf Basis der Evidenzevaluation eine Hypothese bestätigt oder widerlegt werden kann (Sodian & Mayer, 2013). Das übergreifende Wissenschaftsverständnis lenkt jeden einzelnen dieser Schritte, indem die Differenzierung zwischen Theorie und Evidenz gelingt und dadurch stimmige Schlussfolgerungen gezogen werden. Fähigkeiten von Vorschulkindern (z.B. Koerber, Sodian, Thoermer & Nett, 2005) sowie Grundschulkindern bei der Koordination von Theorie und Evidenz wurden empirisch überprüft (Bullock & Ziegler, 1999; Sodian & Mayer, 2013; Sodian, Zaitchik & Carey, 1991). Dennoch besteht weiterhin Forschungsbedarf, um die Verzahnung von Wissenschaftsverständnis als übergeordnete Kompetenz und der Fähigkeit zur Evaluation einer Evidenz im Zusammenhang mit einer spezifischen Theorie aufzuzeigen.

Im folgenden Absatz wird die Entwicklung der Fähigkeiten zur Koordination von Theorie und Evidenz erläutert, wobei die Merkmale der einzelnen Formen der Evidenz (bestätigend, widerlegend oder irrelevant) im Zusammenhang mit der Theorie sowie die dazugehörigen Schwierigkeitsstufen dargestellt werden.

Die Evidenz, mit der eine Hypothese bestätigt werden kann, gilt als einfachste Form, da Evidenz und Hypothese übereinstimmen. Bereits Kindern im Alter von vier Jahren gelingt der angemessene Umgang mit bestätigenden Evidenzen (Ruffman, Perner, Olson & Doherty, 1993). Kinder neigen dazu, Evidenzen zu bestätigen (Klahr & Chen, 2003; Kuhn & Pearsall,

2000; Zimmermann, 2007), da hier beide Aussagen zutreffen und auch miteinander übereinstimmen. Im Vergleich zum Umgang mit bestätigenden Evidenzen erscheint die Evaluation von Gegenevidenz oder irrelevanter Evidenz viel schwieriger.

Bei der Gegenevidenz müssen Kinder den Widerspruch zwischen der vorliegenden Beobachtung und der zuvor genannten Aussage in der Theorie erkennen und entsprechend schlussfolgern. Ist eine Evidenz irrelevant im Zusammenhang mit der zuvor aufgestellten Hypothese, besteht zwischen beiden Komponenten (Hypothese und Evidenz) keinerlei Verknüpfung. Dieser Fall scheint besonders schwierig, da Kinder, jünger als neun Jahre, fast automatisch davon ausgehen, dass selbstverständlich ein Zusammenhang zwischen Hypothese und Evidenz besteht. In diesem Fall schätzen sie die Evidenz als passend zur Theorie ein (Kuhn, 1989) bzw. lassen sich von ihrem Vorwissen leiten und ignorieren dabei den fehlenden Zusammenhang von Theorie und Evidenz (Zimmermann, 2007).

Insbesondere Kinder im Kindergarten- und im Grundschulalter haben Schwierigkeiten bei der Interpretation der beiden zuvor erläuterten Formen der Evidenz (Gegenevidenz bzw. irrelevante Evidenz); sie tendieren dazu, diese zu missinterpretieren (Chinn & Brewer, 1998; Klahr & Dunbar, 1988; Kuhn, 1989).

Zur Illustration wird ein Beispiel aus der vorliegenden Querschnittstudie für den Umgang mit irrelevanter Evidenz angeführt. Ist zum Beispiel die Hypothese „Dinge, die mit Luft gefüllt sind, springen“ gegeben, denken Kinder häufig mit einem Flummi (ohne Luft/springt) werde die Hypothese bestätigt. Dabei lassen sie die Voraussetzung zum Bestätigen oder Widerlegen der Evidenz außer Acht und konzentrieren sich bei ihrer Entscheidung nur auf das Sprungverhalten. Sie erkennen noch nicht, dass das Sprungverhalten des Flummis nicht entscheidend ist für die Evaluation der Evidenz im Zusammenhang mit der zuvor genannten Hypothese. Empirische Befunde belegen ein derartiges Antwortverhalten und zeigen, dass Kinder sowie manche Erwachsene dazu neigen, die Evidenz zu ignorieren (Chinn & Brewer, 1998) oder sie mit einer Theorie zu vermischen (Klahr & Dunbar, 1988; Schauble, 1996). Im nächsten Unterkapitel werden empirische Studien zum wissenschaftlichen Denken dargestellt.

#### *4.3 Bereichsübergreifende Fähigkeiten*

Zu Beginn der Entwicklung bauen Kinder ihre naiven Vorstellungen auf eigenen Erfahrungen auf (Ziegler; Hardy, 2015); diese sind anfangs nur aus Sicht des Kindes zu erklären. Empirische Befunde weisen darauf hin, dass Kinder schon früh kognitive Fähigkeiten entfalten, die grundlegend für die Koordination von Theorie und Evidenz sind (Brown, 1989; Sodian et

al., 2005). Dazu gehören bereichsübergreifende Kompetenzen wie analoges (Goswami, 2001), kategoriales sowie kausales Denken (Wilkening & Sodian, 2005).

Mithilfe des analogen Denkens können Kinder auf Grundlage von Ähnlichkeiten zwischen zwei Gegenständen Rückschlüsse ziehen und dadurch eine weitere Beziehung zwischen zwei anderen Gegenständen erklären. Kinder verfügen bereits in jungen Jahren über analoge Denkformen und sind dadurch in der Lage, bekannte Situationen auf neue zu übertragen. Mit diesen Denkformen können Vier- bis Fünfjährige bei einfachen Aufgaben logische Zusammenhänge erfassen (Wilkening & Sodian, 2005). Beispielsweise können Kinder in diesem Alter biologisches Wissen aufbauen, indem sie Gemeinsamkeiten bzw. Unterschiede zwischen Menschen und Tieren (z.B. Hunden) aufbauen (Goswami, 2001). Im Unterschied zu perzeptuell wahrnehmbaren Merkmalen werden strukturelle Merkmale als Basis für das Herstellen einer Analogie zwischen zwei Dingen frühestens von Vorschulkindern erkannt (Goswami, 2001; Holyak & Thagard, 1989). Diese Fähigkeit prägt sich mit Ausweitung des Weltwissens eines Kindes zunehmend aus. Analoges Denken ist wichtig, da Kinder durch Vergleiche zwischen unterschiedlichen Gegenständen lernen können, Argumentationsstränge aufzubauen (Holyak & Thagard, 1989).

In der kognitiven Entwicklung, die die Basis für die Koordination von Theorie und Evidenz darstellt, spielt die Bildung von Kategorien als Komponente bereichsübergreifenden Wissens eine zentrale Rolle. Durch Einteilen von Objekten in Kategorien können Kinder Vorstellungen über deren Eigenschaften und Merkmale gewinnen. Nach Untersuchungen von Gelman & Markmann (1986) und Goswami (2001) verfügen bereits Kinder im Alter von drei Jahren über Einsichten in die Bildung von Kategorien. Spontan und oft unbewusst nehmen Kinder äußerlich sichtbare Eigenschaften von Gegenständen wie beispielsweise Farbe, Größe oder Form wahr (Gelman & Markman, 1986; Goswami, 2001). Äußere Merkmale sind einfach zu erkennen, daher nutzen gerade junge Kinder sie vielfach bei der Kategorienbildung (Gelman & Markman, 1986). Perzeptuelle und strukturelle Kategorienbildung unterscheiden sich grundlegend, denn während die erstere nur durch äußere Merkmale (wie beispielsweise Farbe, Größe, Form) begründet ist, unterliegt letztere allgemein anerkannten Kriterien, z.B. physikalischen Gesetzen (Chi & Feltovich, 1981; Gelman & Markmann, 1986) oder strukturellen Merkmalen (Goswami, 2001). Beispielsweise erfassten bereits Säuglinge perzeptuelle Merkmale von Objekten, auf deren Grundlage sie Kategorien bilden. So wussten sie, dass Federn und Flügel zu Vögeln gehören und diese Merkmale bei allen Vögeln vorhanden sind. Daher konnten bereits Kinder unter einem Jahr grobe Kategorien von Gegenständen unserer Umwelt wie beispielsweise Bäume, Vögel, Blumen etc. unterscheiden (Goswami, 2001).

Weiterhin spielt das kausale Denken als bereichsübergreifende Komponente für die Entwicklung der Fähigkeit naturwissenschaftlichen Denkens eine wichtige Rolle. In einem Experiment wurden drei- und vierjährigen Kindern unvollständige Bildfolgen vorgestellt, die einen einfachen Ursache-Wirkungszusammenhang darstellten (Gelman, Bullock & Meck, 1980). In der Einführung wurde als Beispiel eine vollständige Bilderabfolge präsentiert, die die Abfolge eines Ursache-Wirkungszusammenhangs darstellte: eine vollständige Tasse (erstes Bild) - Hammer (zweites Bild) - kaputte Tasse (drittes Bild). Die drei Bilder einer Aufgabe gehörten immer gleichen Kategorien an: ursprüngliches Objekt-verursachendes Objekt, anderes verändertes Objekt. Bei den Testitems im Anschluss fehlte jeweils ein Bild. Daher wurden den Kindern zu jeder Kategorie drei Bilder präsentiert, aus denen sie ein Bild auswählen sollten.

Bei einem Item war auf dem ersten Bild das ursprüngliche Objekt, zum Beispiel eine Brille, bei dem zweiten Bild das verursachende Objekt, beispielsweise ein Eimer Wasser, zu sehen. Das dritte Bild mit dem veränderten Objekt fehlte. Hier sollten die Kinder aus drei Bildern (zerbrochene Brille, Brille mit Wassertropfen, Erdbeere mit Wassertropfen) das passende auswählen. In der Studie stellten Gelman und Kollegen (1980) fest, dass 92% der Dreijährigen richtig antworten konnten, wenn das verursachende Bild fehlte, 83%, wenn das letzte Bild nicht vorhanden war und 66% wenn das erste Bild fehlte (ebenda). Jeweils etwas schlechter waren die Lösungsraten, wenn die dargestellten Inhalte auf den Bildern in nicht logischer Reihenfolge gezeigt wurden. Insgesamt zeigten die Befunde, dass bereits Kinder im Alter von circa drei Jahren einfache Ursache-Wirkung-zusammenhänge begreifen und Aufgaben richtig lösen können.

Weitere empirische Befunde weisen darauf hin, dass bereits Vorschulkinder bei einfachen Aufgaben im Rahmen eines alltagsnahen Kontextes ein basales Verständnis zur Koordination von Theorie und Evidenz zeigen (Koerber et al., 2005; Ruffman et al., 1993). So gelang fünfjährigen Kindern beispielsweise die Differenzierung zwischen Evidenz und Theorie als unterschiedlichen Kategorien. Bei diesen Studien wurden die Aufgaben im Rahmen eines Kontextes angeboten, da empirischen Befunden zufolge Kinder Aufgaben mit einer inhaltlichen Verknüpfung leichter lösen konnten (Carey, 1989; Sodian, 1991). Gab es nur eine Variable, die für das Annehmen oder Widerlegen einer Hypothese verantwortlich war, dann konnten bereits vierjährige Kinder Evidenz und Theorie adäquat verknüpfen und den bestehenden Zusammenhang identifizieren. Diese Erkenntnisse stammten aus der *Theorie of Mind*-Forschung (Wimmer & Perner, 1983).

In der *Theorie of Mind*-Forschung wurden Kompetenzen zur Koordination von Theorie und Evidenz durch die „Maxi-Aufgabe“ (Perner & Lang, 1999) belegt. Ein Junge, Maxi, legte

eine Tafel Schokolade in einen grünen Schrank. Während der Junge sich auf dem Spielplatz aufhielt, legt die Mutter die Schokolade in den blauen Schrank. Die Frage an die Probanden lautete: „Wo wird Maxi die Schokolade suchen, wenn er vom Spielplatz zurückgekehrt ist?“ (Perner & Lang, 1999, S.338). Kinder unter vier Jahren antworteten meist, dass er sie im blauen Schrank suchen würde, wohingegen ältere Kinder sich in den Jungen hineinversetzen konnten und daher antworteten, dass er sie im grünen Schrank suchen würde (Perner & Lang, 1999). Auf Basis empirischer Studien ist davon auszugehen, dass Kinder bereits ab einem Alter von vier Jahren (Perner, 1991; Ruffman et al., 1993) Evidenzen in *False Belief Tasks* (Perner, 1991; Wimmer & Perner, 1983) richtig beurteilen können. Sie verstehen, dass Menschen Überzeugungen auf Grundlage der ihnen zur Verfügung stehenden Informationen äußern. Die erläuterten empirischen Befunde zeigten, dass Kinder im Alter von drei bis fünf Jahren bereits Fähigkeiten besitzen, sich in andere hineinzuversetzen (Perner, 1991; Ruffman et al., 1993; Wimmer & Perner, 1983) und auf dieser Grundlage Aussagen treffen. Voraussetzung hierfür ist, dass sie einen einfach verständlichen Kontext vorgestellt bekommen, zu dem sie sich äußern sollen.

In diesem Unterkapitel wurden Erkenntnisse aus der psychologischen Forschung zum Analogem Denken, zum Kausalen Denken und zur *Theory of Mind*-Forschung aufgezeigt, da diese relevante Kompetenzen im Vorfeld von wissenschaftlichem Denken darstellen und häufig als Vorläuferfähigkeiten bezeichnet werden (Brown, 1989; Sodian et al., 2005). Die dargestellten Kompetenzen junger Kinder sind relevant für die Koordination von Theorie und Evidenz.

#### 4.4 Empirische Studien zur Koordination von Theorie und Evidenz

Im Folgenden werden relevante empirische Studien zum wissenschaftlichen Denken mit Fokus auf Hypothesenprüfung, Experimentieren und Argumentieren vorgestellt. Die Probanden stammen aus dem Elementar-, Grundschul- sowie dem Sekundarschulbereich. Häufig handelt es sich dabei um Querschnittstudien, die wissenschaftliches Denken in unterschiedlichen Altersstufen erfassen. Es ist nötig, den Blick auf Studien mit älteren Kindern auszuweiten, da es nur sehr wenige Studien gibt, die wissenschaftliches Denken im Elementarbereich untersuchen (Gropen, Clark-Ciarelli, Hoisington, Stacy & Ehrlich, 2011, Leuchter et al., 2011; Steffensky & Lankes, 2011). Beim Untersuchungsgegenstand (siehe Tabelle 2) werden die unterschiedlichen Bereiche naturwissenschaftlichen Arbeitens dargestellt, innerhalb eines Bereiches sind die Studien nach dem Alter der Probanden in absteigender Reihenfolge geordnet.

Tabelle 2

*Studien zur Koordination von Theorie und Evidenz*

Ausrichtung der Studie	Autoren	Art der Studie	Alter/Anzahl der Probanden
<b>Hypothesenprüfung</b>			
	Inhelder & Piaget, 1958	Aufgaben	Schüler 11–12 Jahren
	Kuhn et al., 1988	Aufgaben	68 Schüler im Alter 9, 12, 15 Jahren
	Sodian et al., 1991	Aufgaben	48 Schüler im Alter 6,7,8 Jahren
	Ruffman et al., 1993	Aufgaben	32 Kinder im Alter von 4, 5 Jahren
	Kuhn & Pearsall, 2000	Aufgaben	45 Kinder im Alter 4, 6 Jahren; Erwachsene
	Koerber et al., 2005	Aufgaben	76 Kinder im Alter 4, 6 Jahren
	Sodian & Koerber, 2011	Aufgaben	107 Schüler im Alter 6, 7, 9, 11 Jahren; Erwachsene
<b>Experimentierfähigkeiten</b>			
	Klahr & Dunbar, 1988	Aufgaben	20 Erwachsene
	Schauble, 1996	Training	10 Schüler im Alter von 11, 12 Jahren
	Bullock et al., 1999	Training	143 Schüler im Alter von 9, 10, 11 Jahren
	Strand-Cary & Klahr, 2008	Training	72 Schüler im Alter von 9, 10, 11 Jahren
	Koerber et al., 2011	Aufgaben	73 Schüler im Alter von 7, 10 Jahren
	Gropen et al., 2011	Aufgaben	Kinder im Alter von 5 Jahren (keine Angabe über Anzahl)
	Steffensky et al., 2012	Training	245 Kinder im Alter von 5 Jahren
<b>Wissenschaftsverständnis /Experimentierfähigkeiten</b>			
	Chinn & Brewer, 2001	Aufgaben	168 Studenten im Alter von 19 bis 26 Jahren
	Toulmin, 2003	Aufgaben	keine Probanden

Ausrichtung der Studie	Autoren	Art der Studie	Alter/Anzahl der Probanden
	Mercer et al., 2004	Training	407 Schüler im Alter von 10 Jahren
	Tytler et al., 2005	Training	12 Kinder im Alter von 5 Jahren
	Sodian et al., 2006	Training	49 Schüler im Alter von 9 bis 10 Jahren
	Grygier, 2008	Training	53 Schüler im Alter von 9 bis 10 Jahren
	Aufschnaiter et al., 2008	Aufgaben	43 Schüler im Alter von 12 bis 13 Jahren

Am Beispiel einiger Studien werden die relevanten Rubriken wissenschaftlichen Arbeitens, Hypothesenprüfung durch Einschätzung der Evidenz, Experimentierfähigkeiten und Fähigkeiten im Argumentieren erläutert. Befunde aus früheren Studien zeigten, dass Kinder erst im Alter von elf Jahren in der Lage sind, Variablen systematisch zu kontrollieren und schlüssige von nicht schlüssigen Experimenten zu unterscheiden (Inhelder & Piaget, 1958; Kuhn et al., 1988).

Demgegenüber wurde in der Studie (Sodian, Zaitchik & Carey, 1991) die Fähigkeit überprüft, inwieweit sieben- bis achtjährige Kinder zwischen einem schlüssigen und einem nicht schlüssigen Experiment differenzieren können. So konnte anhand einer Gruppe von Erstklässlern sowie einer Gruppe von Zweitklässlern gezeigt werden, dass Kinder bereits ab einem Alter von sechs Jahren über wissenschaftliche Kompetenzen verfügen und basale Elemente zum Testen einer Hypothese verstehen und zwischen dem Testen einer Hypothese und dem Produzieren eines Effektes unterscheiden können.

Den Kindern wurde eine Geschichte von zwei Brüdern erzählt, die sich über die Größe einer Maus nicht einig sind. Sie wollten entweder durch das Prüfen einer Hypothese oder durch Füttern der Maus herausfinden, ob es sich um eine kleine oder große Maus handelte. Dazu standen ihnen zwei Schachteln mit unterschiedlich großen Öffnungen zur Verfügung. In jeweils eine Öffnung wurde ein Stück Käse gelegt. In der ersten Frage, der *Find out question* (Sodian et al., 1991, S. 757), prüften die Kinder mithilfe der Größe der Schachtelöffnung, ob die Maus groß oder klein ist. Nur bei einer kleinen Öffnung der Schachtel konnten die Kinder im Rahmen eines schlüssigen Tests die Größe der Maus feststellen, denn nur eine kleine Maus konnte den Käse essen, wenn er im Haus lag. Bei der zweiten Frage, der *Feed question* (ebenda, S.757), ging es darum, den gewünschten Effekt zu erzielen; die Maus sollte

unabhängig von ihrer Größe an den Käse gelangen können (ebenda). Nur wenn die Öffnung groß war, kam die Maus, egal ob sie dick oder dünn ist, in jedem Fall an den Käse. Während die erste Frage auf die Hypothesenprüfung abzielte, stand bei der zweiten Frage das Produzieren des Effektes im Fokus. Mehr als die Hälfte der sechs- bis siebenjährigen Kinder (Erstklässler) sowie 86% der sieben- bis achtjährigen Kinder (Zweitklässler) konnten zwischen beiden Aufgabenstellungen differenzieren; sie wählten die passende Größe der Öffnung und begründeten ihre Wahl entsprechend (Sodian & Mayer, 2013). Die Befunde zeigten, dass bereits junge Grundschul Kinder Grundlagen des Hypothesenprüfens verstehen können und Evidenz zur Evaluation einer Hypothese heranziehen (Sodian et al., 1991; Sodian et al., 2006; Sodian & Mayer, 2013). Ebenso wurde deutlich, dass junge Grundschul Kinder häufig schon zwischen einem schlüssigen und einem nicht schlüssigen Test unterscheiden können, da sie in der Studie je nach Aufgabenstellung die erforderliche Öffnung der Schachtel wählten (Sodian & Mayer, 2013). Befunde der Studie von Sodian et al. (1991) belegten die Kompetenz von Grundschulkindern bei der Differenzierung von Experimenten (schlüssig/nicht schlüssig).

In den folgenden beiden Absätzen werden die Studien von Koerber et al. (2005) und Ruffman et al. (1993) vorgestellt, da die Autoren im Vergleich zu den vorher erläuterten Studien wissenschaftliche Kompetenzen bei Kindergartenkindern untersuchten.

Das Experiment der Studie von Koerber et al. (2005) zeigte Erkenntnisse über die Fähigkeiten von fünfjährigen Kindern bei der Lösung kovariierender und nicht kovariierender Evidenzen und verdeutlichte zudem den Einfluss domänenspezifischen Wissens auf kausale Überzeugungen von Kindern. Dazu wurden Kindern im Alter von fünf und sechs Jahren zwei Kontexte vorgestellt. In Bezug auf die erste Geschichte hatten die Kinder erwartungsgemäß eine persönliche Überzeugung, da es sich um einen alltagsnahen Kontext handelte; zum Inhalt der zweiten haben die Kinder hingegen im Vorfeld keine persönliche Meinung. Zu jeder der beiden Geschichten wurden drei Aufgaben in einer vorgegebenen Reihenfolge gestellt (imperfekte Übereinstimmung - keine Übereinstimmung - perfekte Übereinstimmung). Bei den Aufgaben mit alltagsnahen Kontexten wurden die Evidenzen so getauscht, dass sie jeweils der persönlichen Überzeugung des Kindes widersprachen. Die eine Hälfte der Kinder bekam zunächst den alltagsnahen, dann den neutralen Kontext präsentiert, während die Reihenfolge bei der anderen Hälfte umgekehrt worden war.

Insgesamt zeigten sich erwartungskonforme Ergebnisse. Aufgaben, die vollständig mit der Theorie übereinstimmten, fielen den Kindern deutlich leichter als Aufgaben mit imperfekten Kovarianzen (Koerber et al., 2005). Weiterhin zeigten die Ergebnisse, dass Kinder in neutralen Kontexten die Evidenzen deutlich besser einschätzen können als bei alltagsnahen



Geschichten, wenn der präsentierte Inhalt den natürlichen Überzeugungen der Kinder widerspricht (Koerber et al., 2005). So konnten sie im neutralen Kontext bei perfekter Übereinstimmung zu 95%, bei nicht übereinstimmenden Evidenzen zu 80% und bei imperfekten Kovarianzen zu 65% richtig antworten. Im Vergleich dazu konnten deutlich weniger Kinder bei den alltagsnahen Kontexten (*Prior-Belief Context*) richtig antworten. Hier gaben sie bei perfekter Übereinstimmung zu 60%, bei nicht übereinstimmender Evidenz zu 53% und bei imperfekter Kovarianz zu 17% richtige Antworten (Koerber et al., 2005). Der Mc Nemar-Test in der Studie von Koerber et al. (2005) zeigte bei allen drei Aufgaben im Vergleich zwischen beiden Kontexten signifikante Unterschiede. Diese sind bei der imperfekten Kovarianz besonders deutlich, da die Probanden der Studie im neutralen Kontext signifikant besser antworteten als im alltagsnahen Kontext. Offensichtlich können sich Kinder in neutralen Kontexten leichter richtig äußern, da sie nicht von ihren eigenen Überzeugungen beeinflusst werden (Koerber et al., 2005).

Ähnliche Befunde zeigten sich in der Studie von Ruffman et al. (1993), bei der ebenso Fähigkeiten zur Koordination von Theorie und Evidenz bei vier- bis fünfjährigen Kindern im Fokus der Untersuchung standen. Dabei reduzierten die Wissenschaftler die Komplexität der Aufgabenstellung im Vergleich zu anderen Studien (Kuhn et al., 1989; Sodian et al., 1991) dadurch, dass sie nur eine zu bewertende Evidenz im Zusammenhang mit einer Hypothese verwendeten; ebenso mussten die Kinder in dieser Studie ihre Wahl einer Evidenz nicht begründen. Beide Maßnahmen minderten die kognitiven Anforderungen der Aufgabe (Ruffman et al., 1993). So konnten die Forscher nachweisen, dass es unter diesen Bedingungen bereits Kindern im Alter von vier bis fünf Jahren gelingen kann, einen Zusammenhang von Theorie und Evidenz zu erkennen. Das Erkennen einer Relation von Theorie und Evidenz wurde traditionell älteren Kindern zugetraut (Inhelder & Piaget, 1958; Kuhn, 1988), da ein metakonzeptuelles Verständnis als Voraussetzung für die Koordination von Theorie und Evidenz galt. Mit den erläuterten Vereinfachungen (Begründen von einer Evidenz, ohne Begründung) gelang es den Wissenschaftlern (Koerber et al., 2005; Ruffmann et al., 1993) zu zeigen, dass ein metakonzeptuelles Verständnis nicht unweigerlich die Voraussetzung für eine Koordination von Theorie und Evidenz darstellt.

Im Rahmen des ersten Experimentes wurden eine *Hypothesis-Evidence Task* (Ruffman et al., 1993, S. 1620) sowie eine *False-Belief Task* (Ruffman et al., 1993, S. 1620) eingesetzt. Bei der ersten Teilaufgabe wurde den Probanden zunächst eine Figur, Sally, vorgestellt, die kurz darauf das Szenario verlässt und so keine Informationen mehr erhalten kann. Den Probanden wurde zuerst ein Bild mit einem Jungen gezeigt, der grünes Essen zu sich nahm und dem mehrere Zähne fehlten. Anschließend wurden den Kindern weitere neun Bilder gezeigt, wobei

bei vier der Bilder (grünes Essen mit fehlenden Zähnen) zueinander passten, während die Jungen auf fünf weiteren Bildern rotes Essen zu sich nahmen und ein vollständiges Gebiss haben. Anschließend wurde den Probanden die Frage gestellt, welche Art von Essen das Ausfallen der Zähne verursache. Im zweiten Schritt wurden für die *Faked-Evidence Task* (Ruffman et al., 1993) die zusammengehörenden Merkmale ausgetauscht, so dass jetzt rote Nahrung mit schlechten Zähnen und grüne Nahrung mit guten Zähnen korrelierte. Diese falsche Evidenz rezipierte die Hauptfigur der Geschichte, Sally, die erst jetzt wieder zurückgekehrt war. Sie musste also zu einer gegenteiligen Annahme kommen, da sie auf Grundlage falscher Evidenz geurteilt hatte (ebenda, S.1621). Die Probanden wurden im Anschluss gefragt, was Sally wohl glaube, welches Essen schlechte Zähne verursachte. Sally musste aufgrund fehlerhafter Evidenz zu einer anderen Überzeugung kommen als die Probanden (ebenda). Im zweiten Teil des ersten Experimentes wurde eine *False-Belief Task* (Ruffman et al., 1993, S. 1621) durchgeführt.

Dabei wurden den Kindern zwei Puppen, Katy und John, präsentiert; John steckte einen Stift in einen Umschlag und ging anschließend auf den Spielplatz. Während der Abwesenheit von John legte Katy den Stift in einen zweiten Umschlag. Durch eine Kontrollfrage darüber, wo sich der Stift aktuell befand, wurde sichergestellt, dass die Kinder den Vorgang verstanden hatten. Nach einer Weile kehrte John vom Spielplatz zurück. Nun sind die Kinder gefragt worden, wo er den Stift suchte.

Das Experiment ergab, dass zwar alle Kinder bei der ersten Teilaufgabe die Korrelation von roter Nahrung-gute Zähne - grüne Nahrung-schlechte Zähne begriffen (ebenda), dass ihnen aber das Tauschen der Evidenzen im zweiten Teil der Aufgabe Schwierigkeiten bereitete. Die vierjährigen Kinder konnten zu 38% und die Fünfjährigen zu 82% die zweite Teilaufgabe mit den getauschten Evidenzen im Rahmen der *Hypothesis-Evidence Task* (ebenda) lösen. Das heißt, die Vierjährigen konnten nicht überzufällig richtig nach dem Tauschen der Evidenzen (grüne Nahrung- gute Zähne-rote Nahrung-schlechte Zähne) antworten; demgegenüber schafften es weitaus mehr Fünfjährige. Erwartungsgemäß zeigten sich bei der *False-Belief Task* (ebenda) deutlich höhere Lösungsraten bei den Vierjährigen im Vergleich zur *Faked-Evidence Task* (ebenda) im Rahmen der ersten Aufgabe.

Bei der *False-Belief Task* (ebenda) hatten alle Kinder verstanden, dass jemand, der den Wechsel des Stiftes von dem einen in den anderen Umschlag nicht mitbekommen hatte, eine falsche Überzeugung haben musste. Offensichtlich war die Frage der *False-Belief Task* (ebenda), bei der nur ein Umschlag als richtige Antwort in Frage kam, für die vierjährigen Kinder deutlich einfacher zu beantworten als die *Hypothesis-Evidence Task* (ebenda).

Insbesondere die große Anzahl von Bildern wurde in der Studie als Grund dafür genannt, dass nur 38% der vierjährigen Kinder die kausalen Zusammenhänge richtig benennen konnten (ebenda).

Die Befunde von Ruffman et al. (1993) und Koerber et al. (2005) zeigten, dass Kinder bereits im Alter von fünf Jahren über ein grundlegendes Verständnis in der Hypothese-Evidenz-Relation im Rahmen einfacher Kontexte verfügen. Dabei waren Evidenzen, die eine Hypothese bestätigten, erwartungsgemäß leichter zu bewerten als solche, die diese widerlegten. Hier hatten besonders junge Kinder mehr Schwierigkeiten, eine plausible Antwort zu finden, da sie zunächst davon ausgingen, dass eine vorgegebene Annahme richtig war. In einem neutralen Kontext, das heißt, wenn Kinder keine eigenen

Überzeugungen im Vorfeld hatten, fiel ihnen die Evidenzevaluation mit nicht kovariierenden Daten deutlich leichter als in alltagsnahen Kontexten (Koerber et al., 2005). Zusammenfassend kann auf Basis der empirischen Befunde festgehalten werden, dass bereits Vorschulkinder bestätigende sowie widerlegende Evidenzen erkennen können. Ebenso stellte sich heraus, dass bereits fünfjährige Kinder über ein grundlegendes Verständnis von Theorie und Evidenz verfügen, wenn sie begreifen, dass der Zugang zu Evidenz entscheidend für die Hypothesenbildung ist (Koerber et al., 2005).

In weiteren Untersuchungen wurden Experimentierfähigkeiten von Kindern im Alter vier bis fünf Jahren (Gropen et al., 2011) bis hin zu Schülern im Alter von neun bis elf Jahren (Bullock & Ziegler, 1999) untersucht.

So belegten empirische Befunde der *LOGIK*-Studie (Bullock & Ziegler, 1999), dass es mehr als einem Drittel der Drittklässler, über 60% der Viert- und Fünftklässler sowie nahezu 80% der Sechstklässler gelang, zwischen einem schlüssigen Experiment und einem nicht schlüssigen Experiment zu unterscheiden. Von den Kindern, die einen kontrollierten Test wählten, konnten mehr als die Hälfte der Viertklässler, 80% der Fünftklässler und fast alle Sechstklässler ihre Wahl begründen (Bullock & Ziegler, 1999). Gleichzeitig stellten nur wenige der Probanden selbst einen kontrollierten Test auf (Bullock & Ziegler, 1999). In der Studie wurde eine große Diskrepanz zwischen der Wahl eines kontrollierten Tests sowie dem Aufstellen eines Tests deutlich.

Die Befunde (Bullock & Ziegler, 1999) deuteten darauf hin, dass bereits Grundschulkinder im Alter von neun bis zehn Jahren über ein basales logisches Verständnis verfügen, was ein kontrolliertes Experiment ausmacht; denn sie konnten zwischen einem schlüssigen und einem nicht schlüssigen Experiment unterscheiden. Sie verstanden, dass jeweils nur eine Variable

pro Testdurchlauf verändert werden darf. Folglich begriffen sie, dass das Testen einer Hypothese einen kritischen Vergleich von mindestens zwei Testdurchläufen beinhaltet.

Weitere empirische Befunde (Strand-Cary & Klahr, 2008) belegten, dass Grundschulkinder durch explizite Instruktionsphasen im Rahmen von vier Sitzungen die Logik der Variablenkontrolle verstehen können. Die Befunde der Studie zeigten einerseits, dass bereits Kinder im Grundschulalter über Experimentierstrategien verfügen, andererseits aber auch, dass sie noch nicht die kognitiven Fähigkeiten besitzen, um kontrollierte Experimente zu entwickeln (Klahr, 2000; Schauble, 1996).

Kompetenzen, die beim Experimentieren benötigt werden, beispielsweise das Differenzieren zwischen zwei Experimenten (schlüssig - nicht schlüssig) und das Entwerfen eines kontrollierten Experiments, können durch verschiedene Maßnahmen gefördert werden. So verbesserte sich die Fähigkeit der Variablenkontrollstrategie durch eine gezielte Förderung (Bullock & Ziegler, 1999; Schauble, 1996; Strand-Cary, 2008). Dabei differierten die Auffassungen in der Forschung über die Art der Vermittlung im Hinblick auf die Lerneffekte. Das Verständnis von Kindern konnte durch direkte Instruktion kurzfristig geschult werden (Klahr & Nigam, 2004; Kuhn & Dean, 2004), wobei signifikante Effekte nach einer kurzen Intervention von zwei Tagen erzielt wurden. Demgegenüber stehen Studien, die darauf verweisen, dass instruktive Maßnahmen eher kurzfristig messbare Effekte nach sich zogen (Siraj-Blatchford & Moriarty, 2004).

In anderen Studien (Breddermann, 1983; Schauble, 1996) wurde das Konzept des entdeckenden Lernens favorisiert; hierbei entwickelten Kinder während der Experimentierphasen eigenständig Ideen und probierten diese im Rahmen von Experimenten aus. Befürworter des entdeckenden Lernens sehen einen Vorteil in dieser Form der Vermittlung, da Kinder so auf lange Sicht leichter Transferleistungen erbringen können (Breddermann, 1983; Schauble, 1996). Kinder, die mit direkter Instruktion unterrichtet wurden, konnten dafür möglicherweise besser mit Inkonsistenzen bzw. Fehlern bei Ergebnissen umgehen (Mayer, 2004). In der Forschung wurden die Effekte untersucht, die die erläuterten Maßnahmen zur Förderung (Instruktion im Gegensatz zu entdeckendem Lernen) im Hinblick auf den Lernzuwachs bei Kindern in Abhängigkeit vom geförderten Zeitraum haben (Klahr & Nigam, 2004; Schauble, 1996).

Einer neueren Studie zufolge konnten bereits Lerneffekte bei kurzer Intervention erreicht werden (Steffensky & Lankes, 2012). In dieser Studie bestand die Förderung aus drei Einheiten von jeweils 90 Minuten, die innerhalb eines Zeitraums von vier Monaten stattfanden. Getestet wurden 245 Kinder im Alter von durchschnittlich fünf Jahren. Dabei wurden insgesamt fünf Gruppen, drei Experimentalgruppen sowie zwei Kontrollgruppen, getestet. Im Rahmen

eines Prä-Post-Designs wurden die Gruppen in Bezug auf ihre naturwissenschaftlichen Kompetenzen miteinander verglichen. Die eine Experimentalgruppe führte Experimente durch, die im Rahmen einer Reflexion besprochen wurden. Nach der Experimentierphase konnten die Kinder über ihre Beobachtungen im Zusammenhang mit eigenen Alltagserlebnissen sprechen (Steffensky, 2012). Bei den anderen beiden Experimentalgruppen wurde jeweils nur eine Form der Unterstützung (Gespräch über Alltagserlebnisse-Experimente) angeboten. Den größten Lernzuwachs hatten die Kinder der Gruppen, in denen beide Unterstützungsmaßnahmen (Gespräch über Alltagserlebnisse-Experimente) verknüpft waren (ebenda). Ähnliche Befunde wurden auch in der Studie von Klahr & Nigam (2004) berichtet.

Die Befunde belegten, dass Experimente gerade dann für junge Kinder besonders lern-effektiv sind, wenn diese mit deren Alltagserlebnissen verknüpft werden können (Steffensky et al., 2012). Dabei schien die Nachhaltigkeit höher, wenn die Kinder ihre Beobachtungen selbst in Worte fassen, die im Anschluss mit der Bezugsperson gemeinsam reflektiert wurden (Steffensky & Lankes, 2012). Durch das Formulieren von Beobachtungen konnten Kinder explizites Wissen aufbauen, das sich an ihr Vorwissen angeschlossen hat. (Saalbach, 2013). Das Reflektieren im Zusammenhang mit einem Experiment scheint empirischen Befunden zufolge sowohl zum Aufbau naturwissenschaftlichen Wissens in spezifischen Themenbereichen (Mayer, 2004) als auch zur Bildung wissenschaftlichen Wissens (Butts, 1994; Hofstein & Lunetta, 2004) zentral.

Aktivitäten wie das Reflektieren über ein Experiment sowie das Formulieren von Begründungen auf Grundlage von Beobachtungen stellen hohe Anforderungen, da hierfür eine angemessene Relation zwischen den einzelnen Facetten wissenschaftlichen Wissens, Theorie-Evidenz-Koordination und Argumentation, hergestellt werden muss (Lawson, 2009). Weitere Studien machten deutlich, dass Kinder im Laufe der Entwicklung stichhaltiger argumentieren können, weil sie sukzessive ein besseres Theorie-Evidenz-Koordination aufbauen und strukturierte Formen des Denkens entwickeln (Tytler & Peterson, 2004; Tröbst et al., 2012b).

Mit mikrogenetischen Studien wurde nachgewiesen, dass Kinder zur Lösung von Problemen verschiedene Strategien gleichzeitig nutzen (Siegler & Crowley, 1991; Siegler & Stern, 1998). Dabei entwickelten sie zunehmend angemessenere Lösungsmöglichkeiten. Befunde zeigten, analog zur Koordination von Theorie-Evidenz, dass Kindern eine Argumentation leichter fiel, wenn Theorie und Evidenz in einem plausiblen Zusammenhang stehen (Koslowski, 2008).

Forschungen zum Argumentieren stammten in der Regel aus dem schulischen Kontext (Furtak, Hardy, Beinbrech, Shavelson & Shemwell, 2010); hier hat sich gezeigt, dass die

Förderung schlüssiger Argumente bereits im Primarschulbereich wichtig ist. (Tytler & Peterson, 2005). In einigen Studien konnte gezeigt werden, wie Lehrer Schüler dabei unterstützen können, einen logischen Argumentationsstrang aufzubauen (Mc Neill et al., 2006; Harlen & Qualter, 2009). Beispielsweise konnten ein Erwachsener und ein Kind-mehrere Kinder im Dialog klären, ob eine beobachtete Evidenz die eingangs aufgestellte Hypothese bestätigte. Sofern dies zutraf, konnte ein weiteres Experiment, das die Hypothese bekräftigte, entwickelt werden (Harlen & Qualter, 2009). Im anderen Fall, wenn vorliegende Evidenz nicht mit der Hypothese übereinstimmte, galt die Hypothese als widerlegt. In der Weiterführung konnte eine Hypothese gefunden werden, die mit der Evidenz übereinstimmte. In beiden Fällen, Bestätigung bzw. Widerlegung der Hypothese, begründeten Kinder im Dialog mit dem Erwachsenen den Sachverhalt (Harlen & Qualter, 2009).

Nach von Aufschnaiter et al. (2008) sind Qualität einer Argumentation und wissenschaftliches Wissen eng miteinander verknüpft; je fundierter ein Schüler oder eine Schülerin über allgemeines wissenschaftliches Wissen verfügt, desto besser kann er bzw. sie argumentieren.

Differenziert werden, ähnlich den Stufen bei Piaget, vier Stufen wissenschaftlichen Verstehens, die beim konkreten Objekt beginnen und je Stufe einen höheren Abstraktionsgrad annehmen. So konzentrieren sich Schüler in ihrer Begründung auf Stufe eins ausschließlich auf ein konkretes Objekt; auf Stufe vier verfügen Schüler sowohl über ein umfassendes konzeptuelles, inhaltspezifisches als auch ein bereichsübergreifendes Verständnis, daher können sie Begründungen mit hohem Abstraktionsgrad abgeben. Die empirischen Befunde der Studie zeigten, dass zwölf- bis dreizehnjährige Schüler die höchste Stufe (Stufe vier) noch nicht erreichten (von Aufschnaiter et al., 2008).

Im Gegensatz dazu konnten Hardy, Kloetzer et al., (2010a) in ihrer Studie mit neun- und zehnjährigen Kindern zeigen, dass Kinder mit geeigneten Formen der Unterstützung signifikant bessere Begründungen abgeben konnten als ohne Hilfestellung. Die Kinder wurden in vier neunzigminütigen Unterrichtseinheiten zum Thema *Schwimmen und Sinken* unterrichtet; dabei wurden zwei Stichproben (neunjährige-zehnjährige Kinder) untersucht, die jeweils aus vier Klassen bestanden. Jeweils zwei Klassen wurden mit hoher instruktionaler Unterstützung und zwei mit niedriger Unterstützung gefördert. Die Kinder der Gruppe mit umfangreicher instruktionaler Unterstützung, wurden über Materialeigenschaften sowie Dichte unterrichtet. Dabei zeigte sich, dass bei den zehnjährigen Schülern der Gruppe mit hoher instruktionaler Unterstützung 10% der Schüler die höchste Stufe beim Begründen in der Abschlusseinheit erreichen. Im Vergleich dazu waren es zu Beginn der Untersuchung nur 2% der Schüler.

In einer weiteren Studie von Brown et al. (2010) wurde ebenso ein System zur Einschätzung der Antworten im Sinne einer konzeptuellen Angemessenheit entwickelt. Brown et al. (2010) hatten in ihrem Modell zur Beschreibung des Bewertungsprozesses einen Zyklus entworfen. Mit ihrem *cycle of assessment* (Brown et al., 2010, S.131), zeigten sie, dass Fragestellung, Aufgabe, Beobachtung sowie die Kodierung der Aussagen in direktem Zusammenhang miteinander stehen. Das stufenförmige Schema reichte von unproduktiven über produktive Misskonzepte bis hin zu evidenzbasierten Konzepten. Mit dem Modell konnten Argumente von Kindern differenziert eingeschätzt werden. Selbst ältere Schüler bzw. Studenten nannten in ihren Argumentationen oft unzureichende bzw. unpassende Evidenzen zum Beleg ihrer Schlussfolgerungen. Sinnvoll erscheint es daher, wissenschaftliches Verständnis in verschiedene Stufen einzuteilen (z.B. Hardy et al., 2010a; Koerber et al., 2005; Kuhn & Pearsall, 2000), um so auch frühe Entwicklungen bei Kindern einschätzen zu können. Auf dieser Grundlage können Kinder auf unterschiedlichen Verständnisniveaus eine angemessene Förderung erhalten und lernen, logische Argumentationen auf Grundlage von Theorie und Evidenz aufzubauen.

Im folgenden Absatz wird eine Studie zur Argumentationsfähigkeit von Vorschulkindern beschrieben. Tytler & Peterson (2005) untersuchten zwölf Kinder, die bei Beginn der Längsschnittstudie fünf Jahren alt waren; sie wurden über einen Zeitraum von fünf Jahren zu acht Messzeitpunkten (ein bis zwei pro Jahr) begleitet. Tytler & Peterson untersuchten den Zusammenhang von konzeptuellem Verstehen und epistemologischen Begründungen. Dazu thematisierten sie in ihren Interviews verschiedene Themen wissenschaftlichen Verständnisses, wie beispielsweise, was ein faires Experiment ausmache oder welche Komponenten zum offenen Experimentieren in Klassen relevant waren. Dabei fokussierten Tytler & Peterson auf vier Entwicklungsstufen kindlicher Begründungen. Sie bildeten auf Grundlage der Kinderaussagen vier Stufen des Verständnisses. Diese reichten von einem einfachen Verständnis, geprägt von sehr wenig konzeptuellem Wissen, bis hin zu einer strategischen Suche nach Evidenz, um Hypothesen zu verfeinern und Aussagen mit Möglichkeiten zur Generalisierung zu finden. Kindern mit einem Verständnis, das auf der höchsten Stufe angesiedelt war, gelang es durch geeignete Formen der Überprüfung konfundierter Variablen zu identifizieren. Auf dieser Stufe konnten Probanden zwischen Theorie und Evidenz differenzieren und entsprechend der Fragestellung plausible und wissenschaftlich begründete Argumente nennen. In der Studie von Tytler & Peterson (2005) wurden Kinder im Abstand von jeweils einem Jahr zu einem bestimmten Thema befragt (z.B. warum kann ein Flugzeug fliegen). Ein Mädchen von sechs Jahren konnte bereits beim zweiten Erhebungszeitpunkt ein kontrolliertes Experiment entwerfen und eine plausible

Begründung nennen, weil sie im Hinblick auf das konkrete Problem förderliche Hinweise erhielt. Dadurch konnte sie die relevanten Variablen herausfinden (Tytler & Peterson, 2005). Insgesamt zeigten die Befunde, dass Kinder mit wachsendem domänenspezifischem Wissen ihre Fähigkeiten beim Argumentieren steigern können (ebenda).

Die Stufen in der Studie von Tytler & Peterson (2005) können mit den zuvor beschriebenen Stufen wissenschaftlichen Verstehens verglichen werden (vgl. von Von Aufschnaiter et al., 2008, Hardy et al., 2006; Hardy et al., 2010a), da die genannten Studien die Entwicklung wissenschaftlichen Verstehens thematisieren.

#### 4.5 Zusammenfassung und Fazit

In diesem Kapitel wurde die Entwicklung der Theorie-Evidenz-Koordination bei Kindern von vier bis zwölf Jahren mit ihren beiden Facetten wissenschaftliches Denken, inhaltliches Wissen und bereichsübergreifendes Wissen dargestellt (Kapitel 4.1). Verschiedene Kompetenzmodelle (Hardy et al., 2010a; Schwarz et al., 2012, Sodian et al., 2006) wurden vorgestellt, um die Entwicklung der Theorie-Evidenz-Koordination genau beschreiben zu können. Zur Erklärung des wissenschaftlichen Prozesses wurden zwei Modelle erläutert: *Scientific Dual discovery Problem Solving*-Modell (Klahr, 2000) und das Modell zum Wissenschaftsverständnis (Sodian & Mayer, 2013) (Kapitel 4.2). Weiterhin wurden bereichsübergreifende Fähigkeiten, wie beispielsweise das analoge bzw. kausale Denken aufgegriffen, da diese Denkformen grundlegend für die Entwicklung der Fähigkeit zur Koordination von Theorie und Evidenz sind (Kapitel 4.3). In diesem Zusammenhang wurde Bezug zur *Theory of Mind*-Forschung (z.B. Perner, 1991) genommen, da aus diesem Bereich der psychologischen Forschung die Erkenntnis über bereichsübergreifende Fähigkeiten als Vorläuferfähigkeiten für die Koordination von Theorie und Evidenz hervorgegangen ist. Anschließend wurden empirische Studien zur Koordination von Theorie und Evidenz aufgeführt, um die Möglichkeiten der Förderung wissenschaftlichen Verständnisses im Elementar- und Primarbereich aufzuzeigen; Dazu gehören die Hypothesenprüfung (z.B. Sodian et al., 1991), das Experimentieren (z.B. Steffensky et al., 2012) sowie das Argumentieren (z.B. von Aufschnaiter et al., 2008).

Zusammenfassend kann festgehalten werden, dass die Fähigkeit zum wissenschaftlichen Denken mit der kognitiven Entwicklung eines Menschen zunimmt (Kuhn, 1988; Kuhn & Pearsall, 2000; Koerber et al. 2005). Dennoch zeigten empirische Belege (z.B. Sodian et al., 1991; Koerber et al., 2005), dass die Koordination von Theorie und Evidenz sowie die Revision eigener Theorien durch Evidenz bereits bei jungen Kindern mit einfachen Aufgaben möglich



ist (Kuhn, 2010). Ebenso zeigen empirische Belege (Sodian et al., 2006; Steffensky et al., 2011), dass sich Fähigkeiten zum wissenschaftlichen Denken durch geeignete Fördermaßnahmen steigern lassen. Im nächsten Kapitel wird die Theorie zum Schlussfolgern ausgeführt, für die eine Theorie-Evidenz-Koordination grundlegend ist.

## 5 Die Schlussfolgerungsfähigkeit in der kognitiven Entwicklung

Im folgenden Kapitel wird auf die Entwicklung der Fähigkeit zum Schlussfolgern als Teil wissenschaftlichen Denkens eingegangen. Schlussfolgerungen enthalten ein konditionales Satzgefüge und verdeutlichen einen Zusammenhang von Möglichkeiten und Folgen (Gauffroy & Barrouillet, 2011). Sie sind gekennzeichnet durch die Worte „wenn [...], .dann [...]“. Ein Beispiel für eine konditionale Aussage ist „Wenn es regnet, wird die Erde nass.“ (Tröbst, Hardy & Möller, 2011, S.9).

Logische Schlussfolgerungen sind nur möglich, wenn die Koordination von Theorie und Evidenz gelingt. Adäquate Schlüsse aus Wenn-Dann-Beziehungen zu ziehen, stellt eine zentrale Fähigkeit im Rahmen der kognitiven Entwicklung dar. Wenn Schlussfolgerungen im Rahmen von alltäglichen Handlungen erfolgen, dann verlaufen sie intuitiv und sind pragmatisch geprägt (Markovits, 2006). Weiterhin gibt es Schlussfolgerungen bei kognitiven Prozessen, die logische Denkprozesse erfordern und analytisch geformt sind (Markovits, 2006, Gauffroy & Barrouillet, 2009). Die vorliegende Arbeit konzentriert sich auf den zuletzt genannten Bereich der analytischen Schlussfolgerungsprozesse.

Die Fähigkeit zu diesen Denkprozessen bildet sich zunehmend im Zuge der kognitiven Entwicklung vom Kind über den Jugendlichen bis hin zum Erwachsenen heraus; dennoch fällt auch erwachsenen Personen das korrekte Schlussfolgern bei komplexen Zusammenhängen schwer (Wason, 1966, Johnson-Laird & Byrne, 2002). Folgerichtige Schlussfolgerungen innerhalb eines wissenschaftlichen Prozesses sind nur möglich auf Basis einer stimmigen Koordination von Theorie und Evidenz; Schlussfolgerungen prägen die Abschlussphase eines wissenschaftlichen Zyklus (z.B. Hardy et al., 2006; Mayer & Sodian, 2013). Auf Basis von angemessenen Schlussfolgerungen stellen Menschen eine Argumentationskette auf (Brown, Nagashima, Fu, Timms & Wilson, 2010). Diese Prozesse sind daher ein wichtiger Bestandteil logischen Denkens und spielen eine Schlüsselrolle beim wissenschaftlichen Denken (Markovits, 2006). Schlussfolgerungen bilden sowohl für das Ableiten von Hypothesen als auch für das Planen von eindeutigen, wissenschaftlich begründeten Experimenten die Basis und sind analytisch geprägt (Markovits, 2006).

Der Prozess der Evidenznutzung wird demzufolge mit den Prinzipien des Schlussfolgerns in Zusammenhang gebracht. Ähnlich wie bei der Koordination von Theorie und Evidenz werden bei Schlussfolgerungsprozessen ausgehend von Prämissen Schlussfolgerungen unter der Voraussetzung gezogen, dass konditionale Aussagen für Antezedenzien [p] und Konsequenzen [q] in der Form wenn p, dann q korrekt sind (Barrouillet & Lecas, 2008; Johnson-Laird & Byrne, 2002; Markovits, 2006). Konditionale Schlussfolgerungen gelten daher als logisch richtige Schlüsse.

Formale Regeln zum Schlussfolgern ermöglichen die Darstellung logisch folgerichtiger Schlussfolgerungen. Dabei werden vier Schlussweisen unterschieden: Modus Ponens, Modus Tollens, Ablehnung des Antezedens, Affirmation der Konsequenz. Diese werden in der nachfolgenden Abbildung (Abbildung 7; Barrouillet, Grosset & Lecas, 2000) mit [p] stellvertretend für die Prämisse und [q] für die Konsequenz dargestellt.

<p><b>Modus Ponens</b></p> <ul style="list-style-type: none"> <li>• wenn p dann q</li> <li>• <math>p \rightarrow q</math></li> </ul>
<p><b>Modus Tollens</b></p> <ul style="list-style-type: none"> <li>• wenn p dann q</li> <li>• <math>\neg q</math> daher <math>\neg p</math></li> </ul>
<p><b>Ablehnung des Antezedens</b></p> <ul style="list-style-type: none"> <li>• wenn p dann q</li> <li>• <math>\neg p</math> daher <math>\neg q</math></li> </ul>
<p><b>Affirmation der Konsequenz</b></p> <ul style="list-style-type: none"> <li>• wenn p dann q</li> <li>• q daher p</li> </ul> <p><i>Anmerkung.</i> p = Antezedens, q = Konsequenz,  <math>\neg</math> = Negation.</p>

Abbildung 7. Vier Schlussweisen (Barrouillet et al., 2000, S. 240).

Nur die ersten beiden Schlussweisen führen zu eindeutigen Schlussfolgerungen; die anderen beiden bringen indeterminierte bzw. nicht schlüssige Folgerungen hervor (Johnson-Laird & Byrne, 2002; Markovits, 2006). Die Erkenntnis, dass aus einer gegebenen Prämisse ohne Konsequenz kein eindeutiger Schluss gezogen werden kann, ist schwierig; daher produzieren besonders junge Kinder bei Aufgaben dieser Art deutlich weniger richtige Antworten (Daniel & Klaczynski, 2006; Evans, 2003; Kempert & Hardy, 2012). Aus der konditionalen Aussage

„wenn p dann q“, ergeben sich vier mögliche Belegungen von Antezedens und Konsequenz, die folgendermaßen dargestellt werden können (Johnson-Laird & Byrne, 2002):

[p q]; [p¬q]; [¬p q]; [¬p¬q].

Betrachtet man nur zwei Fälle, [p q] [¬p¬q], dann stellt das Vorhandensein von [p] eine notwendige Bedingung für das Auftreten von [q] dar (Johnson-Laird & Byrne, 2002). Daher wären die vier Schlussweisen im Zusammenhang mit einer bikonditionalen Interpretation „q genau dann, wenn p“ valide (Johnson-Laird & Byrne, 2002). Bei der konditionalen Aussage „wenn p, dann q“ sind drei Kombinationen von Antezedens und Konsequenz für eine Belegung mit „wahr“ in der Wahrheitstabelle möglich:

[p q]; [¬p q]; [¬p¬q]

Die drei genannten Verknüpfungen stellen nach der logischen Wahrheitstabelle (siehe Tabelle 3, Barrouillet, Gauffroy & Lecas, 2008) wahre Aussagen dar.

In der Wahrheitstabelle (Tabelle 3) sind für jede Behauptung entsprechend der Prämissen Belegungen zugeordnet, die wahr oder falsch sein können (Johnson-Laird & Byrne, 2002; Barrouillet et al., 2008). Die unteren beiden Fälle werden als indeterminiert bezeichnet, da aufgrund einer falschen Prämisse [¬p] keine schlüssige Aussage [q] im Zusammenhang mit der konditionalen Verknüpfung „wenn p, dann q“ getroffen werden kann.

Tabelle 3  
*Wahrheitstabelle*

[p]	[q]	[p] impliziert [q]	Fehlerhafte Angaben
wahr	wahr	wahr	wahr
wahr	falsch	falsch	falsch
falsch	wahr	wahr	indeterminiert
falsch	falsch	wahr	indeterminiert

### 5.1 Theorie der Mentalen Modelle beim schlussfolgernden Denken

Zur Erklärung menschlicher Denkweisen werden in der Kognitionspsychologie die sogenannten mentalen Modelle herangezogen. Diese können sich aus der Wahrnehmung von

Sachverhalten sowie der Vorstellung, dem Weltwissen oder dem individuellen Verständnis ergeben (Johnson-Laird & Byrne, 2002); sie eignet sich die Entwicklung kognitiver Prozesse im Laufe der Entwicklung vom Kind zum Erwachsenen darzustellen (Markovits & Barrouillet, 2002; Gauffroy & Barrouillet, 2011).

Nach dieser Theorie verfügen Menschen mit ausgeprägten mentalen Modellen über die gleichen logischen Strukturen wie die, die bei formaler Logik eingesetzt werden (Barrouillet, Grosset & Lecas, 2000; Markovits, 2006). Die Theorie der Mentalen Modelle (*mental models theory*) (Johnson-Laird & Byrne, 1991; 2002, Gauffroy & Barrouillet, 2009) geht davon aus, dass Menschen auf Basis psychologischer Prozesse logische Aussagen im Zusammenhang mit einem Ereignis entwickeln (Barrouillet et al. 2008). Ein Modell wird im Arbeitsgedächtnis abgespeichert und stellt genau eine mögliche Belegung von Antezedens [p] und Konsequenz [q] dar, die die Wenn-Dann-Aussage zutreffend belegt. Dabei stellt ein Modell eine Möglichkeit dar.

Insgesamt gibt es drei Modelle, denen drei Verständnisniveaus zugeordnet sind (Johnson-Laird & Byrne, 2002). Beim ersten Modell ist ein Ereignis im [p q] Fall im Arbeitsgedächtnis gespeichert. Nach Evans (2008) unterliegt das erste Modell heuristischen Denkprozessen, das meist von schnellen, unbewussten Denkvorgängen begleitet wird. Zudem prägen vorwissensgestützte Überzeugungen dieses Modell, das sich als fehleranfällig erweist (Evans, 2008). Mit der zunehmenden Differenzierung des logischen Verständnisses wird jeweils ein weiteres Modell in einen Entscheidungsprozess während des Schlussfolgerns hinzugefügt. Während das erste Modell [p q] intuitiv abgerufen werden kann, wird das zweite und dritte Modell durch analytische Denkprozesse begleitet (Gauffroy & Barrouillet, 2009; Evans, 2008). Kontrollierte, bewusste sowie sequentielle Prozesse kennzeichnen die kognitiven, analytisch geprägten Vorgänge (Gauffroy & Barrouillet, 2009; Vergauwe, Gauffroy & Morsanyi, 2013).

Das einfachste mentale Modell ist die Kombination von [p] und [q]. Auf dieser Stufe können Fragen zu [p q] und [p¬q] richtig beantwortet werden. Diese Art der Verknüpfung wird auf dem konjunktiven Verständnisniveau geleistet (Barrouillet & Lecas, 2008). Auf diesem Niveau wird außer Acht gelassen, dass das Antezedens nicht immer vorhanden ist. Beim zweiten mentalen Modell werden die Schlüsse auf Grundlage der Annahme „Nur wenn, und genau dann, wenn die Voraussetzung [¬ p] gegeben ist, dann tritt die negative Konsequenz [¬q] ein“ gezogen (Johnson-Laird & Byrne, 2002). Diese Stufe wird mit dem bikonditionalen Verständnisniveau bezeichnet (Barrouillet & Lecas, 2008; Johnson-Laird & Byrne, 1991, 2002). Hier liegt ein Fehlschluss vor, da die gegebene Voraussetzung [¬ p] nicht unbedingt mit einer negativen Konsequenz [¬q] verknüpft sein muss, denn eine positive Konsequenz [q] ist ebenso

möglich. Das dritte mentale Modell besteht aus der Verbindung der Negationen von Antezedens  $[\neg p]$  und Konsequenz  $[\neg q]$ . Nach der Theorie der Mentalen Modelle können Menschen korrekt deduktiv schlussfolgern, wenn die drei möglichen Modelle im Entscheidungsprozess vollständig präsent sind und somit alle Formen von Zuständen korrekt zugeordnet werden können (Johnson-Laird & Byrne, 2002). Empirischen Befunden (Barrouillet et al., 2000; Vergauwe et al., 2013) zufolge benötigen Menschen beim angemessenen Umgang mit  $[\neg p]$  Ereignissen,  $[\neg p q]$ ;  $[\neg p \neg q]$ , wenn sie zwei bzw. drei Modelle beim Schlussfolgern nutzen, eine längere Zeitdauer zum Produzieren der folgerichtigen Antwort. Sie erkennen dann die beiden  $[\neg p]$  Ereignisse als irrelevant in Bezug auf die Konsequenz an (Barrouillet & Lecas, 2008; Laird & Byrne, 1991; 2002; Vergauwe et al., 2013).

In diesem Fall gelingen folgerichtige Kombinationen von Antezedens und Konsequenz einer konditionalen Aussage mit den zuvor genannten vier möglichen Belegungen (siehe Kapitel 5). Setzen Menschen bei Schlussfolgerungsprozessen ein mentales Modell ein, können sie ihr Arbeitsgedächtnis erleichtern, da sie sich auf das Antezedens fokussieren, das wahr bzw. falsch sein kann (Johnson-Laird & Byrne, 1991; 2002). Die Theorie der Mentalen Modelle besagt, dass Schlussfolgerungen nach der Schlussweise des Modus Ponens Kindern und ebenso vielen Erwachsenen leichter fallen (Johnson-Laird & Byrne 2002; Barrouillet et al., 2000) als Schlussfolgerungen nach Modus Tollens. Für die erste Schlussweise muss nur ein Modell im Arbeitsgedächtnis präsent sein, während bei der zweiten Schlussweise zwei Modelle im Arbeitsgedächtnis Voraussetzung sind (Barrouillet et al., 2000; Gauffroy & Barrouillet, 2009). Weiterhin wird nach der Theorie der Mentalen Modelle postuliert, dass die Affirmation der Konsequenz leichter gelingen kann als die Ablehnung des Antezedens (Barrouillet et al., 2000). Die theoretische Grundlage der Mentalen Modelle eignet sich, um entwicklungspsychologische Prozesse beim Schlussfolgern zu beschreiben (Gauffroy & Barrouillet, 2009; Gauffroy & Barrouillet; 2011).

Kombiniert man die Aussagen der Wahrheitstabelle mit der Theorie der Mentalen Modelle (Barrouillet et al. 2000, siehe Tabelle 4), ergibt sich folgender Zusammenhang. Aussagen des Modus Ponens sowie die des Modus Tollens sind korrekte Schlussweisen, wobei die Schlussweise des Modus Ponens wesentlich einfacher ist als die des Modus Tollens (Barrouillet et al., 2000), daher können bereits junge Kinder Schlussweisen des Modus Ponens richtig äußern. Die Schlussweise des Modus Tollens wird erst beherrscht, wenn alle drei mentalen Modelle präsent sind. Im Vergleich dazu wird die Schlussweise der Affirmation der Konsequenz häufig von Kindern gewählt, die ein oder zwei mentale Modelle beim Schlussfolgern

verwenden. Wenn Kinder alle drei mentalen Modelle bilden können, nimmt die nicht angemessene Schlussweise der Affirmation der Konsequenz ab.

Tabelle 4

*Vorhersage der Schlussweisen*

		Mentales Modell			
Anteze- dens	Konse- quenz	Schluss- weise	Konjunktiv [p q]	Bikonditio- nal [p q] [¬p¬q]	Konditional [p q] [¬p¬q]
[p]	[q]	Modus Ponens	+	+	+
[q]	[p]	Affirmation der Konse- quenz	+	+	-
[¬p]	[¬q]	Ablehnung des Anteze- dens	-	+	-
[¬q]	[¬p]	Modus Tol- lens	-	+	+

Anmerkung. += Produzieren von Antworten der konditionalen Aussage wenn p dann q; - = Antworten der konditionalen Aussage wenn p dann q, werden nicht produziert.

5.2 *Empirische Studien zur Entwicklung schlussfolgernden Denkens*

Die Fähigkeit zum Schlussfolgern entwickelt sich im Rahmen der kognitiven Entwicklung mit zunehmendem Alter (z.B. Barrouillet & Lecas, 2008; Gauffroy & Barrouillet, 2009). Daher werden in Tabelle 5 Studien mit Probanden unterschiedlichen Alters vorgestellt, um die Entwicklung der Kompetenzen beim Schlussfolgern zu verdeutlichen. Die Studien sind chronologisch nach ihrem Erscheinungsjahr aufgeführt. Zweites Ordnungskriterium für die Anordnung der Tabelle ist das Alter der Probanden. Weiterhin werden die Inhalte der Aufgaben (abstrakt und konkret) angegeben, da sie für die Aufgabenbearbeitung eine entscheidende Rolle spielen.

Tabelle 5

*Übersicht über Studien zum schlussfolgernden Denken*

Studie	Inhalte	Stichprobe (Probanden/Alter)
Wason, 1966	abstrakt	Erwachsene (Alter nicht angegeben)
Wason & Shapiro, 1971 (2. Experiment)	konkret	32 Erwachsene (Alter nicht angegeben)
Barrouillet & Lecas 1999	abstrakt	30 Schüler (9 Jahre); 30 Schüler (11-12 Jahre) 30 Schüler (15 Jahre)
Barrouillet et al., 2000	konkret	32 Schüler (11-12 Jahre)
Barrouillet & Lecas, 2002	konkret	Studenten (Alter nicht angegeben)
Barrouillet et al., 2008	abstrakt	47 Schüler (9 Jahre; 12 Jahre; 16 Jahre)
Gauffroy & Barrouillet, 2009	abstrakt	20 Schüler (8-9 Jahre) 24 Schüler (11-12 Jahre) 22 Schüler (15-16 Jahre) 26 Schüler (24-25 Jahre)
Gauffroy & Barrouillet, 2009	abstrakt	41 Schüler (8-9 Jahre) 41 Schüler (12 Jahre) 39 Schüler (15-16 Jahre) 51 Erwachsene (23-24 Jahre)
Tröbst et al., 2011	konkret	20 (7-8 Jahre) 23 (10-11 Jahre)
Vergauwe et al., 2013	abstrakt	46 Studenten (22-23 Jahre)
Robisch et al., 2014	konkret	101 Schüler (8-9 Jahre)

Empirische Befunde aus einer frühen Studie mit der *Wason Selection Task* (Wason, 1966) zeigten, dass selbst Erwachsene Schwierigkeiten bei der Evaluation eines Ereignisses hatten, bei denen das Antezedens fehlte. In diesem Fall konnte eine vorgegebene Regel nicht eindeutig bestätigt oder falsifiziert werden (ebenda). Als Aufgabe (siehe Abbildung 8) wurden Erwachsenen vier Karten vorgelegt. Auf jeder Karte waren ein Buchstabe auf der einen Seite sowie eine Zahl auf der anderen Seite abgebildet.

Nach der Erklärung wurde den Probanden folgende konditionale Aussage genannt: „Wenn eine Karte den Buchstaben A auf der einen Seite aufweist, dann hat sie auf der anderen Seite eine 2.“ (Wason, 1966, S.7).

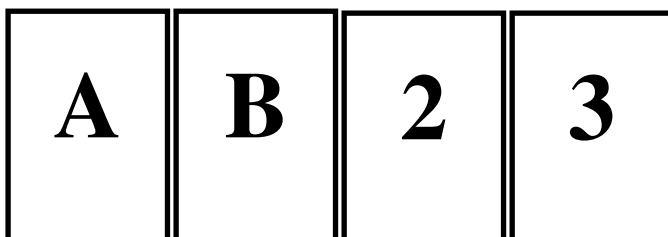


Abbildung 8. Vier Karten der *Wason Selection Task*.

Im Anschluss daran wurden die Probanden aufgefordert, die Karten aufzudecken, die die gegebene konditionale Aussage bestätigte oder widerlegte. Dabei sollten sie darauf achten, nur die zur Prüfung der Aussage notwendigen Karten aufzudecken. Die meisten Erwachsenen wählten die beiden Karten mit dem Buchstaben A und der Zahl 2. Einige wählten auch nur die Karte mit dem Buchstaben A. Mit beiden Antworten konnte die Aufgabe vollständig gelöst werden.

Die Probanden hätten zum richtigen Beantworten der Frage die Karte mit der Zahl 3 aufdecken müssen; denn mit dieser Karte konnte die genannte konditionale Aussage widerlegt werden. Viele Erwachsene berücksichtigten bei dem Experiment mit abstraktem Inhalt nicht, dass die konditionale Aussage durch Aufdecken der Zahl 3 widerlegt werden könnte (Johnson-Laird & Byrne, 2002). Sie konnten daher nicht korrekt schlussfolgern.

In einem weiteren Experiment (Wason & Shapiro, 1971) wurde geprüft, inwieweit Erwachsene in einem plausiblen, alltagsnahen Kontext schlussfolgern können. Dazu wurden zwei unabhängige Gruppen untersucht; eine Gruppe prüfte Aussagen mit abstraktem Inhalt, die andere Gruppe bewertete Aufgaben zum Schlussfolgern mit konkretem Inhalt. Dabei zeigte sich, dass die Teilnehmer bei Aufgaben mit realistischem Inhalt signifikant besser antworten konnten als bei den Aufgaben mit abstraktem Inhalt (Wason & Shapiro, 1971).

Für das bessere Abschneiden beim Schlussfolgern im thematisch gebundenen Kontext werden drei mögliche Gründe genannt: Zum einen wurden in der Aufgabenstellung konkrete Inhalte verwendet, die sich die Teilnehmer besser einprägen konnten als abstrakte Inhalte (ebenda). Weiterhin wirkten sich die konkreten Beziehungen im realistischen Kontext zwischen den Inhalten beim Schlussfolgern positiv aus (Wason & Shapiro, 1971). Zudem formte sich eine inhaltliche Einheit durch die Verknüpfung der Inhalte von Antezedens [p] und Konsequenz [q]. Durch bessere Informationsspeicherung im Arbeitsgedächtnis (ebenda) konnten die Probanden häufiger logische Schlussfolgerungen äußern. Die drei genannten Gründe trugen dazu bei, dass den Probanden das Schlussfolgern im Rahmen eines konkreten Inhaltes leichter fiel (Wason & Shapiro, 1971).



In der Querschnittstudie von Barrouillet, Gauffroy & Lecas (2008) wurden Probanden in verschiedenen Altersgruppen Aufgaben in einer *Truth Testing Task* (Wason, 1966) vorgelegt. Hier wurden die altersbezogenen Kompetenzen beim Schlussfolgern detailliert in den Blick genommen. An der Studie nahmen vier Altersgruppen teil: Schüler im Alter von durchschnittlich achteinhalb Jahren, elfeinhalb Jahren, fünfzehneinhalb Jahren sowie Erwachsene mit einem durchschnittlichen Alter von 28 Jahren. Ihnen wurden abstrakte Aussagen vorgelegt, die sie prüfen sollten. Eine Aussage war beispielsweise „Wenn der Kreis rot ist, dann ist der Stern schwarz“ (Barrouillet et al., 2008). Anschließend wurden die farbigen Symbole (Kreise und Sterne) in der Weise präsentiert, dass sich alle vier Möglichkeiten ergaben:

$[p \ q]$ ;  $[p \neg q]$ ;  $[\neg p \ q]$ ;  $[\neg p \neg q]$

Beispielsweise trafen der blaue Kreis und der grüne Stern für das Ereignis  $[\neg p \neg q]$  zu, denn sowohl das Antezedens als auch die Konsequenz waren nicht gegeben. Barrouillet & Kollegen (2008) konnten zeigen, dass die bestätigenden Aussagen  $[p \ q]$ , bei denen nur ein Modell im Arbeitsgedächtnis vorhanden sein muss, von fast allen Probanden gelöst werden konnten (Barrouillet et al., 2008). Ebenso wurde bei den widerlegenden Fällen  $[p \neg q]$  ein Alterstrend deutlich; so konnten die Probanden die widerlegenden Fälle mit zunehmendem Alter besser erkennen. Weitaus deutlichere altersbedingte Unterschiede beim Schlussfolgern zeigten sich bei den  $[\neg p]$  Ereignissen, die insgesamt irrelevant im Zusammenhang mit der zuvor genannten Aussage zu bewerten waren (Barrouillet et al., 2008). Die jüngeren Teilnehmer (Kinder im Alter von acht und elf Jahren) identifizierten die Ereignisse mit  $[\neg p \ q]$ ;  $[\neg p \neg q]$  als falsch. Dabei schätzten 86% der Achtjährigen und 70% der Elfjährigen die  $[\neg p \neg q]$  Ereignisse als falsch ein; die  $[\neg p \ q]$  Ereignisse wurden zu 66% respektive 68% als falsch gewertet. Bei den Probanden im Alter von fünfzehn Jahren gaben beim  $[\neg p \ q]$  Ereignis noch 61% der Probanden an, die Aussage sei falsch; wohingegen 70% der Fünfzehnjährigen das  $[\neg p \neg q]$  Ereignis als irrelevant einstufen. Die Erwachsenen konnten in beiden  $[\neg p]$  Fällen hohe Lösungsraten erzielen; so antworteten sie zu 90% beim  $[\neg p \neg q]$  Ereignis und zu 82% beim  $[\neg p \ q]$  Ereignis richtig (Barrouillet et al., 2008). Insgesamt zeigte sich in der Querschnittstudie von Barrouillet et al. (2008) eine deutliche Weiterentwicklung der Kompetenz zum Schlussfolgern von den Grundschulern über die Sekundarstufenschüler bis hin zu den Erwachsenen. Dabei wurden aus den vier Ereignissen  $[p \ q]$ ;  $[p \neg q]$ ;  $[\neg p \ q]$ ;  $[\neg p \neg q]$  unterschiedliche Kombinationen bei den Antworten herausgearbeitet und so spezifische Antwortmuster identifiziert. (Barrouillet et al., 2008; Tröbst, Hardy & Möller, 2011). Diese lassen sich in drei aufeinanderfolgenden Stufen abbilden (Barrouillet et al., 2008, siehe Tabelle 6). Im Ergebnis zeigten die Achtjährigen meist ein konjunktives Niveau

auf Stufe 1 (siehe Tabelle 6), wobei sie beim Schlussfolgern nur das mentale Modell  $[p \ q]$  präsent hatten.

Dabei zeigten sie ein Verständnis, bei dem sie das Antezedens der Aussage mit der Konsequenz durch ein einfaches „und“ verknüpften (Barrouillet et al., 2008; Tröbst et al., 2011). So schlussfolgerten sie bei den bestätigenden und widerlegenden Ereignissen richtig, die beiden anderen Ereignisse  $[\neg p \ q]$ ;  $[\neg p \neg q]$  sahen sie als falsch an. Die Elfjährigen antworteten in den meisten Fällen auf dem bikonditionalen Niveau; dabei bewerteten sie das Antezedens als einzigen und ausschließlichen Grund für die Konsequenz (Johnson-Laird/Byrne, 2002; Tröbst et al., 2011). Auf dieser Stufe schätzten sie die drei Ereignisse  $[p \ q]$   $[p \neg q]$ ;  $[\neg p \neg q]$  richtig ein; nur das Ereignis  $[\neg p \ q]$  bewerteten sie falsch. Bei den Erwachsenen zeigte sich ein hauptsächlich konditional geprägtes Interpretationsniveau (Barrouillet et al., 2008; Tröbst et al., 2011).

Tabelle 6

*Interpretationsniveaus beim Schlussfolgern*

	1. Stufe konjunktiv	2. Stufe bikonditional	3. Stufe konditional
$[p \ q]$	+	+	+
$[p \neg q]$	-	-	-
$[\neg p \ q]$	-	-	○
$[\neg p \neg q]$	-	○	○

*Anmerkung.* + = Hypothese bestätigende Aussage, - = Hypothese widerlegende Aussage, ○ = für Hypothese irrelevante Aussage.

In Anlehnung an die zuvor beschriebene Studie haben Tröbst et al. (2011) die Schlussfolgerungskompetenzen von Kindern mit einem durchschnittlichen Alter von siebeneinhalb Jahren sowie von etwa zehn Jahren in einem konkreten Kontext anhand der Frage „Wie kommt es, dass Dinge springen?“ (in Anlehnung an Thiel, 1987) geprüft. Den Kindern wurden typische Äußerungen anderer Kinder als einfache Aussage (z.B. „weiche Dinge springen“) präsentiert. Dabei enthielt jede Aussage entsprechend der Logik des deduktiven Schlussfolgerns jeweils ein Antezedens (Eigenschaft des Gegenstandes) sowie eine Konsequenz (Sprungverhalten). Durch Kombination von Antezedens und Konsequenz ergaben sich genau wie bei der Studie von Barrouillet et al. (2008) vier mögliche Ereignisse  $[p \ q \text{ best}]$   $[p \neg q \text{ wid}]$ ;  $[\neg p \ q \text{ irrel}]$ ;  $[\neg p \neg q \text{ irrel}]$ .

Dementsprechend wurden den Kindern vier zu prüfende Gegenstände präsentiert. Mit diesen sollten die Kinder die gegebene Aussage („weiche Dinge springen“) prüfen. Sie sollten entscheiden, ob ein Gegenstand die Aussage bestätigt, widerlegt oder aber irrelevant für die zu prüfende Aussage war.

In einer multivariaten Analyse zeigte sich ein signifikanter Interaktionseffekt der Faktoren Ereignis und Alter bei den einzelnen Ereignissen. Die Antworten in den beiden Altersstufen der Sieben- und der Zehnjährigen haben sich demzufolge signifikant voneinander unterschieden (Tröbst et al., 2011). Zudem hat sich ein signifikanter Haupteffekt in Bezug auf die einzelnen zu prüfenden Ereignisse gezeigt. So konnten Kinder beider Altersstufen die Ereignisse der Form  $[p \rightarrow q]$  vorwiegend als bestätigend erkennen. Die Ereignisse der Form  $[p \rightarrow \neg q]$  wurden von beiden Altersgruppen meist als widerlegend eingeschätzt (ebenda). Im Gegensatz dazu wurde das Ereignis  $[\neg p \rightarrow q]$  nicht gleichermaßen von den Sieben- und den Elfjährigen eingeschätzt. Während die Elfjährigen das Ereignis eher als bestätigend einstufen, werteten die Siebenjährigen das Ereignis in unterschiedlicher Weise (teilweise bestätigend bzw. teilweise widerlegend). Kinder beider Altersstufen konnten nicht im konditionalen Schlussfolgerungsmuster antworten.

Insgesamt belegen die erläuterten Studien (Barrouillet et al. 2008; Tröbst et al., 2011), dass mit zunehmender kognitiver Entwicklung eine Steigerung der Kompetenzen beim Schlussfolgern zu erkennen ist. Während in einigen Studien vorwiegend Aufgaben mit abstrakten Inhalten genutzt wurden (Wason, 1966; Barrouillet et al., 2008; Gauffroy & Barrouillet, 2009), wurden häufig auch Schlussfolgerungsaufgaben mit konkreten, alltagsbezogenen Inhalten angeboten (Tröbst et al., 2011), um den Probanden das Schlussfolgern zu erleichtern (Wason & Shapiro, 1971). Das Nutzen von konkreten Kontexten, wie beispielsweise die Untersuchung der Frage „Wie kommt es, dass ein Ball springt?“ (in Anlehnung an Thiel, 1987) hatte den Vorteil, dass Kinder bereits Erfahrungen mit den Inhalten des Kontextes gewinnen konnten. In der vorliegenden Studie wurden Hypothesen daher ebenso im Inhaltsgebiet *Elastizität und Plastizität* geprüft. Im folgenden Unterkapitel wird aufgezeigt, inwiefern exekutive Funktionen und Schlussfolgern miteinander zusammenhängen.

### 5.3 Exekutive Funktionen und Schlussfolgern

Exekutive Funktionen spielen eine zentrale Rolle bei kognitiven Prozessen (Röthlisberger, Neuenschwander, Michel & Roebers, 2010), und nehmen daher sowohl beim Schlussfolgern als auch beim naturwissenschaftlichen Lernen insgesamt einen zentralen Stellenwert ein

(Gropen, Clark-Ciarelli, Hoisington, Stacy & Ehrlich, 2011). Der Begriff *exekutive Funktionen* stammte ursprünglich aus dem Bereich der Neuropsychologie. Forschungen im medizinischen Bereich haben gezeigt, dass Gedächtnislücken bei Patienten mit der Beschädigung der frontalen Hirnloben zusammenhängen (Pennington & Ozonoff, 1996). Unter *exekutiven Funktionen* sind selbstregulatorische sowie kognitive Fähigkeiten subsumiert, mit denen es Menschen gelingt, Handlungen zielgerichtet durchzuführen und somit ihr Denken und Handeln zu leiten (Archibald & Kerns, 1999). Dazu gehören Teilprozesse wie das Arbeitsgedächtnis, Planung, Inhibitionskontrolle sowie Steuerung der Aufmerksamkeit (Archibald & Kerns, 1999; Carlson, 2005; Garon, Bryson & Smith, 2008). Mithilfe des Arbeitsgedächtnisses sind Menschen in der Lage, Informationen zu speichern und zu einem anderen Zeitpunkt abzurufen.

Insbesondere die Fähigkeit zur Inhibition ist maßgeblich an dem analytischen Prozess des Schlussfolgerns (Markovits, 2006) im Rahmen des naturwissenschaftlichen Lernens beteiligt (Gropen et al., 2011). Die Inhibitionskontrolle bezeichnet die Fähigkeit, vorschnelle bzw. automatisierte fehlerhafte Reaktionen zu unterdrücken (Dowsett & Livesey, 2000; Miyake, Friedman, Rettinger, Priti & Hegarty, 2001).

Diese Kompetenz ermöglicht es, vorschnelle Antworten zurückzuhalten und sich stattdessen auf die Evidenz, die aus einem Experiment hervorgeht, zu konzentrieren. Die empirischen Befunde zeigen, dass es Kindern im Alter von vier bis fünf Jahren gelingen kann, Teilbereiche der exekutiven Funktionen wie das Arbeitsgedächtnis oder die Fähigkeit zur Inhibition in kognitiven Prozessen einzusetzen. Empirische Befunde belegen, dass Zusammenhänge zwischen der Entwicklung eines Perspektivenwechsels und der Entwicklung von exekutiven Funktionen bestehen (Perner & Lang, 1999; Sodian & Hülken, 2005). Kinder im Vorschulalter können bereits Regeln als inkompatibel erkennen (Röthlisberger et al., 2010; Zelazo, Müller, Frye & Markovits, 2003).

Im Folgenden wird das Modell des Arbeitsgedächtnisses, das ursprünglich von Baddeley & Hitch, 1974 entwickelt worden ist, dargestellt. Im Zentrum des Modells steht die Zentrale Exekutive, die als domänenübergreifende Komponente aufgefasst wird. Damit können Strategien zur Speicherung sowie zum Abrufen von Informationen angewendet werden (Schmid, Zoelch & Roebers, 2008). Der Zentralen Exekutiven mit ihrer Kontrollfunktion sind die beiden Subsysteme, der Visuelle Notizblock (Baddeley & Hitch, 1974) sowie die Phonologische Schleife zugeordnet (siehe Abbildung 9). Während die Zentrale Exekutive domänenübergreifende Kontrollfunktionen übernimmt, sind die beiden Subsysteme für die Aufnahme und Speicherung von spezifischen, visuellen sowie auditiven Informationen verantwortlich (Baddeley & Hitch, 1974; Garon et al., 2008; Schmid et al., 2008). Für den auditiven Bereich ist die

Phonologische Schleife (*phonological loop*) zuständig (Baddeley & Hitch, 1974). Diese übernimmt eine wichtige Funktion beim Lesen lernen sowie beim Aussprechen von Vokabeln (Baddeley, Addeley & Hitch, 2011). Demgegenüber übernimmt der Visuelle Notizblock (Baddeley & Hitch, 1974) die Speicherung von Wortbildern sowie anderen Repräsentationen im bildlich orientierten Speicher (Baddeley, 2000). Beide Subsysteme bestehen jeweils aus einem passiven Speicher sowie einem aktiven Kontrollprozess. Vor allem durch den zuletzt genannten Mechanismus können Informationen leicht zergliedert und durch Wiederholung aufgenommen werden. Hier werden die Einheiten der Zentralen Exekutiven als eigenständige Komponenten betrachtet. Im Modell nach Baddeley & Hitch (1974) und Baddeley (2000) sind die Komponenten wie in Abbildung 9 angeordnet.

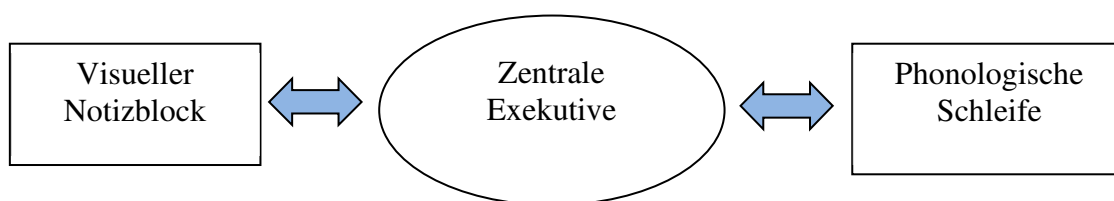


Abbildung 9. Modell der Zentralen Exekutiven nach Baddeley & Hitch (1974).

Das ursprüngliche Modell (siehe Abbildung 9) ist auf Grundlage der Feststellung erweitert worden, dass Gedächtnisleistungen unterschiedlich lange im Gehirn gespeichert werden können (Baddeley, 2000). Dank der Phonologischen Schleife können Inhalte im Langzeitgedächtnis gespeichert und in Sprache umgewandelt werden (ebenda). Die Informationen, die mit Hilfe des Visuellen Notizblocks aufgenommen werden, können ebenso verwendet werden. Allerdings müssen sie vor der Speicherung dekodiert werden; im Vergleich dazu können auditiv erfasste Informationen direkt weiterverwertet werden (Baddeley et al., 2011). In der Studie von Gathercole, Pickering, Ambridge & Wearing (2004) wurden Kindern Aufgaben mit den Bereichen, Phonologische Schleife und Visueller Notizblock vorgelegt; ebenso wurden komplexe Spannaufgaben gezeigt. In der Faktorenanalyse ergab sich ein Drei-Faktoren-Modell, in dem die drei zuvor genannten Komponenten vorgefunden wurden. Damit wurde die Struktur des Arbeitsgedächtnisses nach dem Modell von Baddeley & Hitch (1974) bestätigt (Gathercole et al., 2004).

Später hat Baddeley das Modell zum Arbeitsgedächtnis noch durch ein viertes Element, den Episodischen Buffer (Baddeley, 2000) erweitert (siehe Abbildung 10). Dieser schließt unterschiedliche Informationen der Subsysteme zusammen und fungiert daher als integratives

Element. Der Episodische Buffer kann für kurze Zeit Ereignisse speichern und daraus Informationen, die von der Zentralen Exekutiven stammen, an die benachbarten Systeme weiterleiten (Baddeley, 2000; Baddeley et al., 2011). Daher stellt der Episodische Buffer im Arbeitsgedächtnis ein Verbindungsstück dar, das von der Zentralen Exekutiven kontrolliert wird. Neben der integrierenden Funktion kann der Episodische Buffer auch zu neuen kognitiven Repräsentationen beitragen. Neue Informationen werden durch die Zentrale Exekutive aufgenommen und gegebenenfalls weitergeleitet (Baddeley, 2000; Baddeley et al., 2011). Mit dem erweiterten Konzept zeigte Baddeley, wie Menschen unterschiedliche Informationen speichern und später wieder abrufen können (siehe Abbildung 10). Beispielsweise konnten Menschen gesprochene Wörter auditiv wahrnehmen, speichern und sprachlich wiedergeben oder visuelle Repräsentationen ebenso in sprachliche Formen umsetzen (Baddeley et al., 2011).

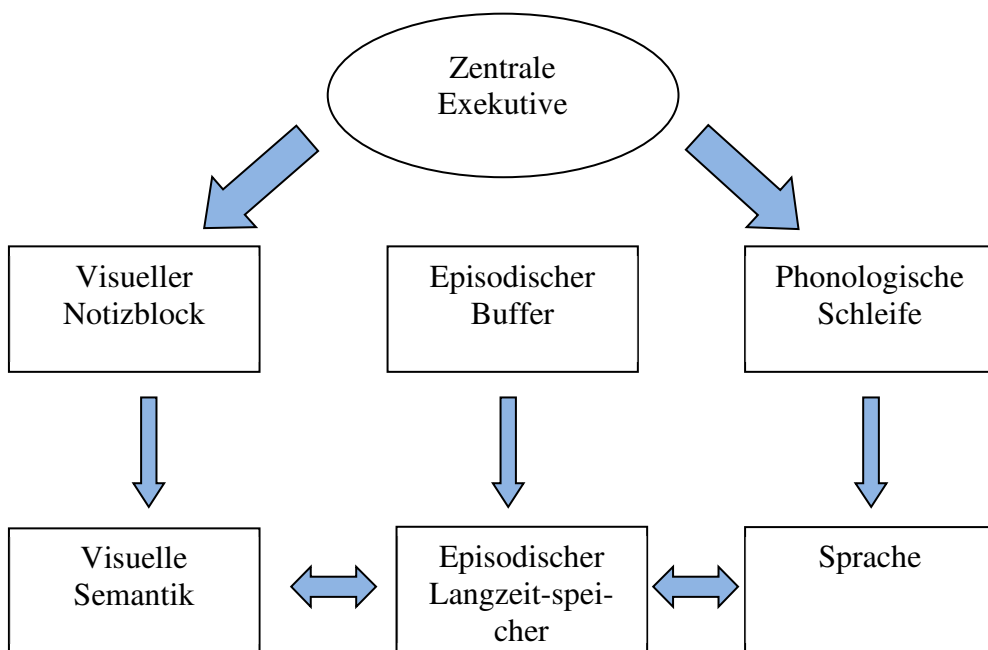


Abbildung 10. Modell der Zentralen Exekutiven nach Baddeley (2000).

Der Episodische Buffer nimmt durch Herstellen von Verknüpfungen, sogenannten *chunks* (Baddeley et al., 2011), beim Integrieren neu gelernter Wissensinhalte eine wichtige Funktion ein. Aufgrund seiner integrativen Funktion befindet sich der Episodische Buffer im Modell (siehe Abbildung 10) in der Mitte zwischen dem visuellen Notizblock und der phonologischen Schleife. Im unteren Bereich des Modells ist das kognitive System abgebildet; mit diesem gelingt es Menschen, Wissen anzusammeln, zu speichern und in Sprache zu modifizieren (Baddeley, 2000). Die drei Bereiche, die im Modell des Arbeitsgedächtnisses dargestellt

sind, entwickeln sich mit zunehmendem Alter (Garon et al., 2008). Sie spielen eine wichtige Rolle innerhalb der kognitiven Entwicklung eines Kindes (Archibald & Kerns, 1999; Hasselhorn et al., 2000). Empirischen Befunden zufolge besteht vor allem bei jungen Kindern eine enge Verknüpfung der drei Hauptelemente (Baddeley & Hitch, 1974).

Englische Forscher der Forschergruppe von Gathercole haben Aufgaben für Kinder von circa sechs Jahren entwickelt, um die Kompetenzen in diesem Alter zu erfassen (Alloway, Gathercole, Willis & Adams, 2004). Weitere Untersuchungen haben gezeigt, dass sich das Arbeitsgedächtnis sowie exekutive Funktionen insgesamt besonders eignen, um Entwicklungsveränderungen im kindlichen Denken aufzuzeigen (Gathercole, 1999; Hasselhorn, Seidler, Brandler & Körner, 2000). Bei diesen Aufgaben ging es zudem darum, den Einfluss der exekutiven Funktionen auf kognitive Entwicklungsprozesse zu zeigen (Schmid, Zoelch & Roebbers, 2008). So sind der visuelle Notizblock, der episodische Buffer sowie die Phonologische Schleife eng miteinander verbunden. Im Laufe der Denkentwicklung zeigten sich vor allem im Bereich der phonologischen Schleife Kompetenzsteigerungen (Schmid et al., 2008). Empirische Befunde belegen, dass exekutive Funktionen zwischen dem dritten und fünften Lebensjahr bedeutend zunehmen (Diamond, Carlson & Beck, 2005; Zelazo et al., 2003). Dies ergab sich durch Verbesserung der Inhibitionskontrolle (Diamond, 2002). Außerdem gewannen Kinder in diesem Alter verstärkt Einsicht in Strukturen von regelorientierten Systemen (Zelazo et al., 2003). Auch die Perspektivenübernahme gelang ihnen sukzessive immer besser (Perner, 1991). Ebenso wuchs mit dem Alter die Dauer der Aufmerksamkeitsspanne (Garon, Bryson & Smith, 2008). Einige andere Wissenschaftler sehen exekutive Funktionen ebenso als einheitliches System (Diamond & Prevor, 1997; Friedman & Miyake, 2004; Hughes, 1998; Lehto, Juujärvi, Kooistra & Pulkkinen, 2003).

Andere Forscher sind überzeugt, dass exekutive Funktionen mit ihren zentralen Komponenten, dem Arbeitsgedächtnis sowie der Inhibitionskontrolle als eigenständige Komponenten fungieren (Diamond, 2002; Letho et al., 2003; Pennington & Ozonoff, 1996). Forschungen aus der Neuropsychologie zeigten, dass Patienten einige Aufgaben aus einem Teilbereich lösen konnten, während sie andere nicht beantworten konnten (Miyake, Friedman, Emerson, Witzki, Howerter, 2000). Daran wurde deutlich, dass exekutive Funktionen kein einheitliches System mit Subfunktionen darstellen (Miyake et al., 2000), sondern eigenständige Teilbereiche beinhalten. In ihrer Studie konnten die einzelnen Komponenten Wechsel von Aufgaben (*Shifting*), Aktualisieren und Kontrollieren von Repräsentationen des Arbeitsgedächtnisses (*Updating*) sowie Unterdrücken von vorschnellen Antwort (*Inhibition*) in Strukturgleichungsmodellen als einzelne latente Variablen identifiziert werden (Miyake et al., 2000). Empirische Befunde

zeigten signifikante Korrelationen, dennoch luden die beiden exekutiven Teilfunktionen, Arbeitsgedächtnis und Inhibition, auf unterschiedlichen Faktoren in der Faktorenanalyse (Lehto et al., 2003). Die Befunde verdeutlichten, dass die Funktionen separiert voneinander bestehen, auch wenn sie miteinander verknüpft sind (Lehto, et al., 2003; Miyake et al., 2000).

In diesem Abschnitt wurden die beiden Forschungsstränge, die exekutive Funktionen entweder als einheitliches System (Baddeley, 2000; Baddeley & Hitch, 2011) oder als eigenständige Teilkomponenten (Diamond, 2002; Lehto et al., 2003; Miyake et al., 2000; Pennington & Ozonoff, 1996) betrachten, beschrieben. Während der erste Forschungsstrang von der kognitiven Psychologie geprägt ist, ist der zweite Ansatz durch neuropsychologische Sichtweisen beeinflusst. Zusammenfassend kann festgestellt werden, dass exekutive Funktionen eine bedeutsame Rolle für die kognitive Entwicklung eines Kindes spielen (Röthlisberger et al., 2010); im Vorschulalter zeigen sich deutliche Fortschritte, insbesondere bei der Inhibitionskontrolle (Archibald & Kerns, 1999) sowie bei der Arbeitsgedächtniskapazität (Zoelch et al., 2005). Die Ausprägung der exekutiven Funktionen eines Kindes sind stark von seiner individuellen sozio-ökonomischen Herkunft geprägt (Noble, Norman & Farah, 2005). Kinder aus sozial benachteiligten Familien verfügten in Testaufgaben über schlechtere Ergebnisse bei den exekutiven Funktionen (Noble, Norman & Farah 2005). Daher gehen diese Autoren davon aus, dass Schlussfolgern und Sprachkompetenz eng miteinander verknüpft sind.

#### 5.4 *Schlussfolgern und Argumentieren*

Der Begriff *Argumentieren* ist geprägt vom Philosophen Toulmin (1958) und beruht auf einer logischen Folge einzelner Schritte. Nach Toulmin (1958) stellt ein Mensch, entsprechend der eigenen Haltung bzw. Überzeugung, eine Behauptung auf. Zur Stärkung derselben führt er Informationen, wie beispielsweise passende Daten bzw. Fakten an und kann so einem Zuhörer seine Behauptung im Idealfall glaubhaft darlegen.

Schlussfolgerungen sind grundlegend beim wissenschaftlichen Argumentieren (Hardy, Möller, Kloetzer & Sodian, 2010 b) und folgen nach Lawson (2010) einem Prozess, der aus vier Schritten besteht. Diese benannte der Autor mit „Abduktion, Retroduktion, Deduktion sowie Induktion“ (ebenda, S. 340). Im ersten Schritt wird eine mögliche Hypothese aufgestellt (Abduktion), im zweiten Schritt wird diese im Rahmen eines ersten Testes überprüft (Retroduktion). Anschließend erfolgt im dritten Schritt die Folgerung bzw. Verfeinerung der ursprünglichen Aussage (Deduktion) und es werden neue Vorhersagen getroffen. Alternativ wird im dritten Schritt die Induktion angewendet. Dabei wird von der Beobachtung eines



Gegenstandes auf weitere Gegenstände geschlossen (ebenda, S. 341). Im vierten Schritt wird die zuvor getroffene Aussage mit der Beobachtung verglichen. Stimmen beide Komponenten überein, wird die eingangs aufgestellte Aussage bestätigt; im anderen Fall wird sie widerlegt. Nach Lawson (2010) berücksichtigt eine gute Argumentation mögliche Alternativen mit dazugehöriger Evidenz, um die Behauptung schließlich zu bestätigen oder zu widerlegen.

### 5.5 Zusammenfassung und Fazit

In diesem Kapitel wurden Theorien zum Schlussfolgern sowie die Entwicklung der Fähigkeit aus konditionalen Aussagen korrekte Schlüsse zu ziehen, in den Blick genommen. Insgesamt werden vier Schlussweisen unterschieden (siehe Kapitel 5.1), davon sind jedoch nur Modus Ponens sowie Modus Tollens formal korrekte Schlussweisen. Der Modus Ponens beschreibt eine Wenn-Dann-Beziehung, die von einer vorhandenen Prämisse ausgeht, während beim Modus Tollens eine nicht gegebene Prämisse die Grundlage zum Schlussfolgern bildet. Das Schlussfolgern auf Basis einer vorhandenen Prämisse ist empirischen Befunden zufolge leichter (Barrouillet & Lecas, 2000; Barrouillet et al., 2008; Tröbst et al., 2011) als das Schlussfolgern auf Grundlage einer nicht vorhandenen Prämisse (Vergauwe et al., 2013). In diesem Fall ist ein Ereignis im Zusammenhang mit einer Hypothese irrelevant und gleichzeitig besonders schwer zu interpretieren (Wason & Shapiro, 1971; Tröbst et al., 2011). Daher neigen Kinder häufig dazu, eine nicht vorhandene Prämisse zu ignorieren oder falsch zu interpretieren (Chinn & Brewer, 1998). Empirischen Befunden (Barrouillet et al., 2008; Leever, 1999; Tröbst et al., 2011) zufolge steigert sich die Schlussfolgerungskompetenz mit zunehmendem Alter. Verschiedene Wissenschaftler (Barrouillet et al., 2008; Tröbst et al., 2011) konnten in ihren Studien eine stufenweise Entwicklung vom konjunktiven über das bikonditionalen hin zum konditionalen Interpretationsniveau zeigen (siehe Kapitel 5.2). Entscheidenden Einfluss auf das Schlussfolgern nehmen exekutive Funktionen (siehe Kapitel 5.3), die sich mit zunehmendem Alter bei kontinuierlicher Förderung exekutiver Funktionen ausprägen (Gropen, et al., 2011; Röthlisberger et al., 2010). Empirischen Befunden zufolge können Entwicklungsveränderungen im Kindesalter sehr gut mit Aufgaben in diesem Bereich dargelegt werden (Gathercole et al., 2004; Hasselhorn et al., 2000). Beide Elemente, Schlussfolgern sowie exekutive Funktionen, prägen sich im Laufe der kognitiven Entwicklung eines Kindes aus (Kapitel 5.3). Weiterhin wurde in diesem Kapitel der Zusammenhang zwischen Schlussfolgern und Argumentieren nach Toulmin (1958) und Lawson (2010) erläutert (Kapitel 5.4). Als Fazit dieses Kapitels lässt sich festhalten, dass

sich die Fähigkeit zum Schlussfolgern ebenso wie die exekutiven Funktionen bei entsprechender Förderung sukzessive im Laufe der Entwicklung eines Kindes herausbilden.

## 6 Forschungslücken und Ziele

Um die Zielsetzungen von der Querschnitt- sowie der Trainingsstudie zu verdeutlichen, werden in diesem Kapitel die Forschungslücken im Bereich frühkindlicher naturwissenschaftlicher Bildung dargestellt. Sie sind relevant für die Fragestellungen sowie für die Konzeption der beiden vorliegenden Studien.

Im Grundschulbereich liegen einige Studien zur Stärkung des Wissenschaftsverständnisses als bereichsübergreifende, naturwissenschaftliche Kompetenz vor. Zudem liegen auch einige Studien vor, die sich mit Kompetenzen der Verknüpfung von Theorie und Evidenz bei Kindern im Elementarbereich befassen. So beforschten Koerber et al. (2005) die Theorie-Evidenz-Koordination bei fünfjährigen Kindern. Sie stellten fest, dass Kinder bereits in diesem Alter über basale Fähigkeiten zur Theorie-Evidenz-Koordination verfügen. Kinder können Zusammenhänge verstehen, wenn diese im Rahmen eines inhaltlichen Kontextes angeboten werden. Die Studien von Ruffman et al. (1993) und Koerber et al. (2005) untersuchten die Fähigkeiten junger Kinder bei der Koordination von Theorie und Evidenz. In den erläuterten Forschungsarbeiten wurden die Kompetenzen zur Theorie-Evidenz-Koordination jedoch nicht für Kinder unterschiedlichen Alters gemessen. Es fehlt bisher ein Test, der über die Altersspanne von vier- bis zehnjährigen Kindern die Kompetenzen in diesem Bereich erfasst. Daher ist im Rahmen der Querschnittstudie ein Wissenschaftsverständnistest entwickelt worden, der die Kompetenzen bei der Theorie-Evidenz-Koordination erfasst. Die Kompetenzen in den verschiedenen Altersstufen wurden in diesem Bereich detailliert durch eine Item-Response-Analyse erfasst.

Weiterhin liegt keine Studie bisher vor, in der die Schlussfolgerungskompetenzen von Kindern in ihrer Entwicklung vom Elementar- über den gesamten Primarbereich hinweg im Rahmen dieses Inhaltsgebietes erfasst werden. Die Querschnittstudie zielte darauf ab, mit empirischen Daten in vier Altersstufen (vier Jahre, fünf bis sechs Jahre, sieben bis acht Jahre, neun bis zehn Jahre) aufzuzeigen, über welche Kompetenzen Kinder zur Theorie-Evidenz-Koordination jeweils verfügen und welche Kompetenzniveaus sie beim Schlussfolgern erreichen. Beide Tests, Schlussfolgerungstest und Wissenschaftsverständnistest, wurden im Rahmen der Vorstudie auf ihre Durchführbarkeit und Validität hin überprüft.

Bisherige Forschungsarbeiten untersuchten zudem nicht den Zusammenhang zwischen Theorie-Evidenz-Koordination, exekutiven Funktionen und der Kompetenz des Schlussfolgerns im Laufe der Entwicklung von Kindern. Daher wurde in der Querschnittstudie untersucht, wie exekutive Funktionen sowie Wissenschaftsverständnis sich auf das Schlussfolgern auswirken. Durch die Querschnittstudie wurden die bestehenden Forschungslücken geschlossen, indem die Entwicklungen von Schlussfolgerungskompetenzen im Zusammenhang mit exekutiven Funktionen und Wissenschaftsverständnis bei Kindern im Alter von vier bis zehn Jahren aufgezeigt wurden.

Die Erkenntnisse aus der Querschnittstudie dienten als Grundlage für die Entwicklung der Trainingsmaßnahmen der Trainingsstudie. Mit dem Wissen, über welche Kompetenzen Kinder in den Altersstufen (vier Jahre, fünf bis sechs Jahre, sieben bis acht Jahre, neun bis zehn Jahre) verfügen, konnten altersgemäße Trainingsmaßnahmen entwickelt werden. In der vorliegenden Arbeit wurden diese für Kinder im Elementarbereich, im Alter von fünf bis sechs Jahren erarbeitet und untersucht.

Insgesamt liegen nur wenige empirische Studien zur bereichsspezifischen Förderung naturwissenschaftlicher Kompetenzen bei Vorschulkindern vor. Die wichtige Frage, welche Maßnahmen sich zur Förderung von Denkprozessen, wie der Koordination von Theorie und Evidenz bei Kindern im Elementarbereich eignen, ist bislang nicht umfassend beantwortet worden. Der Einsatz unterschiedlicher Trainingsmaßnahmen zur gezielten Förderung einer Theorie-Evidenz-Koordination im Zusammenhang mit schlussfolgerndem Denken ist bislang nur bei Kindern im Grundschulalter untersucht worden. Daher wurden in der Trainingsstudie Maßnahmen zur Förderung der Theorie-Evidenz-Koordination im Hinblick auf ihre Effektivität geprüft.

Die Studie analysierte Lernprozesse einzelner Kinder in Anlehnung an ein mikrogenetisches Design vor und nach dem Training im Rahmen eines Prä-Post-Designs. In der vorliegenden Studie wurde ermittelt, wie sich der Kompetenzzuwachs im Hinblick auf die Schlussfolgerungskompetenzen kurzfristig im Post-Test und auch nachhaltig im Transfer-Test in beiden Trainingsgruppen im Vergleich zur Kontrollgruppe darstellte. Dazu wurden die drei Experimentalgruppen (1. Trainingsgruppe, 2. Trainingsgruppe, Kontrollgruppe) im Post-Test miteinander verglichen und geprüft, ob sich die Trainingsgruppen im Hinblick auf den Lernzuwachs deutlich von der Kontrollgruppe unterscheiden. Zudem wurde untersucht, welche Gruppe in den beiden Tests (Post-Test, Transfer-Test) nach dem Training einen höheren Lernzuwachs aufgewiesen hat. Die Häufigkeiten angemessener Antworten im Vergleich untereinander machten deutlich, welche Gruppe effektiver trainiert worden ist. Der Trainingsstudie wurde eine

Vorstudie vorgeschaltet, da die neuentwickelten Fördermaßnahmen im Hinblick auf ihre Durchführbarkeit getestet wurden.

## 7 Querschnittstudie

Die Querschnittstudie zielte darauf ab, die Kompetenzen beim Schlussfolgern bei Kindern in vier Altersstufen (vier bis zehn Jahren) im Elementar- und Primarbereich festzustellen (Stephan-Gramberg, Robisch, Löhner, Hardy, Tröbst & Möller, 2012), um Entwicklungen kindlichen Denkens über einen Zeitraum von sechs Jahren darzustellen.

### 7.1 Vorstudie

In der Vorstudie der Querschnittstudie wurden die Schlussfolgerungskompetenzen sowie die wissenschaftlichen Kompetenzen der Probanden in den unterschiedlichen Altersstufen untersucht. Die Vorstudie wurde durchgeführt, um herauszufinden, wie die Testinstrumente (Schlussfolgerungstest, Wissenschaftsverständnistest) bei Kindern im Elementar- und Primarbereich funktionieren. Es wurde genau darauf geachtet, ob die Probanden die Formulierungen bei den Vermutungen im Schlussfolgerungstest verstanden. Zudem wurde geschaut, ob der neu entwickelte Wissenschaftsverständnistest in allen Altersstufen (vier bis zehn Jahre) sinnvoll durchführbar war.

#### 7.1.1 Fragestellungen und Hypothesen

Das zentrale Anliegen der Querschnittstudie war es, die Voraussetzungen und Kompetenzen beim Schlussfolgern junger Kinder im Inhaltsgebiet *Elastizität und Plastizität* im Elementar- und Primarbereich zu identifizieren. Dazu wurde die Entwicklung beim domänenspezifischen Denken, dem schlussfolgernden Denken mittels der *Truth Testing Task* (Wason, 1966) bei Kindern aus vier Altersstufen (vier Jahre, fünf bis sechs Jahre, sieben bis acht Jahre, neun bis zehn Jahre) untersucht. Entwicklungsspezifische kognitive Aspekte, wie beispielsweise die Fähigkeit zur Koordination von Theorie und Evidenz, die Kompetenzen zum Schlussfolgern, die Kapazität des Arbeitsgedächtnisses ebenso wie die Inhibition wurden in den unterschiedlichen Altersstufen in den Blick genommen. Aufgrund empirischer Befunde ist davon auszugehen, dass sich die Kompetenz zum Schlussfolgern mit zunehmendem Alter verbessert. (Barrouillet et al., 2008; Tröbst et al., 2011). Daraus ergaben sich folgende Forschungsfragen und Hypothesen für die Querschnittstudie.

Frage 1a: Welche Kompetenzen zeigen Kinder beim Evaluieren von Einzelereignissen in den genannten Altersstufen?

Hypothese 1a: Junge Kinder im Kindergartenalter schlussfolgern bei den irrelevanten Ereignissen  $[\neg p \ q]$ ,  $[\neg p \neg q]$  signifikant schlechter als Grundschul Kinder.

Frage 1b: Welche Kompetenzen zeigen Kinder im Kindergartenalter beim Schlussfolgern auf den Kompetenzniveaus (konjunktiv, bi-konditional, konditional, RFRF, RFFR, Sonstige) im Vergleich zu Grundschulkindern?

Hypothese 1b: Junge Kinder im Kindergartenalter schlussfolgern auf dem Niveau „Sonstige“, während Neun- bis Zehnjährige auf dem höheren Niveau, dem konjunktiven oder bikonditionalen Niveau antworten (Beschreibung der Kompetenzniveaus siehe Kapitel 5.1). Die Kompetenzniveaus werden bei der Berechnung als abhängige Variable genutzt.

Die zweite Frage zielte auf die Überprüfung der Kompetenzen beim Wissenschaftsverständnis:

Frage 2: Welche Kompetenzen zeigen Kinder bei der Koordination von Theorie und Evidenz im Wissenschaftsverständnis?

Hypothese 2: Junge Kinder zeigen weniger gut ausgeprägte Fähigkeiten beim Wissenschaftsverständnis als ältere Kinder, daher erzielen sie ein geringeres Kompetenzniveau beim Test als ältere Kinder.

### 7.1.2 Design

Das Design der Vorstudie umfasste zwei Tests, den Wissenschaftsverständnis test sowie den Schlussfolgerungstest. Beide Erhebungen wurden an einem Tag durchgeführt.

### 7.1.3 Analysestrategien bei der Auswertung

Im folgenden Unterkapitel werden die Analysestrategien bei der Auswertung der Daten im Zusammenhang mit den Fragestellungen beschrieben. Die Vorstudie prüfte, ob der Schlussfolgerungstest sowie der Wissenschaftsverständnis test sich als Testinstrumente in ihrer Durchführung sinnvoll waren. Um der Frage der Kompetenzen der Kinder in den vier Altersstufen (vier, fünf bis sechs, sieben bis acht und neun bis zehn Jahre) nachzugehen, wurden auf deskriptiver Ebene die relativen Häufigkeiten sowie die Mittelwerte (mit Standardabweichungen) dargestellt.

Die Einzelereignisse ( $[p \ q]$ ,  $[p \neg q]$ ,  $[\neg p \ q]$ ,  $[\neg p \neg q]$ ) dienen als Grundlage zur Bildung der Interpretationsniveaus (konditional, bikonditional, konjunktiv, RFFR, RFRF, Sonstige), die

in Kapitel 5.1 beschrieben wurden. Die Interpretationsniveaus wurden nach dem Entwicklungsmodell gebildet (Barrouillet et al., 2008), das drei aufeinanderfolgende Entwicklungsstufen in Bezug auf die Gültigkeit konditionaler Aussagen umfasst (siehe Kapitel 5.1). Das konditionale Niveau stellt die angemessene Schlussfolgerungsweise dar, gefolgt vom bikonditionalen und dem konjunktiven Niveau. Beim konditionalen Interpretationsniveau, dem bestmöglichen Niveau, wird das Antezedens als ein möglicher Grund der Konsequenz verstanden (Tröbst et al., 2011b). Beim bikonditionalen Interpretationsniveau wird das Antezedens als einziger und ausschließlicher Grund für die Konsequenz aufgefasst. Beim konjunktiven Verständnisniveau verknüpfen Kinder beim Schlussfolgern das Antezedens (z.B. Eigenschaft des Gegenstandes) mit der Konsequenz (z.B. Sprungverhalten), indem sie diese logisch durch ein „und“ miteinander verbinden.

Zusätzlich sind zwei Interpretationsniveaus, das RFFR- und das RFRF-Niveau (siehe Tabelle 7) bei jungen Grundschulkindern gefunden worden (Tröbst et al., 2011b). Beim RFFR-Niveau überprüfen Kinder die gegebene Aussage nicht, sondern gehen davon aus, dass diese gültig ist. Daher antworten sie regelorientiert. Bei diesem Niveau können sie zwischen Theorie und Evidenz noch nicht differenzieren. Beim RFRF-Niveau fokussieren Kinder in ihrer Schlussfolgerung auf das Sprungverhalten [q]. Sie bestätigen eine Vermutung für einen Gegenstand, wenn die zweite Eigenschaft (der Ball springt) zutrifft.

Tabelle 7

*Erweiterte Interpretationsniveaus*

	RFFR	RFRF	1. Stufe konjunktiv	2. Stufe bikondi-tio- nal	3. Stufe konditional
[p q]	+	+	+	+	+
[p¬q]	-	-	-	-	-
[¬p q]	-	+	-	-	○
[¬p¬q]	+	-	-	○	○

*Anmerkung.* + = Hypothese bestätigende Aussage, - = Hypothese widerlegende Aussage, ○ = für Hypothese irrelevante Aussage.

Folglich bestätigen Kinder auf diesem Niveau eine Vermutung, wenn der Gegenstand springt; springt dieser nicht, gehen sie davon aus, dass damit die Vermutung widerlegt wird.

Auf Grundlage der Fähigkeiten der Probanden beim Schlussfolgern wurden die Interpretationsniveaus gebildet.

Eine multivariate Analyse, *MANOVA*, bei der mehrere abhängige Variablen gleichzeitig miteinander verglichen werden, wurde zuerst durchgeführt. Die *MANOVA* prüfte, ob signifikante Unterschiede im Hinblick auf die Interpretationsniveaus beim Schlussfolgern zwischen den Altersgruppen vorlagen.

Beim Wissenschaftsverständnistest wurden die Aussagen der Kinder mit Hilfe eines Kodierschemas (siehe Anhang B\_4) kodiert und von zwei unabhängigen Ratern bewertet. Die Höhe der Übereinstimmung der Rater wurde mit der Ermittlung des Kappa-Wertes vorgenommen. Mit der Vorstudie wurde geprüft, ob die eingesetzten Tests (Schlussfolgerungstest, Wissenschaftsverständnistest) funktionierten bzw. inwiefern noch Änderungen vorgenommen werden sollten.

#### 7.1.4 Methoden

Im folgenden Kapitel wird der Schlussfolgerungstest nach der *Truth Testing Task* (Wason, 1966) mit vier Vermutungen aufgeführt. Zusätzlich wurde in der Querschnittstudie ein Wissenschaftsverständnistest verwendet, um die Fähigkeiten der Theorie-Evidenz-Koordination in den einzelnen Altersstufen zu ermitteln. Weiterhin wurden exekutive Funktionen als Kontrollvariablen erhoben. Der Inhibitionstest (Früchte Gemüse Stroop) wird in diesem Kapitel vorgestellt (Jansen, Mannhaupt, Marx, Skowronek, 1999). Auf die anderen Tests (Wortlistentest, Farb-Benennung-Rückwärts) wird nur verwiesen (ebenda), da sie nicht so zentral wie der Inhibitionstest in der vorliegenden Studie sind. Mit der Vorstudie wurden die Instrumente im Hinblick auf Durchführbarkeit und Sensibilität im Zusammenhang mit den drei Altersgruppen (vierjährige, sieben- bis achtjährige und neun- bis zehnjährige Probanden) überprüft. Die Erhebungen dazu wurden von jeweils einer wissenschaftlichen Mitarbeiterin sowie einer wissenschaftlichen Hilfskraft oder zwei wissenschaftlichen Hilfskräften durchgeführt. Der Vorwisentest wird in der Vorstudie nicht beschrieben, da er bei den Forschungsfragen nicht berücksichtigt wurde. Eine detaillierte Beschreibung ist nachzulesen bei Robisch (2012).

##### *Schlussfolgerungstest*

In Anlehnung an die *Truth Testing Task* (Wason, 1966; Barrouillet et al., 2008; Tröbst et al., 2011) wurde der Schlussfolgerungstest durchgeführt. Dieser wurde sowohl in der Querschnittstudie als auch in der Trainingsstudie eingesetzt. Im Kindergarten beantworteten Kinder die



Fragen des Versuchsleiters im Rahmen von Einzelsitzungen; in der Schule schrieben die Schüler die Antworten in Form eines Gruppentestes in ein Testheft. Der Test wurde sowohl im Kindergarten als auch in der Schule mit Hilfe eines Interviewleitfadens durchgeführt, um vergleichbare Versuchsbedingungen herzustellen.

In der Vorstudie überprüften alle Probanden vier Vermutungen (siehe Tabelle 8). Mit einem Gegenstand konnte eine Vermutung bestätigt oder widerlegt werden bzw. konnte sie ebenso irrelevant im Zusammenhang mit derselben sein.

Tabelle

8

*Vier Vermutungen*

Vermutung	Gegenstände	bestätigend [p q]	widerlegend [p¬q]	irrelevant [¬p q]	irrelevant [¬p¬q]
Alle runden Bälle springen.	Tischtennisball	x			
	Tonkugel		x		
	Flummie			x	
	Luftballon mit Sand gefüllt				x
Alle Bälle mit Luft springen.	Spielball	x			
	Hohler Fensterkittball		x		
	Eisenkugel			x	
	Luftballon ohne Luft				x
Alle leichten Bälle springen.	Styroporkugel	x			
	Wattebausch		x		
	Bocciakugel			x	
	Körnerkissen				x
Alle weichen Bälle springen.	Softball	x			
	Murmel		x		
	Jonglierball			x	
	Knetkugel				x

Ein Teilnehmer probierte aus, ob der Gegenstand sprang oder nicht; im Anschluss daran evaluierte er diesen im Zusammenhang mit der zuvor genannten Vermutung; Beispielsweise wurde die Vermutung „Alle runden Bälle springen“ (siehe Tabelle 8) genannt. Die Antwortmöglichkeiten wurden den Probanden jeweils im Anschluss präsentiert. Dazu wurden drei Smileys in unterschiedlichen Farben (grün, gelb, rot) verwendet (siehe Tabelle 9):

„Du hast drei Möglichkeiten zu antworten. Wenn du glaubst, Tims Vermutung stimmt, legst du den Spielball auf den grünen Smiley.“ (Auf den Smiley zeigen) „Wenn du glaubst, Tims Vermutung hat damit nichts zu tun, legst du den Spielball auf den gelben Smiley.“ (Auf den Smiley zeigen) (siehe Anhang A 4.) „Wenn du glaubst, Tims Vermutung stimmt nicht, legst du den Spielball auf den roten Smiley.“ (Auf den Smiley zeigen) (siehe Anhang A 2).

In gleicher Weise wurden alle vier Gegenstände im Hinblick auf eine Vermutung (siehe Tabelle 9) geprüft und dem jeweiligen Smiley zugeordnet. Die Probanden bekamen keinerlei Rückmeldung in Bezug auf die Richtigkeit ihrer Aussage.

Tabelle 9

*Zuordnung der Antworten zu Symbolen*

Symbol	grüner Smiley	roter Smiley	gelber Smiley	gelber Smiley
Antwort	bestätigend	widerlegend	irrelevant	irrelevant
	[p q]	[p¬q]	[¬p q]	[¬p q]

Die vier Antworten zu einer Vermutung wurden genutzt, um Interpretationsniveaus zu bilden. Eine ausführliche Beschreibung der Interpretationsniveaus ist in Kapitel 5.1 zu finden. Das Antwortverhalten auf den einzelnen Niveaus ist für die Vermutung „Alle runden Bälle springen“ in Tabelle 10 zusammengestellt.

Auf dem konjunktiven Niveau sieht ein Kind den Tischtennisball als Bestätigung der Vermutung „Alle runden Bälle springen“ und die übrigen Bälle als Widerlegung an. Ein Kind auf dem bikonditionalen Niveau sieht den Tischtennisball als Bestätigung der Vermutung an, die Tonkugel sowie das Flummieci widerlegen diese aus seiner Sicht, während es den mit Sand gefüllten Luftballon als irrelevant wertet. Ein Kind antwortet auf dem bestmöglichen, dem konditionalen Interpretationsniveau, wenn zusätzlich zum Luftballon das Flummieci als irrelevant im Zusammenhang mit der oben genannten Vermutung gesehen wird. In der Studie von Tröbst et al. (2011) sind zwei zusätzliche Niveaus gefunden worden. Ein Kind agiert auf dem RFFR-

Niveau<sup>3</sup> (Tröbst et al., 2011), wenn es den Tischtennisball und den Luftballon mit Sand als Bestätigung, die Tonkugel und das Flummieci zur Widerlegung der Vermutung nutzt (siehe Tabelle 10). Beim RFRF-Niveau<sup>4</sup> fokussiert das Kind auf das Sprungverhalten und ordnet deswegen Tischtennisball und Flummieci als bestätigend und Tonkugel und mit Sand gefüllten Luftballon als widerlegend ein (siehe Tabelle 10). Die beschriebenen Kodierungen der Interpretationsniveaus waren Grundlage für die Analysen in der Querschnittstudie sowie in der Trainingsstudie. Beantwortete ein Kind einer Frage nicht, wurde dies als fehlender Wert aufgenommen und in SPSS entsprechend kodiert und in den Analysen nicht berücksichtigt. In der Querschnittstudie ist dieser Fall nur bei zwei Kindern im Vorschulalter aufgetreten. Diese wurden in der Berechnung ausgeschlossen. In der Grundschule gab es keinen fehlenden Wert. Die Kinder haben bei allen Aufgaben eine Antwort gegeben.

Tabelle 10

*Kodierung nach Interpretationsniveau*

Vermutung: Alle runden Bälle springen				
Antworten auf dem konjunktiven Niveau				
	bestätigend [p q]	widerlegend [p¬q]	irrelevant [¬p q]	irrelevant [¬p¬q]
Tischtennisball	x			
Tonkugel		x		
Flummieci		x		
Luftballon mit Sand gefüllt		x		
Vermutung: Alle runden Bälle springen				
Antworten auf dem bikonditionalen Niveau				
	bestätigend [p q]	widerlegend [p¬q]	irrelevant [¬p q]	irrelevant [¬p¬q]
Tischtennisball	x			
Tonkugel		x		
Flummieci		x		

<sup>3</sup> RFRF-Niveau = Evaluation von vier Ereignissen im Zusammenhang mit einer Hypothese: Richtig-Falsch-Falsch-Richtig.

<sup>4</sup> RFRF-Niveau = Evaluation von vier Ereignissen im Zusammenhang mit einer Hypothese: Richtig-Falsch-Richtig-Falsch.

Luftballon mit Sand gefüllt					x
Vermutung: Alle runden Bälle springen					
Antworten auf dem konditionalen Niveau					
	bestätigend [p q]	widerlegend [p¬q]	irrelevant [¬p q]	irrelevant [¬p¬q]	
Tischtennisball	x				
Tonkugel		x			
Flummie			x		
Luftballon mit Sand gefüllt					x
Vermutung: Alle runden Bälle springen					
Antworten auf dem RFFR-Niveau					
	bestätigend [p q]	widerlegend [p¬q]	irrelevant [¬p q]	irrelevant [¬p¬q]	
Tischtennisball	x				
Tonkugel		x			
Flummie		x			
Luftballon mit Sand gefüllt	x				
Vermutung: Alle runden Bälle springen					
Antworten auf dem RFRF-Niveau					
	bestätigend [p q]	widerlegend [p¬q]	irrelevant [¬p q]	irrelevant [¬p¬q]	
Tischtennisball	x				
Tonkugel		x			
Flummie	x				
Luftballon mit Sand gefüllt		x			

### *Wissenschaftsverständnistest*

Der Wissenschaftsverständnistest diente der Überprüfung der Fähigkeiten zur Theorie-Evidenz-Koordination bei Kindern von vier bis zehn Jahren. Der Test wurde in Anlehnung an die Aufgaben zum Wissenschaftsverständnis aus einem Projekt zur Entwicklung naturwissenschaftlicher Kompetenz (Koerber, Kropf, Mayer, Sodian & Schwippert, 2009) entwickelt. Zwei der drei Kontexte im neu konzipierten Wissenschaftsverständnistest sind vom Projekt Science-

P adaptiert (ebenda) worden. Im Test wurden alltagsnahe Kontexte benutzt, um die Fähigkeiten der Kinder in Verbindung mit Situationen aus ihrer Lebenswelt zu prüfen (siehe Anhang B\_2; Anhang B\_3). Der Fokus bei der Testentwicklung lag darauf, Fähigkeiten sowohl von Kindern aus dem Elementar- als auch aus dem Primarbereich bei der Unterscheidung von Theorie und Evidenz zu ermitteln. Dazu dienten sowohl Aufgaben zur Produktion von Theorie und Evidenz als auch zum Erkennen von Evidenz. Insgesamt wurden vier Bereiche erfasst: das eigenständige Produzieren von Evidenz und Theorie sowie das Erkennen von Evidenz und Theorie. Empirische Befunde (Koerber et al., 2009) zeigten, dass das Produzieren von Evidenz bzw. Theorie Kindern schwerer fiel als das Erkennen von Evidenz bzw. Theorie.

Der Wissenschaftsverständnistest umfasste drei Bildergeschichten aus dem Alltag von Kindern. Die Geschichten setzten sich aus jeweils vier Bildern zusammen. Entsprechend der vier Aufgabenbereiche (Produzieren von Evidenz - Erkennen von Evidenz - Produzieren von Theorie - Erkennen von Theorie) enthielt jeder Kontext vier Fragen, wobei jedes Bild auf eine Facette abzielte (z.B. auf das Produzieren von Evidenz): Für jede richtige Antwort gab es einen Punkt, sodass bei vier richtigen Antworten und drei Kontexten insgesamt maximal zwölf Punkte erreicht werden konnten.

Zu keinem Zeitpunkt des Tests erhielten die Probanden Unterstützung oder Rückmeldung auf ihre Antworten. Zu Beginn der Interviews lagen nur Bild eins und Bild vier offen vor den Probanden (siehe Abbildung 11), während Bild zwei und Bild drei verdeckt waren. Dieses Vorgehen wurde gewählt, um einem Kind die Produktion von Evidenz zu ermöglichen.

Ein Proband erhielt die Aufgabe, sich zunächst in die Lage der Hauptperson der Geschichte hineinzusetzen und mögliche Belege für einen situationsbezogenen Inhalt des Kontextes zu nennen.

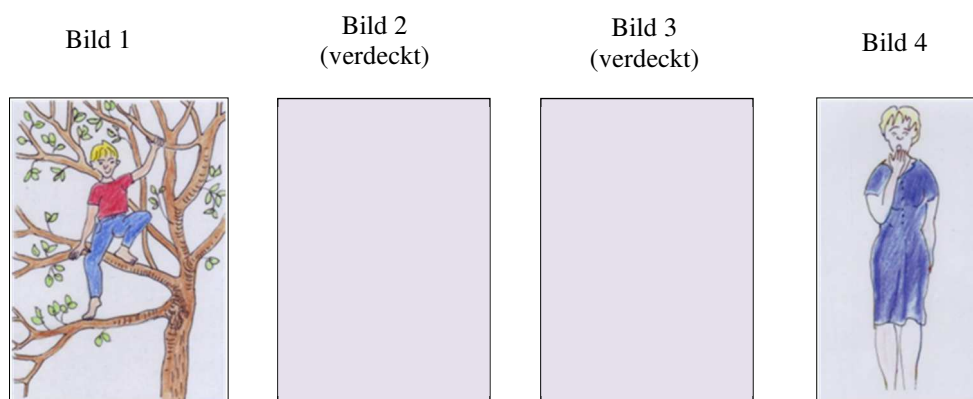


Abbildung 11. Produktion von Evidenz.

Die Aufgabe zur ersten Geschichte („Marius auf dem Baum“) wird im Folgenden exemplarisch vorgestellt. Die Hauptpersonen der ersten Geschichte waren ein Junge, Marius, und seine Mutter. Die Kinder sollten sich bei dieser Aufgabe zunächst in die Lage der Mutter hineinversetzen und anhand des ersten Bildes - Marius auf dem Baum - und des vierten Bildes der Geschichte - die erschreckt blickende Mutter - überlegen, welchen sichtbaren Beleg es für Marius Sturz geben hätte geben können (siehe Abbildung 11); das heißt, sie sollten auf Basis eigener Erfahrungen Evidenz produzieren.

Im nächsten Schritt wurde Bild drei aufgedeckt (siehe Abbildung 12); auf diesem Bild saß Marius auf dem Boden neben einem abgebrochenen Ast. Hier sollten die Kinder Evidenz erkennen und benennen. Als Grundlage für das passive Erkennen der Evidenz diente Bild 3, denn dort waren verschiedene Belege für die Verletzung des Jungen erkennbar (Tränen, zerrissene Hose, Blutfleck, abgebrochener Ast).

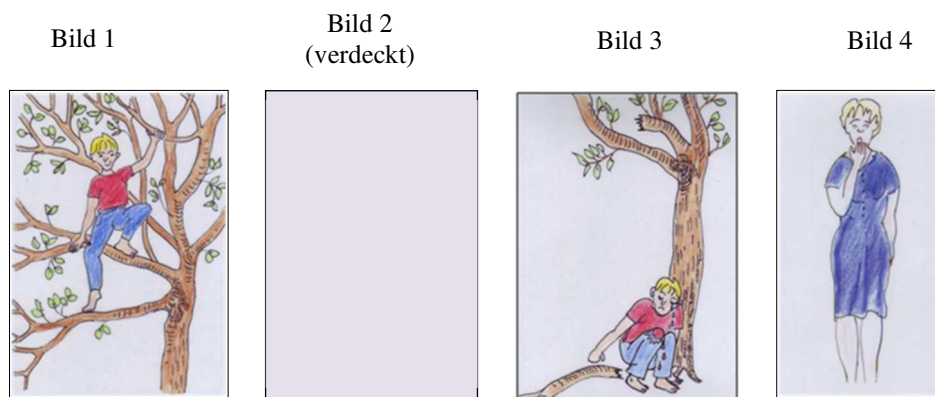


Abbildung 12. Erkennen von Evidenz.

Direkt im Anschluss an diese Aufgabe sollte ein plausibler Grund für die Verletzung des Jungen gefunden werden. Die Kinder hatten die Aufgabe, eine sinnvolle Theorie zu produzieren. Zum Schluss wurde das vierte Bild - mit Marius auf einem abbrechenden Ast - aufgedeckt (siehe Abbildung 13). Mithilfe des Bildes sollten die Kinder die Theorie interpretieren und den Grund für den Sturz nennen.

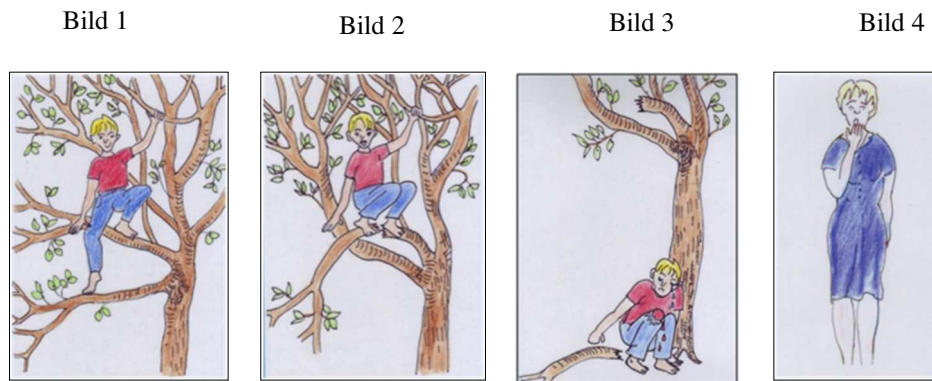


Abbildung 13. Interpretation von Theorie

Die zuvor beschriebene Form der Durchführung des Wissenschaftsverständnis-tests wurde gewählt, um die Kinder im Rahmen eines Kontextes von der Produktion hin zum Erkennen von Evidenz bzw. Produktion von Theorie hin zum Erkennen von Theorie zu führen. Im zweiten und dritten Kontext wurden die Aufgaben in gleicher Weise und Reihenfolge wie im ersten Kontext gestellt.

#### *Auswertung des Wissenschaftsverständnis-tests*

Eine Kodierung zum Wissenschaftsverständnis-test aus dem Kontext „Marius“ ist beispielhaft aufgeführt (siehe Tabelle 11). Antworten wurden mit 0 kodiert, sofern sie die Frage nicht zielführend beantworteten; mit 1 wurden Antworten kodiert, die zeigten, dass Kinder den Beleg für Marius Sturz vom Baum benennen konnten. Ein Kodierschema für alle 12 Items des Wissenschaftsverständnis-tests wurde entwickelt und bildete die Grundlage für die Kodierungen der Antworten (siehe Anhang B\_4).

Tabelle 11

#### *Beispiel der Kodierung zu Frage 1 Kontext Marius*

	0	1
Frage 1: Was glaubst du, woran erkennt die Mutter, dass Marius beim Klettern runtergefallen ist?	<p>...weil er sich nicht festhält. ...vom Baum. ...dass der runtergefallen ist. ...das kann ich nicht sehen. ...weil sie es gesehen hat. ...weil sie weiß, dass da Matsch ist. ...weil er auf den Baum geklettert ist.</p>	<p>...weil er auf dem Boden liegt. ...an Kratzern, Schürfwunden. ...weil ein Loch in der Hose ist. ...weil er weint. ...weil seine Hose zerrissen ist.</p>

0	1
...weil der Baum wackelt.	...hat sich wehgetan, sitzt unter Baum.

Die Antworten der Probanden wurden dichotom (0 ⇔ falsch; 1 ⇔ richtig) auf Basis des festgelegten Kodierschemas (siehe Anhang B\_4) kodiert. Anschließend wurden die Antworten von zwei unabhängigen Beobachtern geratet. Aus den zwei Werten zu jeder Antwortkategorie wurde mithilfe einer Kreuztabelle die Übereinstimmung von verschiedenen Beobachtern desselben Sachverhaltes berechnet.

Im folgenden Absatz wird erläutert, wie zuverlässig eine Übereinstimmung von zwei Beobachtern sein kann. Als einfaches Maß kann man den Anteil an Übereinstimmungen (Janssen & Laatz, 2013, S. 281) von Beobachtungen an der gesamten Zahl der Vergleiche berechnen. M gibt die Zahl der Übereinstimmungen an, N steht für die Zahl der Vergleiche.

$$\ddot{U} = \frac{M}{N}$$

Daraus lässt sich der Kappa-Wert mit folgender Gleichung berechnen,  $\ddot{U}$  bildet dabei den Anteil der beobachteten Übereinstimmungen.

$\ddot{U}_E$  den Anteil der erwarteten Beobachtungen (ebenda, S. 281). Der Anteil der erwarteten Übereinstimmungen lässt sich mit folgender Formel berechnen:

$$\ddot{U}_E = \sum_{i=1}^k (p_i)^2$$

Der Kappa-Wert wird mit folgender Formel berechnet (Janssen & Laatz, 2013, S. 281):

$$\text{kappa} = \frac{\ddot{U} - \ddot{U}_E}{1 - \ddot{U}_E}$$

Der Kappa-Wert berechnet sich aus dem Quotienten der Differenz der Übereinstimmung und der erwarteten Übereinstimmung ( $\ddot{U} - \ddot{U}_E$ ) sowie der Differenz von 1 und der erwarteten Übereinstimmung ( $1 - \ddot{U}_E$ ). Mit der Formel des Kappa-Wertes wird der Zusammenhang von wirklicher und erwarteter Übereinstimmung berechnet (Janssen & Laatz, 2013). Der Kappa-Wert gibt einen Nachweis für die Reliabilität (Zuverlässigkeit) einer gebildeten Skala (ebenda, S. 280) an.

*Item-Response-Theorie*



Das Rasch-Modell ist ein probabilistisches Modell, das der Item-Response-Theorie (IRT<sup>5</sup>) angehört (Moosbrugger & Kevala, 2007). Es wurde in der vorliegenden Arbeit zur Analyse des Wissenschaftsverständnistests eingesetzt. Das Rasch-Modell (Rasch, 1960; Rost, 2004) wird im Folgenden kurz erläutert werden. Im Modell wird davon ausgegangen, dass die Antworten auf Testitems Indikatoren für dahinter liegende Variablen (= latente Variablen) darstellen. Diese werden durch Items mit dichotomem Antwortformat (gelöst, nicht gelöst) bestimmt.

Dabei wird eine beobachtete Antwort als manifeste Variable genutzt, um daraus eine dahinterliegende, latente Variable zu bilden. Diese stellt daher eine Merkmalsausprägung einer nicht direkt beobachtbaren Fähigkeit einer Person dar (Moosbrugger & Kevala, 2007) und wird auf Basis manifester, dichotomer Variablen in Form von Antworten auf Items (gelöst, nicht gelöst) geschätzt. Voraussetzung für die Durchführung des Rasch-Modells stellt das Vorhandensein der lokalen stochastischen Unabhängigkeit dar. Sie ist gegeben, wenn keine signifikanten Korrelationen zwischen Antworten innerhalb von latenten Variablen vorliegen (Kolen & Brennan, 2004). Eine weitere zentrale Voraussetzung ist die Eindimensionalität, was bedeutet, dass beim dichotomen Rasch-Modell alle Variablen auf einer latenten Variablen zusammengefasst werden (Kolen & Brennan, 2004).

Grundgedanke des Modells von Rasch (Rasch, 1960; Rost, 2004) ist, dass eine Person (=  $v$ ) mit einer spezifischen Fähigkeit (=  $\theta$ ) eine bestimmte Lösungswahrscheinlichkeit zum Lösen einer Aufgabe hat. Berechnet wird die Lösungswahrscheinlichkeit durch Differenzbildung zwischen Personenfähigkeit (=  $\theta$ ) und Itemschwierigkeit  $\sigma_i$  als logistische Funktion.

Nach der folgenden Modellgleichung wird die Lösungswahrscheinlichkeit berechnet (Rost 2004, S. 119):

$$p(x_{vi}) = \frac{\exp(x_{vi}(\theta_v - \sigma_i))}{1 + \exp(\theta_v - \sigma_i)}$$

Dabei stellt  $x_{vi}$  die Wahrscheinlichkeit einer Person  $v$  das Item  $i$  zu lösen dar. Sie berechnet sich aus dem Quotienten der Exponentialfunktion der Differenz von Personenfähigkeit und Itemschwierigkeit.

Die Funktionsgleichung kann graphisch als Item Charakteristik-Kurve (= ICC) dargestellt werden. Dabei werden die Ausprägungen der latenten Variablen auf der Abszisse und die Lösungswahrscheinlichkeit auf der Ordinate dargestellt (Moosbrugger & Kevala, 2007).

---

<sup>5</sup> Item = Aufgabe in einem Test.

Die Trennschärfeparameter sind bei der eindimensionalen Raschanalyse für alle Items in einem Test gleich, daher verlaufen Item Charakteristik-Kurven (=ICC) alle parallel zueinander und sind monotonsteigend (Moosbrugger & Kevala, 2007). Dabei bestimmen die beiden Modellparameter, Personenfähigkeit und Itemschwierigkeit, den Verlauf der Funktionskurve.

Die beiden Modellparameter werden beim Rasch-Modell auf einer Skala in Logit-Einheiten abgebildet (Wilson, 2005). Die Skala wird in einer Abbildung, auch Wright-Map (Wilson, 2005) genannt, dargestellt. Dabei sind die Antworten der Personen entsprechend ihrer Fähigkeit auf der linken Seite und die Items, geordnet nach der Schwierigkeit, auf der rechten Seite der Wright Map dargestellt. Die Weighted Likelihood Estimation (= WLE) ist ein Parameter, mit dem die Fähigkeit einer Person eingeschätzt wird (Warm, 1989). Mit den beiden Modellparametern, Personenfähigkeit und Itemschwierigkeit kann die Lösungswahrscheinlichkeit eines Items vorhersagt werden.

Aus dem Zusammenhang von Personenfähigkeit und Itemschwierigkeit wird die individuelle Lösungswahrscheinlichkeit einer Person berechnet. Dabei wird der Schwierigkeitsparameter  $\sigma_i$  als jene Merkmalsausprägung  $\theta_v$  bezeichnet, bei der die Lösungswahrscheinlichkeit für ein Item genau 50% beträgt (Moosbrugger & Kevala, 2007). Das heißt, wenn die Itemschwierigkeit  $\sigma$  gleich der Personenfähigkeit  $\theta$  ist, dann kann eine Person ein Item mit einer Wahrscheinlichkeit von 50% lösen. An dieser Stelle sind Personenfähigkeit und Itemschwierigkeit gleich. Liegt die Personenfähigkeit höher als die Schwierigkeit des Items, so wird die Lösungswahrscheinlichkeit des Probanden höher als 50% sein (Wilson, 2005); liegt die Personenfähigkeit niedriger als die Schwierigkeit des Items, dann ist die Lösungswahrscheinlichkeit geringer als 50%. Die Ausprägungen auf der latenten Variablen werden durch Personenparameter und Itemschwierigkeit geschätzt. Mit der Maximum-Likelihood-Methode werden die beobachteten Daten unter bestimmten Modellannahmen durch einen iterativen Vorgang geschätzt. Je höher der Wert der Likelihood ist, desto wahrscheinlicher werden die empirischen Daten mit den geschätzten Werten zueinander passen (Moosbrugger & Kevala, 2007). Die Parameter werden auf Basis der geltenden Modellannahmen so geschätzt, dass die beobachteten Daten möglichst gut zum Modell passen.

Der Item-Fit oder *Mean Weighted Square (MNSQ)* zeigt, ob die empirischen Daten mit den geschätzten Werten übereinstimmen. Mit dem Programm Conquest (Wu, Adams & Wilson, 1998) wird die Lösungswahrscheinlichkeit für jedes Item und jede Person berechnet. Liegen die Werte für den *MNSQ* zwischen 0,8 und 1,20, befinden sie sich in einem akzeptablen Rahmen (Wu et al., 1998). Bond & Fox (2007) hingegen geben einen Wertebereich von 0.75 bis 1.33 an. Eine exakte Übereinstimmung zwischen beobachteten bzw. manifesten und

geschätzten bzw. latenten Variablen liegt bei einem *Mean Weighted Square* von 1 vor. Bei Werten größer als 1 ist die Varianz zwischen beobachteten Daten und geschätzten Lösungshäufigkeiten höher, als durch die Modellschätzung vorhergesagt worden ist. In diesem Fall resultieren die Werte in einer flachen Kurve; bei Werten kleiner als 1 ist die Varianz zwischen beobachteten Daten und geschätzten Lösungshäufigkeiten geringer, als durch die Modellgleichung vorhergesagt worden ist, und es ergibt sich eine steile Kurve.

Zusätzlich werden für jeden *Mean Weighted Square* auch T-Werte als Prüfwerte angegeben. Sind die T-Werte größer als +2 oder kleiner als - 2, liegt eine signifikante Abweichung zwischen beobachteten und geschätzten Werten vor. Weiterhin wird die EAP/PV-Reliabilität<sup>6</sup> als erwarteter Schätzwert (Rost, 2004) für die individuellen Messwerte angegeben. Sie gibt, ähnlich wie die Reliabilität von Cronbachs Alpha, Auskunft über die Zuverlässigkeit eines Tests. Ein weiterer Wert ist die punktbiseriale Korrelation, die den Zusammenhang eines Items mit der Gesamtskala darstellt (Rost, 2004). Mit den erläuterten Werten, Lösungswahrscheinlichkeit, MNSQ, WLE sowie der Korrelation, kann die Passung zwischen einzelnen Aufgaben und Probanden ermittelt werden.

#### *Aufgaben zur Kontrolle exekutiver Funktionen*

Die Aufgaben zur Kontrolle exekutiver Funktionen (Wortliste, Test zur Farb-Benennung-Rückwärts) sowie der Inhibitionstest wurden in der Vorstudie noch nicht eingesetzt, da es dort ausschließlich um die Erprobung der beiden Testinstrumente (Schlussfolgerungstest, Wissenschaftsverständnistest) ging.

Die beiden Aufgaben zur Kontrolle exekutiver Funktionen, der Wortliste-Test (*word list recall task*) sowie der Test zur Farb-Benennung Rückwärts (*backward colour recall task*), wurden aus der Studie von Zoelch, Seitz & Schumann-Hengsteler (2005) adaptiert (siehe Anhang\_E2; siehe Anhang C\_1). Die Reliabilität in beiden Tests war gut (siehe Zoelch et al., 2005 für eine detaillierte Beschreibung).

Die Aufgaben zur Inhibition wurden aus Jansen et al. (1999) übernommen. Die Fähigkeit zur Inhibition wurde mittels einer *Früchte-Gemüse Stroop-Aufgabe* (Jansen et al., 1999) erhoben. Der Test wurde im Rahmen des Bielefelder Screenings, eines Verfahrens zur Früherkennung von Lese-Rechtschreibschwierigkeiten bei Vorschulkindern, entwickelt. Hierbei wurden dem Kind im Rahmen der Einzeltestung falsch eingefärbte Bilder von Obst- und

---

<sup>6</sup> EAP/PV Reliabilität = erwartete (expected a priori) Reliabilität.

Gemüsesorten vorgelegt. Insgesamt enthielt der Test 28 Bilder, je vier Bilder in sieben aufeinanderfolgenden Reihen. Die erste Reihe wurde dem teilnehmenden Kind als Beispiel erklärt; dazu nannte der Versuchsleiter die realitätsgetreuen (und nicht die abgebildeten) Farben der ersten Reihe. Das Kind hatte im Anschluss die Möglichkeit, Verständnisfragen zu stellen. Nach dieser Erklärungsphase des Versuchsleiters wiederholte das Kind die erste Reihe zur Einführung. Diese wurde bei der Bewertung nicht berücksichtigt. Die Aufgabe für den Probanden bestand darin, bei den darauffolgenden sechs Reihen so schnell wie möglich selbstständig die korrekte Farbe der jeweiligen Obst- oder Gemüsesorte zu nennen; so wie er diese beispielsweise beim Einkaufen vorfindet (siehe Anhang D\_2).

Für jede richtige Antwort gab es einen Punkt, daher gab es bei sechs Reihen mit jeweils vier Bildern maximal 24 Punkte. Die benötigte Bearbeitungszeit bildete das Maß für die Inhibitionsfähigkeit. Im Durchschnitt bearbeiteten Kinder (von 4,5 bis 5,2 Jahren) diese Aufgabe in 64 bis 73 Sekunden. Mit zunehmendem Alter sinkt die Bearbeitungszeit (Jansen et al., 1999), weil die Kinder mit steigenden kognitiven Fähigkeiten kürzere Zeit zum Unterdrücken einer vorschnellen falschen Antwort bzw. zum Äußern der richtigen Antwort benötigten.

#### *7.1.5 Stichprobe*

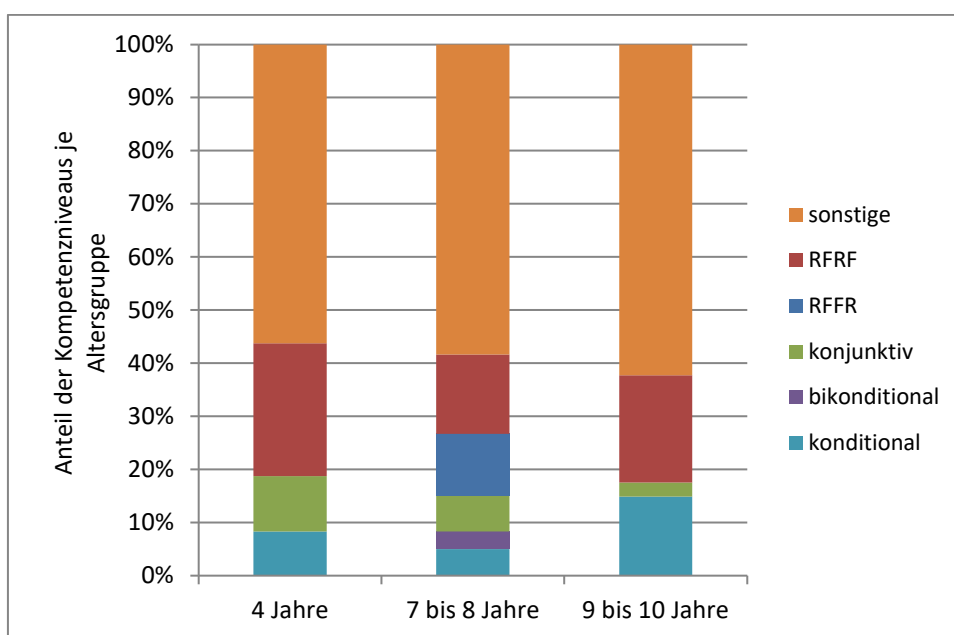
In der Vorstudie wurden 37 Kinder getestet. Die Stichprobe setzte sich aus neun Kindern je Gruppe mit einem durchschnittlichen Alter von etwa vier Jahren, zehn Kinder mit einem durchschnittlichen Alter von knapp sieben Jahren sowie neunzehn Kindern mit einem Altersdurchschnitt von knapp zehn Jahren zusammen.

Die Vorstudie wurde in den einzelnen Altersgruppen in unterschiedlicher Form durchgeführt. Die jüngste Altersgruppe der vierjährigen sowie die knapp siebenjährigen Kinder wurden in Einzelsitzungen befragt. Die Einzelinterviews eigneten sich für Kinder beider Altersgruppen, da Kinder meist mit vier bzw. sieben Jahren für die Testung noch zu wenige Lese- und Schreibkenntnisse haben. Die neun- bis zehnjährigen Kinder wurden aus Gründen der Zeitersparnis mit Hilfe eines Testheftes befragt. Dieses enthielt sowohl die Aufgaben des Wissenschaftsverständnis- als auch die Aufgaben des Schlussfolgerungstests.

#### *7.1.6 Ergebnisse*

Die erste Forschungsfrage untersuchte die Fähigkeiten der Kinder im Schlussfolgern in den drei Altersstufen. Im folgenden Absatz werden kurz zentrale Ergebnisse dargestellt.

Wie anhand der Abbildung 14 ersichtlich ist, nahmen mit dem Alter der Kinder die konditionalen Antworten zu, ebenso steigerten sich die Antworten mit bikonditionalem Interpretationsniveau. Gleichzeitig erhöhten sich die Antworten mit konjunktivem Interpretationsniveau von den vierjährigen hin zu den sieben- bis achtjährigen Kindern und verringerten sich dann wieder bei den neun- bis zehnjährigen Kindern. Die Interpretationsniveaus RFRF und RFFR verringerten sich von den Vierjährigen hin zu den Neun- bis Zehnjährigen. Die nicht bestimmbareren Antworten der Kategorie „sonstige“ nahmen mit steigendem Alter der Kinder leicht zu.



Anmerkung. MANOVA  $F(5, 30) = 16.54; p < 0.001; \epsilon^2 = 0.73$

Abbildung 14. Interpretationsniveaus zum Schlussfolgern.

Mit dem Kolmogorov-Smirnov-Test wurden die Daten auf Normalverteilung überprüft. Es zeigte sich, dass eine Normalverteilung der Daten vorlag. In einer multivariaten Analyse mit MANOVA ergab sich mit  $F(5, 30) = 16.54; p < 0.001; \epsilon^2 = 0.7$  ein signifikanter Altersunterschied zwischen den drei Altersgruppen. Bei der Analyse wurden die Kompetenzniveaus der Probanden jeweils als abhängige Variable gewählt.

Mit der zweiten Forschungsfrage wurde untersucht, welche Fähigkeiten, Kinder in den verschiedenen Altersstufen bei der Koordination von Theorie und Evidenz im Wissenschaftsverständnis zeigten.

Die Antworten der Probanden im Wissenschaftsverständnistest wurden mit Hilfe eines Kodierschemas ausgewertet: die Übereinstimmung von zwei unabhängigen Ratern wurde mit dem Kappa-Wert berechnet. Insgesamt ergab sich mit  $Kappa = 0.74$  eine gute Übereinstimmung über die zwölf Testitems (Beschreibung vom Wissenschaftsverständnistest siehe Kapitel 7.1.4) hinweg.

In Tabelle 12 sind die Mittelwerte mit den Standardabweichungen aus dem Wissenschaftsverständnistest für die einzelnen Altersstufen dargestellt. Bei den Aufgaben zur Produktion von Evidenz zeigte sich ein deutlicher Alterstrend, so konnten die Vierjährigen diese Aufgaben kaum lösen<sup>7</sup>. Die sieben- bis achtjährigen Kinder erreichten einen Mittelwert von .69 (0.95), die neun- bis zehnjährigen Kinder erzielten mit 2.4 (0.94) einen weitaus höheren Mittelwert als die nächst jüngeren Probanden.

Beim Erkennen von Evidenz konnte ebenso eine deutliche Kompetenzsteigerung über die einzelnen Altersstufen festgestellt werden. Die Mittelwerte reichten von 1.4 bei den Vierjährigen bis 2.3 bei den Neun- bis Zehnjährigen.

Bei der Produktion von Theorie ergaben sich Mittelwerte von 2.1 bis 2.8 über alle drei Altersstufen hinweg. Die Aufgaben zur Interpretation von Theorie waren am leichtesten zu lösen. Diese Aufgabenrubrik konnten die Jüngsten gut beantworten, sie hatten einen Mittelwert von 2.4. In diesem Aufgabenbereich war daher die Kompetenzsteigerung von den Vierjährigen bis hin zu den Neun- bis Zehnjährigen kaum sichtbar. Die Mittelwerte stiegen von 2.4 (0.87) auf 2.9 (0.22) an.

---

<sup>7</sup> Produktion von Evidenz bei Vierjährigen ist mit 0.1 so gering, dass der Wert im Diagramm nicht abgebildet wird.

Tabelle 12

*Mittelwerte (SD) im Wissenschaftsverständnistest*

Rubriken Wissenschaftsverständnis	Alter MW (SD)		
	4 Jahre	7-8 Jahre	9-10 Jahre
Produktion von Evidenz MW (SD)	0.1 (0)	0.7 (1.0)	2.4 (0.9)
Erkennen von Evidenz MW (SD)	1.4 (0.8)	1.2 (1.2)	2.3 (0.7)
Produktion von Theorie MW (SD)	2.1 (0.9)	1.9 (1.0)	2.8 (0.4)
Interpretation von Theorie	2.4 (0.8)	2.4 (0.9)	2.9 (0.2)

Bei der Prüfung der Mittelwerte auf Signifikanz ergab sich mit  $F(4,30)=302,06$ ;  $p < 0.001$  ein signifikanter Altersunterschied zwischen den Probanden im Wissenschaftsverständnistest. Insgesamt war die interne Konsistenz des Wissenschaftstests als eindimensionales Konstrukt mit Cronbachs alpha  $\alpha = 0.76$  gut.

Beide Tests, Schlussfolgerungstest und Wissenschaftsverständnistest funktionierten in der Vorstudie, da sie die Kompetenzen der Kinder altersspezifisch messen konnten. Daher wurden die Tests in der Hauptstudie in gleicher Weise durchgeführt, wobei die Formulierungen im Schlussfolgerungstest leicht abgeändert wurden.

## 7.2 Hauptstudie

Bei der erneuten Realisierung des Schlussfolgerungstests in der Hauptstudie wurden zwei Formulierungsänderungen bei den Vermutungen vorgenommen. Sie wurden den Kindern nicht mehr als Allaussage präsentiert, da diese in der Vorstudie hohe Erwartungen bei den Kindern in Bezug auf das Sprungverhalten von runden Gegenständen hervorgerufen hatten. Daher wurde die Formulierung der Vermutung „Alle runden Bälle springen“ aus der Vorstudie in „Dinge, die rund sind, springen“ bei der Hauptstudie geändert.

Bei der neuen Formulierung in der Hauptstudie wurde ebenso der Begriff „Ball“ nicht verwendet, da dieser für einige der verwendeten Gegenstände (den Luftballon mit Sand, das Flummie sowie die Tonkugel) nicht treffend ist (siehe Kapitel 7.1.4). Stattdessen wurde der Begriff „Dinge“ genutzt, da dies ein allgemeiner Begriff ist und Gegenstände mit sehr

unterschiedlichen Eigenschaften beschreibt. Zusätzlich wurden in der Hauptstudie zwei weitere, negative-Formulierungen (z.B. Dinge, die ohne Luft sind, springen nicht“) bei Kindern des Primarschulbereichs hinzugefügt, um deren erweiterte Kompetenzen sensibler erfassen zu können (siehe Kapitel 7.1.4).

Im folgenden Absatz werden die Forschungsfragen und Hypothesen der Hauptstudie vorgestellt.

### 7.2.1 Fragestellungen und Hypothesen

Frage 1: Welche Kompetenzen zeigen Kinder in den genannten Altersstufen beim Schlussfolgern?

Hypothese 1: Kinder im Kindergartenalter schlussfolgern auf dem Niveau „Sonstige“, während Grundschulkinder auf den höheren Niveaus, dem konjunktiven oder bikonditionalen Niveau antworten. Die Kompetenzniveaus werden bei der Analyse als abhängige Variablen genutzt.

In der zweiten Frage wurde das Wissenschaftsverständnis in den einzelnen Altersstufen geprüft.

Frage 2: Welche Kompetenzen zeigen Kinder beim Wissenschaftsverständnis?

Hypothese 2: Kinder im Kindergartenalter zeigen geringere Kompetenzen beim Wissenschaftsverständnis als Grundschulkinder.

Die dritte Frage zielte auf die empirische Überprüfung dieses Zusammenhangs beider Komponenten ab.

Frage 3: Zeigt sich ein Zusammenhang zwischen dem Wissenschaftsverständnis und der Schlussfolgerungskompetenz in kontextnahen Aufgaben im Bereich Naturwissenschaften?

Hypothese 3: Kinder, die im Wissenschaftsverständnistest eine hohe Kompetenz erreichen, zeigen im Prä-Test häufiger richtige Antworten im konditionalen Sinne als Kinder mit einer geringeren Kompetenz.

### 7.2.2 Methoden

In der Hauptstudie wurden zusätzlich zu den vier Vermutungen der Vorstudie (siehe Kapitel 7.1) zwei negativ formulierte Vermutungen hinzugefügt (siehe Tabelle 13) und den Probanden der drei älteren Altersgruppen (sechsjährige, sieben- bis achtjährige sowie neun- bis zehnjährige



Kinder) zur genaueren Erfassung ihrer Kompetenzen vorgelegt. In der Tabelle sind die Gegenstände nach folgender Systematik eingeordnet:  $[p \ q]$ ,  $[p \neg q]$ ,  $[\neg p \neg q]$ ,  $[\neg p \ q]$ . Bei den Erhebungen wurden die Gegenstände den Probanden nicht immer in derselben Reihenfolge präsentiert, um einen Reihenfolgeeffekt zu vermeiden (siehe Anhang A\_1).

Tabelle 13

*Weitere Vermutungen für sieben bis zehnjährige Probanden*

Vermutung	Gegenstände	bestätigend $[p \ q]$	widerlegend $[p \neg q]$	irrelevant $[\neg p \ q]$	irrelevant $[\neg p \neg q]$
Dinge, die ohne Luft sind, springen nicht.	Knetkugel	x			
	Würfel		x		
	Luftpolsterkissen			x	
	Tischtennisball				x
Dinge, die schwer sind, springen nicht.	Tonkugel	x			
	Eisenkugel		x		
	Federsäckchen			x	
	aufgeblasener Luftballon				x

Die Antworten der Kinder wurden im Hinblick auf ihre Angemessenheit hin bewertet (siehe Tabelle 13). Dabei wurden jeweils vier Gegenstände, mit denen eine Vermutung überprüft wird, betrachtet. Es wurden Interpretationsniveaus gebildet (siehe Kapitel 5.1). Das konjunktive Interpretationsniveau ergab sich, wenn ein Kind die Tonkugel als Gegenstand zur Bestätigung der Vermutung „Dinge, die schwer sind, springen nicht“ angesehen und die anderen drei Gegenstände (Eisenkugel, Federsäckchen, aufgeblasener Luftballon) als Beispiel zum Widerlegen der zuvor genannten Vermutung erkannt hatte (siehe Tabelle 13). Am Standort Münster werden die Daten für den Grundschulbereich von der wissenschaftlichen Mitarbeiterin Christin Robisch (Robisch, et al., 2014) erhoben und ausgewertet. Hierzu werden inferenzstatistische Verfahren eingesetzt, um Altersunterschiede der Kinder im Elementar- und Grundschulbereich zu ermitteln (Tröbst et al., 2012).

### 7.2.3 Stichprobe

Die Stichprobe der Querschnittstudie bestand aus 142 Kindern. 82 Kinder des Primarbereichs stammten aus dem Raum Münster (Nordrhein-Westfalen) und 60 Kinder des Elementarbereichs aus dem Rhein-Main-Gebiet (Hessen). An beiden Standorten wurden Kinder der umliegenden Grundschulen bzw. Kindergärten gewählt. Am Standort Münster wurden Schüler und Schülerinnen des Primarbereichs befragt, am Standort Frankfurt Kinder des Elementarbereichs.

Sowohl für die Bearbeitung der Schlussfolgerungsaufgaben als auch für die Aufgaben zur Koordination von Theorie und Evidenz sowie die der exekutiven Funktionen ist ein gutes Sprachvermögen (Verständnis und Produktion) grundlegend. Im Grundschulbereich beteiligten sich 49 Kinder mit einem durchschnittlichen Alter von 7,8 Jahren ( $SD = 4,2$ ) und 33 Kinder mit einem durchschnittlichen Alter von 10,0 Jahren ( $SD = 5,7$ ). Im Kindergartenbereich haben sich 29 Kinder mit einem Durchschnittsalter von 4,4 Jahren ( $SD = 3,5$ ) beteiligt und 31 Kinder im Durchschnittsalter von 5,7 Jahren ( $SD = 5,7$ ). Die Stichprobe mit den Angaben zur Anzahl der Probanden sowie zu deren Alter ist in Tabelle 14 aufgeführt.

Tabelle 14

#### *Stichprobe Querschnittstudie*

Alter der Probanden	Anzahl	Alter [Jahre] (SD)
4-jährig	29	4,4 (0,3)
5- bis 6-jährig	31	5,7 (0,6)
7- bis 8-jährig	49	7,8 (0,4)
9- bis 10-jährig	33	10,0 (0,5)

### 7.2.4 Design

Die Querschnittstudie umfasste drei Testtage, an denen jeweils zwei Tests durchgeführt wurden (siehe Tabelle 15). Insgesamt wurden sechs Tests mit allen Probanden an den Standorten Münster (Primarbereich) sowie Frankfurt (Elementarbereich) realisiert.

Tabelle 15

*Testdesign Querschnittstudie*

Tag 1	Tag 2	Tag 3
Farb-Benennung-Rückwärts-Test Wissenschaftsverständnis-test	Vorwissenstest Früchte-Gemüse-Stroop (Inhibitionstest)	Wortliste Schlussfolgerungstest im Kontext <i>Elastizität und Plastizität</i>

*7.2.5 Ergebnisse zur ersten Forschungsfrage: Fähigkeiten beim Schlussfolgern*

Mit dem Kolmogorov-Smirnov-Test wurden die Daten auf Normalverteilung überprüft. Es zeigte sich, dass eine Normalverteilung der Daten vorlag.

Um den ersten Teil der Forschungsfrage (Forschungsfrage 1a), welche Fähigkeiten Kinder in den einzelnen Altersstufen beim Evaluieren der Einzelereignisse zeigten, zu beantworten, sind die Antworten der Probanden in allen vier Altersstufen aufgeführt.

Dazu wurden zunächst zur Beantwortung die deskriptiven Ergebnisse der vier Ereignisse der Schlussfolgerungsaufgabe in Tabelle 16 dargestellt. Dabei wird ein Ereignis [p q]; [p¬q]; [¬p q]; [¬p¬q] pro Zeile mit den drei Antwortmöglichkeiten (bestätigend, irrelevant, widerlegend) in den vier Altersstufen gezeigt. In blau sind die Antworten gekennzeichnet, die im konditionalen Sinne angemessen bewertet wurden ([p q bestätigend]; [p¬q widerlegend]; [¬p q irrelevant]; [¬p¬q irrelevant]).

Kinder aller vier Altersgruppen konnten die bestätigenden Ereignisse vorwiegend als solche erkennen; eine deutliche Zunahme der korrekten Antworten zeigte sich bei den sieben- bis achtjährigen sowie bei den neun- bis zehnjährigen Kindern (siehe Tabelle 16). Dieser Trend stellte sich ebenso heraus bei den Ereignissen [p¬q wid], die als widerlegend einzuschätzen waren. Bei Ereignissen der Form [¬p q irrel], die als irrelevant zu bewerten waren, zeichnete sich ein monotoner Anstieg von Beurteilungen ab. Die vierjährigen Kinder bewerteten am seltensten die Ereignisse der Form [¬p q irrel] als irrelevant, wohingegen die Viertklässler diese Antwortoption am häufigsten wählten. Die gleiche Tendenz ergab sich für Ereignisse der Form [¬p¬q irrel], allerdings war die Beurteilung als irrelevant hier nicht die beliebteste Antwortkategorie. Insgesamt zeigte sich deskriptiv ein deutlicher Anstieg von Antworten im konditionalen Sinn [p q best]; [p¬q wid]; [¬p q irrel]; [¬p¬q irrel].

Tabelle 16

*Relative Verteilung der Evaluationen der präsentierten Ereignisse*

	Elementarbereich						Primarbereich					
	4 Jahre			5- 6 Jahre			7 - 8 Jahre			9-10 Jahre		
Ereignis	best	irr	wid	best	irr	wid	best	irr	wid	best	irr	wid
[p q]	63%	20%	16%	70%	11%	19%	86%	6%	8%	80%	14%	6%
[p¬q]	37%	15%	48%	12%	20%	67%	7%	14%	79%	4%	16%	80%
[¬p q]	38%	22%	40%	27%	32%	41%	17%	33%	50%	4%	52%	44%
[¬p¬q]	38%	20%	42%	27%	25%	48%	43%	29%	28%	43%	42%	15%

*Anmerkung.* Nicht-Addition zu 100 (pro Zeilenzelle) aufgrund von Rundung. best = bestätigend; irr = irrelevant; wid = widerlegend.

Bei der Berechnung mit einer ANOVA wurden die einzelnen Ereignisse (p q best]; [p¬q wid]; [¬p q irrelevant]; [¬p¬q irrel]) jeweils als abhängige Variable erfasst. Dabei ergab sich ein signifikanter Unterschied beim Ereignis [p¬q wid] zwischen den vierjährigen und den fünf- bis sechsjährigen Kindern mit  $F(1, 60) = 5.6$ ;  $p < .05$ ;  $\epsilon^2 = 0.04$ ; die Effektstärke (=  $\epsilon^2$ ) war sehr gering. Bei allen anderen drei Ereignissen ([p q best]; [¬p q irrelevant]; [¬p¬q irrel]) war kein signifikanter Unterschied zwischen der Altersgruppe der vierjährigen und der fünf- bis sechsjährigen Probanden festzustellen.

In Tabelle 17 sind alle Tests der Querschnittstudie für die vier Altersstufen (vier Jahre, fünf bis sechs Jahre, sieben bis acht Jahre, neun bis zehn Jahre) aufgeführt. Der Schlussfolgerungstest ist mit den vier Einzelereignissen im bestmöglichen, dem konditionalen Antwortniveau aufgeführt. Die Analysen wurden im Rahmen eines Berichtes für die Deutsche Forschungsgemeinschaft durchgeführt (Tröbst et al. 2012a). Auf deskriptiver Ebene zeigte sich ein deutlicher Alterstrend von den vierjährigen bis hin zu den neun- bis zehnjährigen Teilnehmern in allen Bereichen: die Kompetenzen beim Schlussfolgern erhöhten sich ebenso wie bei der Zentralen Exekutiven und bei der auditiven Merkfähigkeit (Phonologische Schleife), da die Mittelwerte anstiegen. Die Inhibitionsfähigkeit verbesserte sich auch, denn die Bearbeitungszeit der Aufgaben nahm mit dem Alter ab (siehe Tabelle 17). In einer ANOVA zeigte sich ein signifikanter Unterschied beim Inhibitionstest zwischen der Gruppe der vierjährigen und der fünf-bis sechsjährigen Probanden mit

$F(1, 61) = 8.6; p=.05; \epsilon^2=0.13$ . Das bedeutete, je älter die Probanden wurden, desto weniger Zeit benötigten sie zum Bearbeiten der Aufgaben.

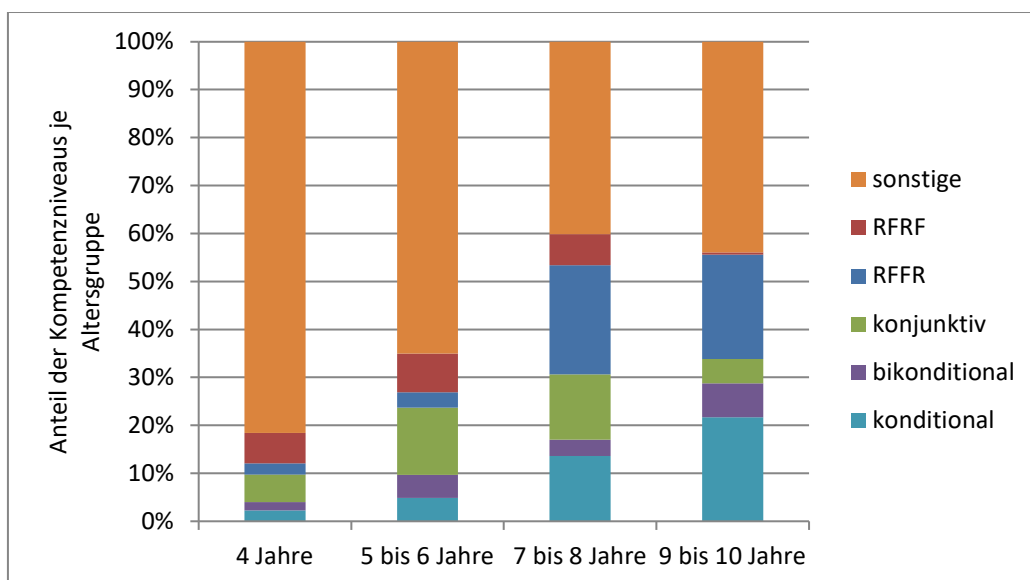
Tabelle 17

*Mittelwerte (Standardabweichung) der Tests im Überblick*

Variable	Alter der Probanden							
	4 Jahre		5-6 Jahre		7 -8 Jahre		9 -10 Jahre	
	M	SD	M	SD	M	SD	M	SD
Alter (Jahre)	4.45	0.30	5.71	0.55	7.80	0.45	10.05	0.47
Korrekte Antworten (relative Häufigkeiten)								
Variable	M	SD	M	SD	M	SD	M	SD
[p q] (best)	0.63	0.37	0.70	0.28	0.86	0.20	0.80	0.24
[p¬q] (wid)	0.48	0.32	0.67	0.31	0.79	0.28	0.80	0.27
[¬p q] (irrel)	0.22	0.23	0.32	0.26	0.33	0.36	0.52	0.34
[¬p¬q] (irrel)	0.20	0.24	0.25	0.24	0.29	0.30	0.42	0.31
Auditive Merkfähigkeit (Phonological Loop)								
Variable	M	SD	M	SD	M	SD	M	SD
	3.46	0.74	3.65	0.71	4.18	0.93	4.45	0.71
Zentrale Exekutive (Spanne)								
Variable	M	SD	M	SD	M	SD	M	SD
	2.17	0.54	2.68	0.83	3.18	1.30	3.83	0.81
Inhibition (Sekunden)								
Variable	M	SD	M	SD	M	SD	M	SD
	117	39.74	83.23	29.81	55.94	14.96	45.33	9.17
Vorwissen (Summenscore)								
Variable	M	SD	M	SD	M	SD	M	SD
	1.80	0.88	2.02	0.90	2.56	0.81	2.69	0.63
Wissenschaftsverständnis (wle)								
Variable	M	SD	M	SD	M	SD	M	SD
	0.36	1.32	0.69	1.10	1.42	1.18	2.24	1.32

Eine detaillierte Vorstellung über die Fähigkeiten von Kindern in den verschiedenen Altersstufen kann man gewinnen, wenn man die einzelnen Interpretationsniveaus zum Schlussfolgern betrachtet. Eine genaue Beschreibung der Interpretationsniveaus sowie deren Bildung kann in Kapitel 5.1 nachgelesen werden.

Die vier Ereignisse über alle vier Vermutungen [ $p \rightarrow q$  best]; [ $p \rightarrow \neg q$  wid], [ $\neg p \rightarrow q$  irrel]; [ $\neg p \rightarrow \neg q$  irrel] bildeten die Grundlage für die Zusammenstellung der Interpretationsniveaus. Zur Beantwortung des zweiten Teil der ersten Forschungsfrage (Forschungsfrage 1b) wurden die relativen Häufigkeiten zum Schlussfolgern entsprechend der Interpretationsniveaus gebildet. Sie sind im Balkendiagramm (Abbildung 15) zu sehen. Deutlich zu erkennen ist, dass mit zunehmendem Alter die konditionalen Antworten anstiegen. Ebenso wuchsen Antworten auf bikonditionalem und konjunktivem Interpretationsniveau. Gleichzeitig reduzierten sich die nicht bestimmbareren Antworten der Kategorie „Sonstige“ von den vierjährigen hin zu den neun- bis zehnjährigen Kindern. Zudem zeigte sich, dass das RFFR-Niveau mit dem Alter der Kinder zunahm, während das RFRF-Niveau abnahm (vgl. Tröbst et al., 2011; Tröbst, Robisch, Stephan-Gramberg, Hardy & Möller, 2012 b). Anschließend wurde eine *MANOVA* berechnet, wobei die verschiedenen Interpretationsniveaus (konjunktiv, bikonditional, konditional, RFFR, RFRF) als abhängige Variablen genutzt wurden. Dabei stellte sich ein signifikanter Unterschied bei den Kompetenzen zwischen den Altersgruppen heraus.



Anmerkung. *MANOVA*  $F(5,132) = 4.76$   $p < 0.001$ ;  $\epsilon^2 = 0.15$

Abbildung 15. Relative Häufigkeiten Interpretationsniveaus.

Die relativen Häufigkeiten der Interpretationsniveaus sind in tabellarischer Übersicht (Tabelle 18) aufgeführt.

Tabelle 18

*Relative Häufigkeiten der Interpretationsniveaus*

Interpretations-niveaus	Alter der Kinder			
	4 Jahre	5-6 Jahre	7-8 Jahre	9-10 Jahre
konditional	2%	5%	14%	22%
bikonditional	2%	5%	3%	7%
konjunktiv	6%	14%	14%	5%
RFFR	2%	3%	23%	22%
RFRF	6%	8%	6%	1%
Sonstige	82%	65%	40%	44%

*Anmerkung.* Nicht-Addition zu 100 (pro Spalte) aufgrund von Rundung.

Bei Betrachtung der relativen Häufigkeiten (siehe Abbildung 15, Tabelle 18) sowie der Mittelwerte (Tabelle 19) zeigte sich ein deutlicher Alterstrend. Die fortgeschrittenen Niveaus, vom konjunktiven hin zum konditionalen Niveau, nahmen mit dem Alter deutlich zu; gleichzeitig verringerten sich die Antworten der Kategorie „Sonstige“. Bei den vierjährigen Probanden kam diese Kategorie mit 82% sehr häufig vor; wohingegen sie sich bei den neun- bis zehnjährigen Probanden auf 44% reduzierte.

Vergleicht man ausschließlich die Gruppen der vierjährigen und der fünf- bis sechsjährigen Probanden, so verringerte sich die Kategorie „Sonstige“ von 82% auf 65%; in einer ANOVA ergab sich aber kein signifikanter Unterschied zwischen beiden Altersgruppen.

Tabelle 19

*Mittelwerte (Standardabweichungen) der Interpretationsniveaus*

Interpretations-niveaus	Alter der Kinder			
	4 Jahre	5-6 Jahre	7-8 Jahre	9-10 Jahre
konditional	0.10 (0.35)	0.29 (0.69)	0.81 (1.5)	1.30 (1.5)
bikonditional	0.10 (0.40)	0.29 (0.58)	0.20 (0.4)	0.42 (0.7)
konjunktiv	0.48 (1.02)	0.84 (1.04)	0.82 (1.05)	0.30 (0.53)
RFFR	0.28 (0.84)	0.19 (0.47)	1.37 (1.63)	1.85 (1.3)
RFRF	0.38 (0.73)	0.48 (1.06)	0.39 (0.76)	0.30 (1.7)
sonstige	4.62 (1.68)	3.9 (1.8)	2.4 (1.85)	2.64 (2.06)

*Anmerkung. MANOVA  $F(5, 132) = 4.76; p < 0.001; \varepsilon^2 = 0.15$ .*

Der Kolmogorov-Smirnov-Test ergab, dass eine Normalverteilung der Daten vorlag. In einer multivariaten Varianzanalyse *MANOVA* stellte sich der Alterseffekt mit  $F(5,132) = 4.76; p = 0.00; \varepsilon^2 = 0.15$  als signifikant heraus. Dabei wurde bei der Einstellung der Kontraste „wiederholt“ gewählt, damit die vier Altersgruppen der Stichprobe jeweils miteinander verglichen werden konnten.

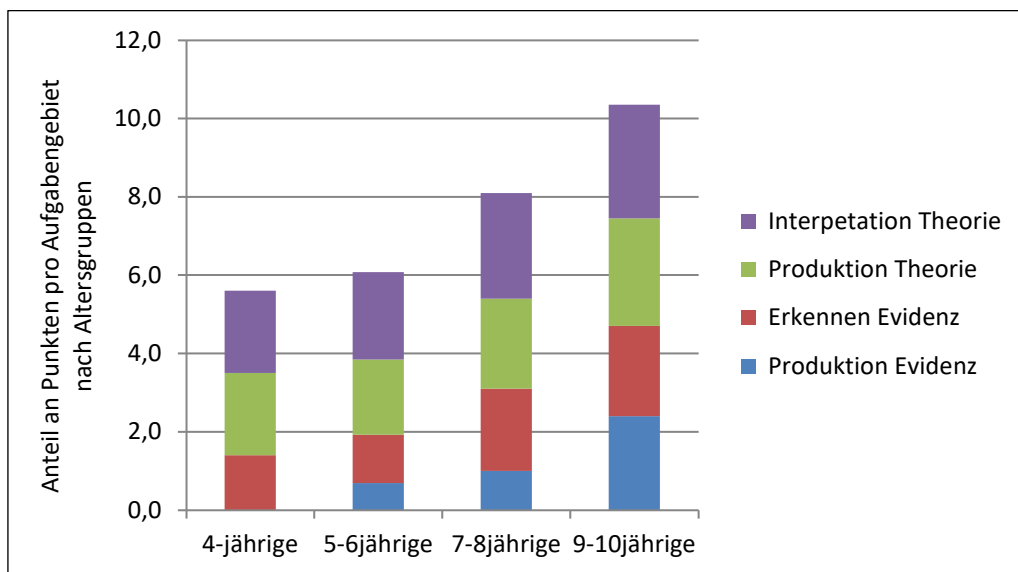
### 7.2.6 Ergebnisse zur zweiten Forschungsfrage: Kompetenzen beim Wissenschaftsverständnis

Die folgenden Analysen beantworten die zweite Forschungsfrage, in der untersucht wurde, welche Fähigkeiten Kinder unterschiedlichen Alters beim Wissenschaftsverständnis aufgewiesen hatten. Die Analysen wurden sowohl auf deskriptiver als auch auf inferenzstatistischer Ebene durchgeführt.

Die relativen Häufigkeiten der mittleren Punktzahlen der Antworten bei den zwölf Items zum Wissenschaftsverständnis sind in Abbildung 16 sowie die Mittelwerte in Tabelle 20 zusammengestellt. Es zeigten sich ähnliche Ergebnisse wie in der Vorstudie. Die Kompetenzsteigerung mit zunehmendem Alter wurde deutlich, da Kinder ihre Fähigkeit zur Theorie-Evidenz-



Koordination mit zunehmendem Alter steigerten. Das bedeutet, je älter die Kinder waren, desto mehr Punkte erreichten sie im Mittel bei den einzelnen Aufgaben. Bei den Aufgaben zur Produktion von Evidenz erzielten Kinder aller Altersgruppen am wenigsten Punkte, bei den Aufgaben zur Interpretation von Theorie am meisten Punkte. Die vierjährigen Probanden konnten bei den Items zur Produktion von Evidenz noch keine Punkte erreichen.



MANOVA  $F(4,134) = 1041.28; p < 0.05; \epsilon^2 = 0.98$

Abbildung 16. Mittlere Punktzahlen beim Wissenschaftsverständnistest.

Tabelle 20

Mittelwerte (Standardabweichungen) im Wissenschaftsverständnistest

Rubriken Wissenschaftsverständnis	Alter der Kinder			
	4 Jahre	5–6 Jahre	7–8 Jahre	9–10 Jahre
Produktion von Evidenz	1.1(1.5)	2.3 (2.0)	3.7 (1.9)	5.0 (1.9)
Erkennen von Evidenz	2.4 (2.0)	3.7 (1.8)	4.7(1.5)	5.3(0.8)
Produktion von Theorie	3.3 (2.1)	4.6 (1.5)	5.3 (0.9)	5.8 (0.6)
Interpretation von Theorie	4.3 (1.5)	5.4 (1.1)	5.5 (0.8)	5.8 (0.5)

Anmerkung. MANOVA  $F(4,134) = 1041.28; p < 0.05; \epsilon^2 = 0.98$ .

Wie aus Tabelle 20 zu entnehmen ist, zeigte sich ein großer Zuwachs an richtigen Antworten bei der *Produktion von Evidenz* sowie beim *Erkennen von Evidenz* beim Vergleich der vierjährigen und der neun- bis zehnjährigen Probanden. Die vierjährigen Kinder erreichten bei der Produktion von Evidenz einen Mittelwert von 1.1 (SD = 1.5), während die neun- bis zehnjährigen Probanden im Mittel 5.0 (SD = 1.9) erlangten. Ähnlich groß war der Unterschied im Mittel bei der Aufgabe zum Erkennen von Evidenz zwischen beiden Altersgruppen. Hier lag der Mittelwert bei 2.4 (SD = 2.0) bei den vierjährigen und 5.3 (SD = 0.8) bei den neun-bis zehnjährigen Kindern. Ähnlich sah es beim Vergleich der vierjährigen mit den neun- bis zehnjährigen Probanden bei der Produktion und Interpretation von Theorie aus.

Betrachtet man die vier Aufgabenrubriken (*Produktion von Evidenz*, *Produktion von Theorie*, *Produktion von Theorie*, *Interpretation von Theorie*) innerhalb einer Altersstufe (Tabelle 20), so ließ sich bei den Vierjährigen ein deutlicher Anstieg der Mittelwerte von der Produktion von Evidenz hin zur Interpretation von Theorie feststellen. Daran ließ sich der Schwierigkeitsgrad der Aufgaben ersehen. Während die Produktion von Evidenz besonders jungen Kindern schwer fiel, kamen sie sehr gut mit den Aufgaben zur Produktion von Theorie zurecht. Die vierjährigen Probanden hatten im Bereich Produktion von Evidenz mittlere Punktzahlen von 1.1 (SD = 1.5) und bei der Interpretation von Theorie mittlere Punktzahlen von 4.3 (SD = 1.5).

Die Neun- bis Zehnjährigen hatten bei der Produktion von Evidenz mittlere Punktzahlen von 5.0 (SD = 1.9) und bei der Interpretation von Theorie mittlere Punktzahlen von 5.8 (SD = 0.5).

Vergleicht man die vierjährigen mit neun- bis zehnjährigen Teilnehmern bei den Aufgaben zur Interpretation von Theorie, so erkennt man eine geringere Unterscheidung zwischen den Altersgruppen als beim Vergleich mit den Aufgaben zur Produktion von Evidenz und Produktion von Theorie. Bei der Produktion von Theorie erreichten die jüngsten Teilnehmer bereits mittlere Punktzahlen von 4.3 (SD = 1.5) und die neun- bis zehnjährigen erzielten einen Mittelwert von 5.8 (SD = 0.5).

Insgesamt ergab sich in einer multivariaten Varianzanalyse ein signifikanter Alterstrend mit  $F(4, 134) = 1041.28$ ;  $p < 0.05$ ;  $\epsilon^2 = 0.98$ . Vergleicht man ausschließlich die Gruppe der vierjährigen mit denen der fünf- bis sechsjährigen Probanden in einer *MANOVA*, so stellte sich mit  $F(1,5 6) = 211.5$   $p < 0.01$   $\epsilon^2 = 0.94$  ebenso ein signifikanter Unterschied zwischen den beiden Altersgruppen heraus.

*Item-Response-Analyse für den Wissenschaftsverständnistest*

Im Rahmen der zweiten Forschungsfrage, die die Kenntnisse der Kinder beim Wissenschaftsverständnistest in den vier Altersstufen untersuchte, wurde eine eindimensionale Rasch-Skalierung durchgeführt (siehe Tabelle 21). Die Item-Kennwerte wurden für jeden Aufgabentyp in folgender Weise angegeben: Lösungshäufigkeit, Schwierigkeit, Trennschärfe sowie Infit und T-Wert. Die Rasch-Skalierung wurde in gleicher Weise wie im Bericht für die Deutsche Forschungsgemeinschaft (Tröbst, Robisch, Hardy, Möller & Stephan-Gramberg, 2012a) durchgeführt. Die Items sind in absteigender Reihenfolge nach ihrer Schwierigkeit abgebildet; das bedeutet, dass das schwierigste Item als erstes und das leichteste als letztes aufgeführt ist. Das methodische Vorgehen zur Item-Response-Theorie wird in Kapitel 7.1.4. beschrieben.

Tabelle 21  
*Item-Kennwerte Wissenschaftsverständnistest*

Items	Aufgabentyp	N	Lösungshäufigkeit (%)	Schwierigkeit	Trennschärfe	Infit MNSQ	T-Wert
1_1 Produktion Evidenz Marius	Produktion von Evidenz	142	44.37	2.93	0.63	1.09	0.8
3_1 Produktion Evidenz Struppi	Produktion von Evidenz	142	62.94	1.58	0.77	0.84	-1.1
2_1 Produktion Evidenz Elena	Produktion von Evidenz	142	69.23	1.09	0.70	0.94	-0.5
3_2 Erkennen Evidenz Struppi	Erkennen von Evidenz	142	69.23	1.09	0.72	0.9	-0.8
2_2 Erkennen Evidenz Elena	Erkennen von Evidenz	142	78.17	0.33	0.68	0.99	1.2
2_3 Produktion Theorie Elena	Produktion von Theorie	142	82.39	-0.09	0.58	1.09	0
1_2 Erkennen Evidenz Marius	Erkennen von Evidenz	142	85.52	-0.012	0.51	1.18	1.2
2_4 Interpretation Theorie Elena	Interpretation von Theorie	142	82.52	-0.12	0.59	1.08	0.6

Items	Aufgabentyp	N	Lösungshäufigkeit (%)	Schwierigkeit	Trennschärfe	Infit MNSQ	T-Werte
3_3 Produktion Theorie Struppi	Produktion von Theorie	142	86.71	-0.61	0.53	1.07	0.5
1_3 Produktion Theorie Marius	Produktion von Theorie	142	88.10	-0.80	0.63	0.91	-0.5
1_4 Interpretation Theorie Marius	Interpretation von Theorie	142	94.40	-1.96	0.43	1.02	0.2
3_4 Interpretation Theorie Struppi	Interpretation von Theorie	142	97.90	-3.33	0.32	1.08	0.3

*Anmerkung.* Skala EAP/PV - Reliabilität: 0.79;  
 Varianz: 2.7;  
 Schwierigkeit: M = -0.01, SD = 1.63, MIN = -3.3, MAX = 2.93;  
 Trennschärfe: M = 0.59, SD = 0.13, MIN = 0.32, MAX = 0.77;  
 Infit: M = 1.02, SD = 0.10, MIN = 0.84, MAX = 1.18.

In der Wright Map (Abbildung 17) sind die Items des Wissenschaftsverständnistests in Logit-Einheiten auf einer Skala abgebildet. Ganz oben ist das schwierigste Item (1\_1 Produktion von Evidenz), ganz unten ist das leichteste Item (3\_4 Interpretation von Theorie) aufgeführt.

Die Fit-Statistiken waren gut. Der Infit lag mit 1.01 (SD = 0.09) innerhalb der Grenzen von 0.8 und 1.2 (Adams, 2002; Wu & Adams, 2007).

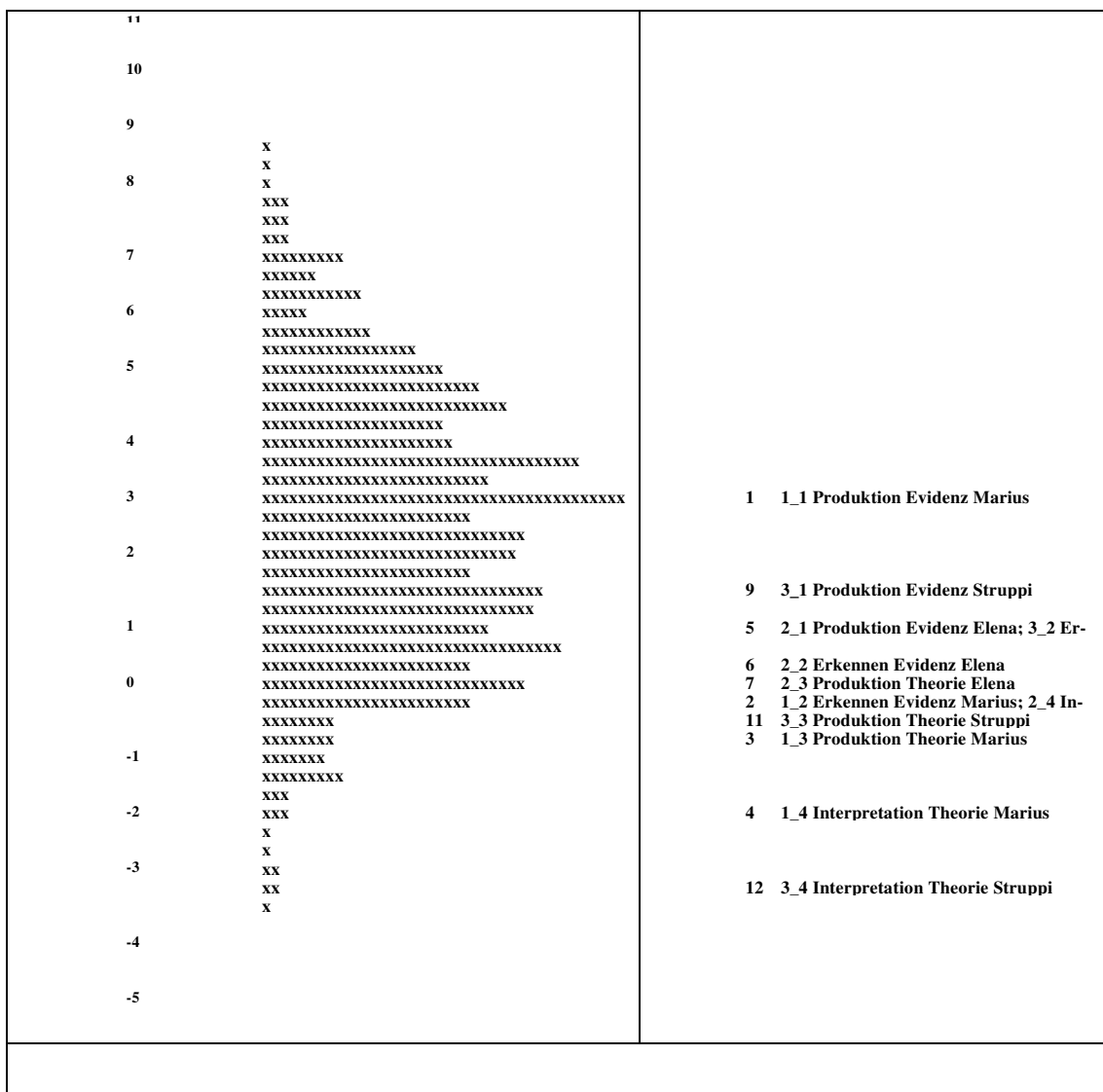


Abbildung 17. Wright Map des eindimensionalen Rasch-Modells.

Die punktbiserialen Korrelationen bzw. klassischen Trennschärfen befanden sich bei 0.43 und 0.77 mit einer mittleren Trennschärfe von 0.59 (SD= 0.13) in einem zufriedenstellenden Bereich (Moosbrugger & Kevala, 2008). Die Varianz der Variablen wurde mit 2.8 ermittelt. Insgesamt war die EAP/VP-Reliabilität, die ähnlich der Reliabilität von Cronbachs Alpha ist, mit einem Wert von 0.79 zufriedenstellend (siehe Tabelle 21). Die akzeptablen Reliabilitätswerte zeigten, dass eine gute interne Konsistenz für den Wissenschaftstest vorlag. Es gab keine signifikanten T-Werte ( $-2.0 < T\text{-Werte} < + 2.0$ ); das bedeutet die Abweichung des beobachteten Wertes vom erwarteten Wert war nicht signifikant. Die *Weighted Likelihood Estimation* (WLE) nahm mit dem Alter der Kinder zu (siehe Tabelle 22).

Tabelle 22

*Weighted Likelihood Estimation*

	4 Jahre	5–6 Jahre	7–8 Jahre	9–10 Jahre
Weighted Likelihood Estimation (WLE)	-0.36 (1.3)	0.68 (1.2)	1.42 (1.2)	2.24 (1.3)

### 7.2.7 Ergebnisse zur dritten Forschungsfrage: Zusammenhang von Schlussfolgern und Wissenschaftsverständnis

Um der dritten Forschungsfrage nachzugehen, ob es einen Zusammenhang zwischen einer Schlussfolgerungskompetenz und dem Wissenschaftsverständnis gab, wurden logistische Regressionen mit dem Programm R (Field, 2009) errechnet (Tröbst et al., 2012a). In der Arbeitsgruppe von Frau Prof. Dr. Möller wurden Regressionsanalysen durchgeführt, um den Zusammenhang zwischen Schlussfolgern und Wissenschaftsverständnis zu prüfen (DFG-Bericht, 2012). Dabei waren die angemessenen Evaluationen der einzelnen Ereignisse  $[p \neg q \text{ wid}]$ ;  $[\neg p q \text{ irrel}]$ ;  $[\neg p \neg q \text{ irrel}]$  im konditionalen Sinne die abhängigen Variablen und wurden mit der mittleren Personenfähigkeit des Wissenschaftsverständnisses sowie dem Alter in einer schrittweisen Erweiterung berechnet. Als feste Effekte gingen die Haupteffekte (Ereignis, Wissenschaftsverständnis und Alter) ein (Tröbst et al., 2012a). Das Ereignis  $[p q \text{ best}]$  wurde dabei als Referenzkategorie genutzt. Es ergab sich mit  $B = 0.44$  und  $SE = 0.18$ ;  $p < 0.05$  ein signifikanter Zusammenhang zwischen Schlussfolgerungskompetenz und Wissenschaftsverständnis (Tröbst et al., 2012a). Bei Verknüpfung der einzelnen Ereignisse jeweils mit dem Wissenschaftsverständnis zeigte sich, dass sich bei den Ereignissen  $[\neg p q \text{ irrel}]$  sowie bei  $[\neg p \neg q \text{ irrel}]$  der positive Zusammenhang des Wissenschaftsverständnisses reduzierte; beim Ereignis  $[\neg p q \text{ irrel}]$  lag er bei  $B = -0.04$ ;  $SE = 0.08$ ,  $p < 0.05$ , beim Ereignis  $[\neg p \neg q \text{ irrel}]$  lag er bei  $B = -0.14$ ,  $SE = 0.08$   $p < 0.05$ . Im Vergleich dazu gab es einen positiven Wert beim Ereignis  $[p \neg q \text{ wid}]$  mit  $B = 0.09$ ,  $SE = 0.08$   $p < 0.05$ .

### 7.3 Diskussion

In der folgenden Diskussion werden die Ergebnisse der Querschnittstudie im Hinblick auf die drei Fragestellungen zusammengefasst und interpretiert.

Die erste Fragestellung zielte auf die Schlussfolgerungskompetenzen der Kinder in den vier Altersgruppen (vier Jahre, fünf bis sechs Jahre, sieben bis acht Jahre, neun bis zehn Jahre) ab. Zur Erfassung der Kompetenzen beim Schlussfolgern wurden den älteren Probanden sechs Vermutungen, den vierjährigen Kindern vier Vermutungen präsentiert. Die Reihenfolge der präsentierten Ereignisse [p q best]; [p¬q wid]; [¬p q irrel]; [¬p¬q irrel] variierte zwischen den einzelnen Vermutungen, sodass die Kinder beim Antworten keine Reihenfolge entdecken konnten.

Für die Ermittlung der Schlussfolgerungskompetenzen wurden die Einzelereignisse über vier Vermutungen jeweils zusammengefasst. Die deskriptive Analyse zeigte einen signifikanten Anstieg der Kompetenzen im Schlussfolgern mit dem Alter (siehe Tabelle 16). Diese Zunahme bestätigte sich in einer multivariaten Analyse mit *MANOVA*. In einer weiteren Analyse mit *ANOVA* wurden die Unterschiede der Kinder unterschiedlichen Alters bei den Einzelereignissen untersucht. Dabei zeigte sich ein signifikanter Unterschied zwischen den vierjährigen und den fünf- bis sechsjährigen Kindern beim Ereignis [p¬q wid]; die fünf- bis sechsjährigen Probanden konnten signifikant besser die widerlegenden Ereignisse im Schlussfolgerungstest erkennen. Die Kompetenzsteigerung mit dem Alter wurde ebenso in einer logistischen Regressionsanalyse deutlich (Tröbst et al., 2012a). So zeigten die Regressionsgewichte einen positiven Effekt des Alters auf das korrekte Evaluieren von Evidenzen. Besonders klar war der Anstieg korrekter Evaluationen bei den irrelevanten Ereignissen, hier zeigen die neun- bis zehnjährigen Kinder signifikant bessere Ergebnisse als die anderen drei Altersgruppen. Den jüngeren Kindern (vier- bis achtjährige Kinder) fiel die Evaluation von irrelevanten Ereignissen im Vergleich zu den bestätigenden oder widerlegenden Ereignissen schwer. Insgesamt nahmen die korrekten Einschätzungen mit dem Alter signifikant zu, das heißt mit dem Alter konnten sehr viel mehr Kinder richtig antworten. Insgesamt kann die Kompetenzzunahme beim Schlussfolgern mit dem Alter als lineare Zunahme beschrieben werden.

Entsprechend der Evaluationen der Einzelereignisse antworteten die Kinder mit zunehmendem Alter auf einem fortgeschrittenen Interpretationsniveau (siehe Tabelle 18, Abbildung 15). Die drei Interpretationsniveaus nach Barrouillet und Kollegen (2008), sowie das

RFRF<sup>8</sup>- und das RFFR<sup>9</sup>-Niveau (Tröbst et al., 2011) stellten sich in den Altersgruppen folgendermaßen dar: Das konjunktive Interpretationsniveau nahm von den vierjährigen zu den sieben- bis achtjährigen Kindern zu, bei den neun- bis zehnjährigen Kindern nahm es wieder ab. Im Vergleich dazu steigerte sich sowohl das bi-konditionale als auch das konditionale Interpretationsniveau mit dem Alter. Die Häufigkeit des Auftretens des RFRF- und des RFFR-Niveaus (Tröbst et al., 2011) veränderte sich ebenso mit dem Alter. Während das RFRF-Niveau mit zunehmendem Alter deutlich abnahm, war die Häufigkeit der Antworten auf dem RFFR-Niveau bei sieben- bis achtjährigen Teilnehmern höher als bei den Vierjährigen. Bei neun- bis zehnjährigen Probanden sanken die Antworten auf dem RFFR-Niveau wieder. In einer multivariaten Varianzanalyse zeigte sich im Hinblick auf die Schlussfolgerungskompetenzen ein signifikanter Alterseffekt. Insgesamt konnte die erste Hypothese bestätigt werden, da die inferenzstatistischen Analysen sowohl bei den Einzelereignissen als auch bei den Interpretationsniveaus signifikante Zuwächse zwischen den Altersstufen aufgewiesen hatten. Die vorliegenden empirischen Befunde bestätigten die Ergebnisse der vorangegangenen Studien (Barrouillet et al., 2008; Tröbst et al., 2011).

Die zweite Forschungsfrage untersuchte die Kompetenzen der Kinder in den vier Altersstufen (vierjährig, fünf- bis sechsjährig, sieben- bis achtjährig sowie neun- bis zehnjährig) beim Wissenschaftsverständnis. Er prüfte die Kompetenzen der Probanden und umfasst insgesamt zwölf Items mit vier unterschiedlichen Aufgabenbereichen (Produktion von Evidenz, Erkennen von Evidenz, Produktion von Theorie, Interpretation von Theorie). Ähnlich der zuvor beschriebenen Schlussfolgerungskompetenzen (siehe Frage 1) stiegen die Fähigkeiten beim Wissenschaftsverständnis mit dem Alter der Probanden. Die vier Aufgabenrubriken (Produktion von Evidenz, Erkennen von Evidenz, Produktion von Theorie, Interpretation von Theorie) unterschieden sich im Hinblick auf ihre Schwierigkeit. Am schwierigsten waren jeweils die Aufgaben zum Produzieren von Evidenz. Diese Aufgaben lösen die Vierjährigen in allen drei Kontexten nur selten (Item 1\_1 Produktion von Evidenz Marius; Item 2\_1 Produktion von Evidenz Elena; Item 3\_1 Produktion von Evidenz Struppi). Sie konnten selten plausible Evidenzen nennen (siehe Kapitel 7.2.6). In einer *MANOVA*, in der die vier Aufgabenbereiche als abhängige Variable fungierten, stellte sich ein signifikanter Unterschied zwischen beiden Altersgruppen

---

<sup>8</sup> RFFR-Niveau=Evaluation von vier Ereignissen im Zusammenhang mit einer Hypothese: Richtig-Falsch-Falsch-Richtig.

<sup>9</sup> RFRF-Niveau= Evaluation von vier Ereignissen im Zusammenhang mit einer Hypothese: Richtig-Falsch-Richtig-Falsch.



heraus. Das heißt die fünf- bis sechsjährigen Probanden konnten signifikant besser die Aufgaben des Wissenschaftsverständnistest lösen als die vierjährigen Probanden.

Betrachtet man die Ergebnisse der eindimensionalen Raschanalyse (siehe Kapitel 7.2.6), so ergab sich in Bezug auf die Lösungswahrscheinlichkeit für die vier Aufgabenbereiche folgendes Bild. Die Produktion von Evidenz stellte sich im Vergleich zu den anderen drei Aufgabenbereichen (Erkennen von Evidenz, Produktion von Theorie, Interpretation von Theorie) als anspruchsvollster Aufgabenbereich heraus. Die Lösungswahrscheinlichkeit bei der Produktion von Evidenz im ersten Kontext (Item 1\_1 Produktion von Evidenz) war am niedrigsten (siehe Kapitel 7.2.6), bei den Aufgaben zur Produktion von Evidenz aus den beiden anderen Kontexten ist sie deutlich höher. Die Schwierigkeit innerhalb der drei Kontexte differierte auch bei den anderen Aufgabenbereichen. Die Steigerung der Lösungswahrscheinlichkeiten vom ersten Kontext (Marius auf dem Baum) über den zweiten (Elena und Sophie beim Mittagessen) hin zum dritten Kontext (Struppi und Marie auf ihrem Spaziergang) hing vermutlich damit zusammen, dass die Teilnehmer sich beim ersten Kontext mit der Aufgabenstellung erst vertraut machen mussten, während sie diese beim dritten Kontext bereits kannten (siehe Kapitel 7.2.6). Der Wissenschaftsverständnistest ist im Anhang zu finden. Der Protokollbogen ist im Anhang\_B1, der Interviewleitfaden im Anhang\_B2 sowie die Bilder sind in Anhang\_B3 aufgeführt.

In allen drei Kontexten fielen den Probanden die Aufgaben im zweiten Aufgabenbereich, dem Erkennen von Evidenz, deutlich leichter als im ersten Aufgabenbereich, der Produktion von Evidenz. Beim zweiten Aufgabenbereich, dem Erkennen von Evidenz, erzielten die Vierjährigen im Vergleich zum ersten Bereich eine höhere Lösungsrate.

Beim dritten Aufgabenbereich (Produktion von Theorie) erreichten alle Probanden einen höheren Wert als beim zweiten. Hier ergab sich in der Gruppe der Neun- bis Zehnjährigen ein Deckeneffekt. Dementsprechend lagen hier die Lösungsraten sehr hoch (siehe Kapitel 7.2.6). Am leichtesten waren die Aufgaben zur Interpretation von Theorie, daher erzielten Kinder aller Altersgruppen hier die höchsten Lösungsraten. In der Gruppe der Neun- bis Zehnjährigen ergab sich ebenso ein Deckeneffekt (siehe Kapitel 7.2.6).

Mit dem Wissenschaftsverständnistest wurden die Kompetenzen der Kinder aller vier Altersstufen altersspezifisch erfasst. Die interne Konsistenz des Testes kann als gut bezeichnet werden. Bei der jüngsten Altersgruppe gab es einen Bodeneffekt im ersten Aufgabenbereich bei der Produktion von Evidenz. Allerdings wurden Kompetenzen der neun- bis zehnjährigen Kinder mit sehr gutem Wissenschaftsverständnis im dritten und vierten Aufgabenbereich, der Produktion von Theorie sowie der Interpretation von Theorie, nicht erschöpfend ermittelt. Hier trat ein Deckeneffekt auf. Eine komplexere Aufgabenstellung in diesen beiden Rubriken hätte

vermutlich besser die Identifikation von Kompetenzen der neun- bis zehnjährigen Kinder ermöglicht.

Zusammenfassend kann festgestellt werden, dass Kinder aller Altersstufen Kompetenzen beim Wissenschaftsverständnis demonstrieren konnten. Die älteren Kinder zeigten erwartungsgemäß höhere Kompetenzen als die jüngeren Kinder (Stephan-Gramberg et al., 2012b). Demzufolge konnte die zweite Hypothese bestätigt werden.

In der Studie wurde im Rahmen der dritten Forschungsfrage geprüft, ob ein Zusammenhang zwischen Wissenschaftsverständnis und Schlussfolgern besteht. Dafür wurden logistische Regressionsanalysen mit gemischten linearen Modellen (Tröbst et al., 2012a) berechnet. Als abhängige Variable wurden die im konditionalen Sinne korrekten Antworten genutzt. Dabei wurden die Faktoren Ereignis, Wissenschaftsverständnis und Alter schrittweise in ein Modell aufgenommen. Es ergab sich ein signifikanter Effekt für Wissenschaftsverständnis und Schlussfolgern. In weiteren Analysen wurden Interaktionen von Wissenschaftsverständnis und Einzeleignissen herausgebildet. Mit dem Alter nahm der Effekt des Wissenschaftsverständnisses beim Schlussfolgern ab. Insgesamt konnte die dritte Hypothese bestätigt werden, da sich ein signifikanter Zusammenhang zwischen Schlussfolgerungskompetenzen und Wissenschaftsverständnis herausstellte (siehe auch Stephan-Gramberg et al., 2012b).

Der Zusammenhang zwischen den drei Maßen der exekutiven Funktionen, Phonologische Schleife, Zentrale Exekutive sowie der Inhibition, und dem Schlussfolgern wurde im Rahmen des Berichtes für die Deutsche Forschungsgemeinschaft untersucht (Tröbst et al., 2012a). Dabei stellte sich die Inhibition (mit dem Früchte-Gemüse-Stroop-Test) bei einem Vergleich der drei Maße der exekutiven Funktionen als einflussreichste Prädiktorvariable dar. Die Ergebnisse der Querschnittstudie dienten als Grundlage für die Entwicklung der Trainingsmaßnahmen der Trainingsstudie.

## 8 Trainingsstudie

Die empirischen Befunde der Querschnittstudie haben die Kompetenzen von jungen Kindern beim Schlussfolgern dargelegt und dienten als Grundlage, um in einer zweiten Studie unter Laborbedingungen Erfolg versprechende Trainingsmaßnahmen zur Unterstützung von Denkprozessen beim Schlussfolgern zu entwickeln. Die Ergebnisse verdeutlichten, dass Kinder im Alter von fünf bis sechs Jahren Schwierigkeiten haben, eine Hypothese mit entsprechender Evidenz zu widerlegen. Weiterhin machten die Befunde deutlich, dass Kinder Probleme mit dem

Erkennen irrelevanter Evidenz sowie mit der Schlussfolgerung im Zusammenhang mit einer Hypothese hatten. Bei widerlegender und irrelevanter Evidenz neigten Kinder dazu, diese zu ignorieren bzw. diese falsch zu interpretieren. In der Regel tendierten besonders junge Kinder dazu, eine Hypothese durch Evidenz zu bestätigen. Der angemessene Umgang mit Evidenz stellt eine wichtige Grundlage beim naturwissenschaftlichen Lernen dar und sollte daher mit den Trainingsmaßnahmen effektiv gefördert werden. Die Ergebnisse der Querschnittstudie machten eine detaillierte Kompetenzeinschätzung von Kindern im Alter von fünf bis sechs Jahren möglich. So konnten die Trainingsmaßnahmen genau passend zum Kompetenzniveau der Probanden entwickelt werden.

In der Trainingsstudie wurde analog zur Querschnittstudie ebenso das Inhaltsgebiet *Elastizität und Plastizität* gewählt, damit in beiden Studien mit dem Schlussfolgerungstest als zentralem Instrument ein einheitliches Testformat mit identischem Inhaltsgebiet zur Verfügung stand. Anknüpfend an Pea (2004) und Van de Pol (2010) wurden die beiden Trainingsmaßnahmen entwickelt und ihre Effektivität mit Hilfe eines Prä-Post-Designs geprüft.

Zunächst sollen kurz die Grundzüge mikrogenetischer Studien skizziert werden, da die durchgeführte Trainingsstudie mit ihrem Prä-Post-Design und den beiden zwischengeschalteten Trainingssitzungen in Anlehnung an mikrogenetische Studien konzipiert worden ist. Mikrogenetische Studien eignen sich, um Lernprozesse und insbesondere konzeptuelle Umstrukturierungen bei der Verwendung von Strategien darzustellen, da das Vorgehen bei Durchführung und Analyse feingranular gestaltet ist (Siegler, 2006; Siegler & Stern, 1998). Dabei werden gleichartige Aufgaben in unterschiedlichen Sequenzen Kindern über einen Zeitraum von mehreren Tagen angeboten.

Die Beobachtungen werden detailliert analysiert, um dadurch konzeptuelle Umstrukturierungen und daraus resultierende Strategiewechsel herauszufiltern (Siegler, 1995; Siegler & Stern, 1998). Innerhalb der verschiedenen Sequenzen erfolgen detaillierte Beobachtungen, um beispielsweise Lösungsstrategien von Kindern zu erkennen. Während der Durchführung kommt es zu einer hohen Beobachtungsdichte im Vergleich zur Rate der konzeptuellen Veränderungen (Siegler, 2006; Siegler & Crowley, 1991; Siegler & Stern, 1998). Sind die Beobachtungsabstände gering, werden häufig keine oder nur kleine Entwicklungsschritte wahrgenommen. Vergrößert man die Beobachtungsabstände, so werden umfassendere Entwicklungsschritte festgestellt (Kuhn, 2006; Siegler, 2006). Die Lernentwicklung kann dabei sowohl bei einem Kind individuell beobachtet werden als auch bei einer Gruppe von Kindern (Kuhn, 2006).

Fünf charakteristische Dimensionen kognitiver Umstrukturierungsprozesse werden bei mikrogenetischen Analysen berücksichtigt. Dazu gehören der Prozess, die Häufigkeit der Anwendung, die Breite, die Variabilität sowie die Quelle des Wechsels (Siegler, 1995; Siegler, 2006): "Within this framework, distinctions are made among five dimensions of change: the path, rate, breadth, variability and source of change." (Siegler, 1995, S. 228).

Bei mikrogenetischen Studien zum Lernverhalten von Kindern ist es wichtig, zunächst den Verlauf der kognitiven Veränderung im Rahmen einer sequentiellen Analyse zu betrachten. Dabei wird untersucht, wie schnell eine begonnene Veränderung stattfindet und wann sie zu einem konsistenten Gebrauch einer fortgeschrittenen Strategie führt (Siegler & Stern, 1998). Dabei ist der Entwicklungsprozess als variabler Prozess anzusehen, bei dem unterschiedliche Kinder verschiedene Lösungswege wählen (Kuhn, 2006).

Häufig erfolgt ein Wechsel von Strategien als gradueller Prozess, wobei meist Strategien parallel verwendet werden; erst nach einiger Zeit entscheidet sich ein Kind für die komplexere Strategie (Siegler, 2006). Diese Entwicklung wird im Rahmen der Breite und Variabilität des Wechsels (Siegler, 1995, 2006) betrachtet. Beim Schlussfolgern im Zusammenhang mit Vermutungen geht es auch um Strategieanwendung, da Menschen auf Grundlage von Voraussetzungen (z.B. Eigenschaften von Gegenständen) sich für eine logische Konsequenz entscheiden. Empirische Befunde mikrogenetischer Studien belegten, dass unterschiedliche Faktoren Lernprozesse begünstigen, so zum Beispiel Feedback, direkte Instruktion sowie kooperatives Lernen (Siegler, 1995). In mikrogenetischen Studien werden Lernprozesse oft als Graph in einem Koordinatensystem dargestellt, wobei die Zeit auf der Abszisse und die Entwicklung auf der Ordinate eingetragen werden (Siegler, 2006).

In der vorliegenden Trainingsstudie wurden die Entwicklungsverläufe vom Prä-Test über die beiden Trainingssitzungen und den Post-Test bis hin zum Transfer-Test in Anlehnung an mikrogenetische Studien ebenfalls graphisch dargestellt. In den Studien von Siegler (1995, 2006) handelte es sich meist um sich überlappende Prozesse bei der Strategieanwendung. Im Unterschied zu den Studien von Siegler (1995, 2006) lag der Fokus in der Trainingsstudie entsprechend der zweiten Forschungsfrage auf der Überprüfung der Wirkung von Fördermöglichkeiten. Dazu wurden die Antworten der Kinder vom Prä-Test mit denen im Post-Test verglichen. Die Kinder wurden in ihren Denkprozessen beim Einschätzen von Wenn-Dann-Beziehungen mit drei unterschiedlichen Trainingsmaßnahmen unterstützt. Dabei handelte es sich um verschiedene Formen von Scaffolding-Maßnahmen, die adaptive Lerner-Unterstützung (Decristan et al., 2015), Kognitive Strukturierung (Einsiedler & Hardy, 2010; Hardy, 2012) sowie

Modellierung (Pea et al., 2004). Im Theorieteil (Kapitel 3) werden die Maßnahmen detailliert beschrieben.

### 8.1 Vorstudie

In der Vorstudie wurden die Wirkungen der Trainingsmaßnahmen, die für diese Studie entwickelt worden sind (Unterstützung durch Visualisierung, Scaffolding durch adaptives Nachfragen, Unterstützung durch Modellierung und adaptives Nachfragen) untersucht.

Beide Maßnahmen der Trainingsstudie wurden adaptiv verwendet, das heißt im Sinne des Kontingenzprinzips (Van de Pol et al., 2010) reduzierte der Versuchsleiter die Unterstützung, wenn ein Kind richtig antwortete; gleichzeitig übernahm ein Kind zunehmend Verantwortung für sein eigenes Lernen. Der Erwachsene agierte gemäß des „*Fading*“ (Van de Pol et al., 2010, S. 278; siehe Kapitel 3.2).

Ziel der vorliegenden Untersuchung war es, in einem Prä-Post-Design beide Maßnahmen anhand der Lernzuwächse von Kindern beider Trainingsgruppen (Scaffolding durch adaptives Nachfragen, Unterstützung durch Modellierung und adaptives Nachfragen) im Post-Test miteinander zu vergleichen. Die Antworten von Kindern in beiden Trainingsgruppen wurden im Post-Test mit denen der Kontrollgruppe verglichen. Die Kinder der Kontrollgruppe erhielten keine Unterstützung, sie erfüllten ausschließlich die Aufgaben des Prä- und Post-Tests. Zum einen wurde in den Analysen geprüft, ob die Probanden der Trainingsgruppen im Post-Test deutlich besser als die der Kontrollgruppe abschnitten. Zusätzlich wurde untersucht, ob sich Unterschiede zwischen den Probanden beider Trainingsgruppen identifizieren ließen. Durch die Vorstudie mit ihrem Prä-Post-Testdesign wurde untersucht, ob die entwickelten Fördermaßnahmen geeignete Instrumente sind, um einen nachprüfbaren Lernzuwachs bei Kindern im Elementarbereich hervorzurufen. Falls die Trainingsmaßnahmen effektiv gewesen sind, erreichen die Probanden der Trainingsgruppen höhere Werte bei den fortgeschrittenen Interpretationsniveaus, dem bikonditionalen bzw. konditionalen Niveau als Probanden der Kontrollgruppe.

#### 8.1.1 Fragestellung und Hypothese

Frage 1: Lässt sich die Schlussfolgerungsfähigkeit von Kindergartenkindern durch ein Training mit adaptiver Unterstützung bzw. mit Modellierung fördern?

Hypothese 1: Kinder der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) sowie Kinder der Trainingsgruppe 2 (Unterstützung durch Modellierung und zusätzliches

Nachfragen) sind Kindern der Kontrollgruppe überlegen. Daher äußern Probanden beider Trainingsgruppen häufiger Antworten auf einem fortgeschrittenen Niveau (bikonditional bzw. konditional) als Probanden der Kontrollgruppe. Die Kompetenzniveaus (konjunktiv, bikonditional, konditional, RFFR, RFRF) werden als abhängige Variablen bei der Berechnung mit *MANOVA* genutzt

### 8.1.2 Methoden

#### *Vorwissenstest*

Beim Vorwissenstest wurden den Kindern insgesamt drei Vermutungen vorgelegt. Die erste Vermutung eines Mädchens war „Dinge, die leicht sind, springen.“ Dazu wurden den Probanden zwei Gegenstände vorgelegt; anhand der Styroporkugel konnte die zuvor genannte Vermutung bestätigt und mit der Watte konnte die Vermutung widerlegt werden. Die Probanden durften die Gegenstände bei diesem Test nicht in die Hand nehmen, um das Sprungverhalten der Gegenstände auszuprobieren. Die Probanden sollten ausschließlich auf Grundlage ihrer Vorannahmen eine Vermutung einschätzen. Diese wurde auf einer vierstufigen Skala (stimmt, stimmt eher, stimmt eher nicht, stimmt nicht) erfasst. Dabei symbolisierten zwei nach oben gerichtete Daumen die Stufe *stimmt*, ein nach oben gerichteter Daumen die Stufe *stimmt eher*, ein nach unten gerichteter Daumen die Stufe *stimmt eher nicht* und zwei nach unten gerichtete Daumen die Stufe *stimmt nicht* (siehe Abbildung 18).



Abbildung 18. Vorwissenstest, Symbole für Einschätzung zu Vermutungen.

Die Probanden konnten auf verschiedenen Daumen ihre Einschätzung demonstrieren. Dazu legten sie ein Kreuz auf die Abbildung, die ihrer Meinung nach am besten im Zusammenhang mit einer zuvor genannten Vermutung passte. Detailinformationen zur Durchführung des Vorwissenstests finden sich im Protokollbogen (Anhang F\_1), im Interviewleitfaden (Anhang F\_2); die Bilder sind in Anhang F\_3 abgebildet.

#### *Scaffolding-Maßnahmen*

In der Trainingsstudie wurden zwei Scaffolding-Maßnahmen (Scaffolding durch adaptives Nachfragen, Scaffolding durch Modellierung und adaptives Nachfragen) in jeweils zwei Trainingssitzungen durchgeführt, wobei jeweils eine Vermutung überprüft wurde. Bei der ersten Sitzung wurde die Vermutung „Dinge, die hart sind, springen“ mit vier Gegenständen (Murmel, Knetkugel, Softball, Jonglierball), bei der zweiten Trainingssitzung die Vermutung „Dinge, die mit Luft gefüllt sind, springen“ wiederum mit vier Gegenständen überprüft. Die Vermutungen waren in beiden Scaffolding-Maßnahmen (Scaffolding durch adaptives Nachfragen, Scaffolding durch Modellierung und adaptivem Nachfragen) identisch.

Jede Maßnahme umfasste einen Prozess aus drei Schritten (siehe Abbildung 19), wobei die ersten beiden Schritte mittels Scaffolding durch Visualisierung in beiden Fördergruppen vollkommen gleich abliefen. Eine Übersicht zur Durchführung beider Trainingsmaßnahmen ist in Tabelle 24 zu sehen.

Tabelle 23

*Übersicht über Vorgehen bei Durchführung der Scaffolding-Maßnahmen*

Scaffolding-Maßnahmen	Visualisierung der Eigenschaft eines Gegenstandes (hart, weich)	Visualisierung des Sprungverhaltens (springt, springt nicht)	Evaluation des Sprungverhaltens in Bezug auf die Vermutung
Scaffolding mit adaptivem Nachfragen	✓	✓	Nachfrage falls Evaluation nicht korrekt, erneute Zuordnung von Eigenschaft, Sprungverhalten, nach erneuter Antwort keine weitere Rückmeldung
Scaffolding durch Modellierung und adaptives Nachfragen	✓	✓	Vollständige Modellierung bei erster Vermutung, Modellierung beim ersten Gegenstand der zweiten Vermutung, Nachfrage falls Evaluation nicht korrekt, erneute Zuordnung von Eigenschaft, Sprungverhalten, nach erneuter Antwort keine weitere Rückmeldung

Im ersten Schritt ordnete ein Kind vier Gegenstände (Murmel, Knete, Softball, Jonglierball) passend zur Vermutung „Dinge, die hart sind, springen.“ nach der Eigenschaft (hart, weich). Die Kategorien wurden mithilfe zweier Platten (Eisenplatte, mit Watte belegte Platte) visualisiert: Eine Eisenplatte stand für harte Gegenstände, eine mit Watte verkleidete Platte symbolisierte die Eigenschaft weich. Beim zweiten Schritt wurde nach Eigenschaft und

Sprungverhalten geordnet (siehe Tabelle 24). Zur Visualisierung des zweiten Schrittes legte der Versuchsleiter jeweils zwei Karten für das Sprungverhalten (springt, springt nicht) auf eine Eisenplatte bzw. auf eine mit Watte belegte Platte. Im ersten Teil des dritten Schrittes erklärte der Versuchsleiter, was unter einer Vermutung zu verstehen war. Im zweiten Teil des dritten Schrittes (siehe Abbildung 19) erfolgte jeweils die Evaluation des Ereignisses durch das Kind (bei Trainingsgruppe 1 nur adaptives Nachfragen in Trainingssitzung 1 und 2) oder durch den Versuchsleiter (bei Trainingsgruppe mit Modellierung in Trainingssitzung 1). Der detaillierte Ablauf einer Trainingssitzung ist in den Interviewleitfäden der beiden Trainingsmaßnahmen nachzulesen. Der Interviewleitfaden zur Trainingsgruppe 1 für die erste Trainingssitzung ist im Anhang H\_2, für die zweite Trainingssitzung im Anhang I\_2 zu finden. Die Interviewleitfäden für die Trainingsgruppe 2 sind in Anhang J\_2; Anhang K\_2 zu finden.





Abbildung 19. Trainingsprozess in drei Schritten.

Im zweiten Schritt (siehe Abbildung 20) wurden die Gegenstände zusätzlich zu ihren Eigenschaften (hart, weich) nach ihrem Sprungverhalten zugeordnet. Dazu legte der Versuchsleiter jeweils zwei Karten für das Sprungverhalten (springt, springt nicht) auf die beiden Platten (hart, weich). Für jede Eigenschaft (hart, weich) eines Gegenstandes gab es zwei Möglichkeiten (springt, springt nicht). Auf der harten bzw. weichen Platte wurden je zwei Karten (springt, springt nicht) platziert. Jeder Gegenstand konnte bei richtiger Zuordnung nur auf einer Platte und Karte liegen, so dass bei vier Gegenständen jeder Platz belegt war.

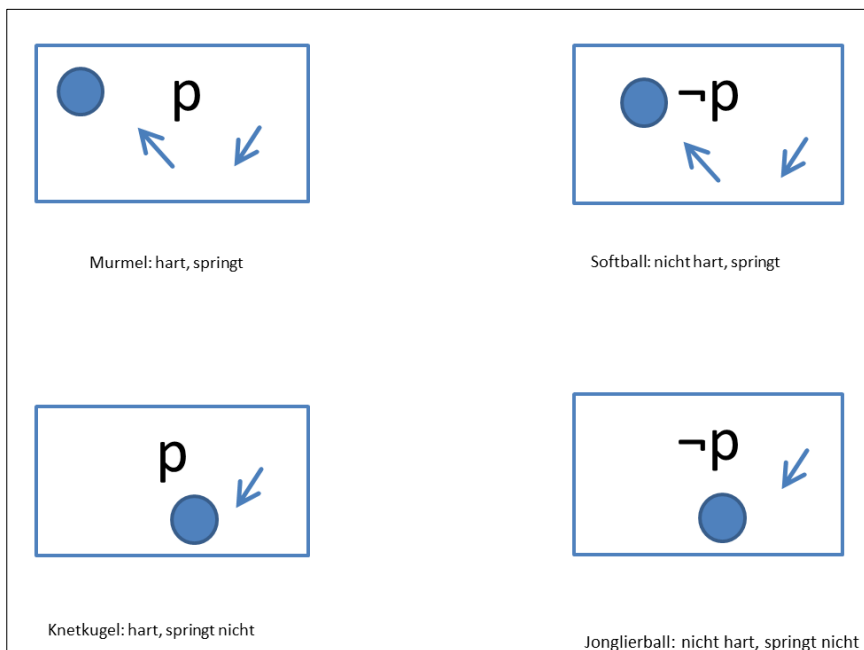


Abbildung 20. Visualisierung bei der Scaffolding -Maßnahme (zweiter Schritt).

In der folgenden Tabelle (Tabelle 25) sind die vier in der Studie verwendeten Gegenstände bei der Vermutung „Dinge, die hart sind, springen“ zusammengestellt. Dabei ist jeweils die mit der Vermutung zu überprüfende Eigenschaft ( $p/\neg p$ ) sowie das Sprungverhalten des einzelnen Gegenstandes ( $q/\neg q$ ) festgehalten.

Tabelle 24

*Evaluation von Ereignissen im Zusammenhang mit einer Vermutung*

Gegenstand	Eigenschaft hart	Ereignis springt	Auswirkung auf Schlussfolgerung
Murmel	$p$	$q$	bestätigt
Knetkugel	$p$	$\neg q$	widerlegt
Softball	$\neg p$	$q$	irrelevant
Jonglierball	$\neg p$	$\neg q$	irrelevant

Im dritten Schritt erfolgte die Evaluation eines Gegenstandes im Zusammenhang mit einer Vermutung. Dabei wurde in der ersten Trainingsgruppe „Scaffolding durch adaptives Nachfragen“ bei einer nicht erwartungsgemäßen Antwort ein Kind erneut gefragt (siehe Interviewleitfaden Anhang H\_2); in der zweiten Trainingsgruppe „Scaffolding durch Modellierung und adaptives Nachfragen“ unterstützte der Versuchsleiter das Denken eines Kindes während der ersten Vermutung vollständig durch Nennen der angemessenen Evaluation (siehe Interviewleitfaden Anhang J\_2).

*Transfer-Test im Inhaltsgebiet Schwimmen und Sinken*

Der Aufbau sowie die Durchführung des Transfer-Testes im Inhaltsgebiet *Schwimmen und Sinken* wurden vom Schlussfolgerungstest übernommen (siehe Anhang L\_1; L\_2). Den Kindern wurden drei Vermutungen, zwei positive Vermutungen sowie eine negativ formulierte mit vier Gegenständen [ $p \ q \ \text{best}$ ], [ $p \neg q \ \text{wid}$ ], [ $\neg p \ q \ \text{irrel}$ ] und [ $\neg p \neg q \ \text{irrel}$ ] präsentiert. Dazu wurden typische Vermutungen von Kindern wie beispielsweise „Dinge, die mit Luft gefüllt sind, schwimmen“ mithilfe von vier Gegenständen (siehe Tabelle 26) überprüft. Dieser Themenbereich eignete sich, da bereits junge Kinder Alltagserfahrungen durch Spielen mit Wasser in diesem Bereich gesammelt haben. In gleicher Weise wie beim Schlussfolgerungstest wurden die Gegenstände bei den Vermutungen getauscht, um einen Reihenfolgeeffekt zu vermeiden.

In diesem Test wurde geprüft, ob Kinder nach dem Training ihre erworbenen Schlussfolgerungskompetenzen in einem Transfer-Kontext anwenden konnten.

Tabelle 25

*Vermutungen des Transfer-Tests*

Vermutung	Gegenstände	bestätigend [p q]	widerlegend [p¬q]	irrelevant [¬p q]	irrelevant [¬p¬q]
„Dinge, die mit Luft gefüllt sind, schwimmen.“	Spielball	x			
	Fensterkitt-Ball		x		
	Styroporstück			x	
	Stein				x
„Dinge, die klein sind, schwimmen.“	Bimsstein	x			
	5-Cent-Stück		x		
	Würfel			x	
	Marmorstein (groß)				x
„Dinge, die schwer sind, gehen unter.“	Eisenstück	x			
	Holzklötz		x		
	Büroklammer			x	
	Wasserball				x

*Argumentationstest*

In diesem Test ging es darum zu prüfen, ob die Probanden in der Lage sind, mit passenden Belegen adäquat zu begründen, ob eine Vermutung stimmt oder nicht.

Im Argumentationstest wurde den Kindern anhand von Spielfiguren eine Gesprächssituation präsentiert. Ein Junge, Tim, war davon überzeugt, dass „Dinge, die leicht sind, springen“. Als Beleg für seine Vermutung zeigte er seiner Freundin, Anna, eine Styroporkugel. Sie glaubte, dass Tim mit seiner Vermutung nicht richtig lag. Sie konnte ihre Ansicht mit einem Wattebausch begründen. Anna hatte einen Beleg für ihre Auffassung, die Tims Vermutung widerlegte. Der Wattebausch war zwar leicht, sprang aber nicht.

In der ersten Aufgabe wurden die Probanden gebeten, ihre Meinung zu äußern; anschließend sollten die Probanden diese auch begründen. Dabei galt eine Begründung als vollständig,

wenn die Probanden in ihrer Argumentation sowohl das Antezedens (die Eigenschaft z.B. leicht) als auch die Konsequenz (das Sprungverhalten z.B. springt) berücksichtigten.

Aufgabe eins und drei sind als dichotome Aufgaben konzipiert (siehe Anhang M\_1). Aufgabe zwei enthielt ein offenes Antwortformat, bei der ein oder zwei Punkte erreicht werden konnten (siehe Anhang M\_1). Dann wurde die Antwort als *richtig* bewertet. Die Antworten wurden dichotom kodiert. Eine richtige Antwort wurde mit *eins* bewertet; mit *null* wurde eine falsche Antwort eingestuft. In der zweiten Aufgabe wurden insgesamt drei Gegenstände präsentiert (siehe Tabelle 27). Die Probanden sollten den Gegenstand auswählen, der die Vermutung „Dinge, die leicht sind, springen“ widerlegte. Mit der leeren Tüte wurde die zuvor genannte Aussage widerlegt.

Tabelle 26

*Übersicht über die Eigenschaften der Gegenstände*

Gegenstand	Entscheidung im Zusammenhang mit Vermutung	
Eigenschaft, Sprungverhalten	leicht	springt
Luftballon	trifft zu	trifft zu
leere Tüte	trifft zu	trifft nicht zu
Luftballon mit Sand	trifft nicht zu	trifft nicht zu

Die Antworten im Argumentationstest wurden kodiert (siehe Tabelle 28). Eine falsche bzw. nicht zu bewertende Antwort wurde mit 0 kodiert. Bei einer richtigen Antwort ohne Begründung wurde die Antwort mit 1 kodiert; eine richtige Antwort mit Begründung wurde mit zwei kodiert (siehe Tabelle 28). Insgesamt konnten vier Punkte im Argumentationstest erreicht werden.

Tabelle 27

*Kodierung der Antworten*

Frage	Antwort (Kodierung)
1. „Denkst du, Anna kann mit dem Wattebausch zeigen, dass Tims Vermutung nicht stimmt?“	0 ⇔ weiß nicht, nein 1 ⇔ ja
2. „Wieso glaubst du das?“	0 ⇔ weiß nicht, 1 ⇔ weil Watte nicht springt 2 ⇔ weil Watte leicht ist und nicht springt
3. „Ist ein Gegenbeispiel dabei, mit dem du zeigen kannst, dass die Vermutung ‚Dinge, die leicht sind, springen‘ nicht stimmt?“	0 ⇔ nein 1 ⇔ ja, Tüte

*8.1.3 Stichprobe*

An der Vorstudie nahmen insgesamt 45 Kinder im Alter von 5,5 Jahren (SD = 0.39) aus Kindergärten des Main-Taunus-Kreises in Hessen teil. Die Kinder wohnten in Kleinstädten mit weniger als 50.000 Einwohnern. Die Probanden der Studie wurden zufällig in drei gleichgroße Gruppen mit jeweils 15 Kindern aufgeteilt.

*8.1.4 Design*

Die Vorstudie wurde an zwei Testtagen durchgeführt. In allen drei Gruppen (Trainingsgruppe 1 und 2, Kontrollgruppe) wurden Prä-Test sowie Post-Test im Kontext *Elastizität und Plastizität*, der Wissenschaftsverständnistest durchgeführt. In den beiden Trainingsgruppen wurde jeweils ein Training an beiden Testtagen durchgeführt (siehe Tabelle 29).

Tabelle 28

*Design der Vorstudie*

	Tag 1	Tag 2
KG	Prä-Test Schlussfolgerungstest im Kontext <i>Elastizität und Plastizität</i> Wissenschaftsverständnistest	Post-Test Schlussfolgerungstest im Kontext <i>Elastizität und Plastizität</i>
TG 1	Prä-Test Schlussfolgerungstest im Kontext <i>Elastizität und Plastizität</i> Wissenschaftsverständnistest Training 1	Training 2 Post-Test Schlussfolgerungstest im Kontext <i>Elastizität und Plastizität</i>
TG 2	Prä-Test Schlussfolgerungstest im Kontext <i>Elastizität und Plastizität</i> Wissenschaftsverständnistest Training 1	Training 2 Post-Test Schlussfolgerungstest im Kontext <i>Elastizität und Plastizität</i>

In Trainingsgruppe 1 wurde die Förderung des Schlussfolgerns in zwei Sitzungen mit Scaffolding durch adaptives Nachfragen durchgeführt (siehe Kapitel 8.1.4).

In Trainingsgruppe 2 wurden die Kinder in der ersten Trainingssitzung mit Scaffolding durch Modellierung und in der zweiten Trainingssitzung bei drei Gegenständen mit Scaffolding durch adaptives Nachfragen unterstützt. Folglich war die Förderbedingung in beiden Trainingsgruppen in der zweiten Trainingssitzung ab dem zweiten Gegenstand identisch (siehe Anhang I\_2, siehe Anhang K\_2).

### 8.1.5 Vorgehen bei der Durchführung

Die Antworten der Kinder der beiden Trainingsgruppen (Scaffolding durch adaptives Nachfragen, Unterstützung durch Modellierung und adaptives Nachfragen) wurden mit denen der Kontrollgruppe im Rahmen der Vorstudie verglichen. Erste Ergebnisse wurden in einem kurzen Artikel veröffentlicht (Stephan-Gramberg & Hardy, 2012a).

Die Studie wurde im Rahmen einer Bachelor- sowie einer Staatsexamensarbeit durchgeführt, daher übernahmen die beiden Examenskandidaten die Durchführung der Fördermaßnahmen (Scaffolding durch adaptives Nachfragen, Scaffolding durch Modellierung und adaptives Nachfragen) in beiden Trainingsphasen. Das Vorgehen in den Trainingsphasen war bei beiden Gruppen in drei Schritte eingeteilt (siehe Kapitel. 8.1.2).

Die Erhebungen des Prä-Tests und Post-Tests wurden entweder von einer studentischen Hilfskraft oder einer wissenschaftlichen Mitarbeiterin durchgeführt. Drei positiv formulierte, jeweils identische Vermutungen wurden im Prä- und Post-Test überprüft. Auf diese Weise wurden möglichst vergleichbare Bedingungen in beiden Testphasen geschaffen. Die Gegenstände zum Prüfen einer Vermutung waren in beiden Testphasen (Prä- und Post-Test) unterschiedlich, um Wiedererkennungseffekte bei den Kindern zu vermeiden (siehe Anhänge G\_2, G\_3).

### 8.1.6 Analysestrategien bei der Auswertung

Verknüpft mit der Forschungsfrage 1 und 2 wurden die Antworten der Probanden im Prä-Test und Post-Test miteinander verglichen. Forschungsfrage 1 zielte darauf ab, zu prüfen, ob sich Unterschiede zwischen den Gruppen (Trainingsgruppe 1, Trainingsgruppe 2, Kontrollgruppe) im Post-Test nachweisen ließen. Dazu wurde für jedes Ereignis  $[p \mid q \text{ best}]$ ;  $[p \mid \neg q \text{ wid}]$ ;  $[\neg p \mid q \text{ irrel}]$ ;  $[\neg p \mid \neg q \text{ irrel}]$  ein Differenzwert vom Post-Test und Prä-Test gebildet. Anschließend wurden relative Häufigkeiten der Interpretationsniveaus vom Prä-Test und Post-Test gebildet. Mit einer *MANOVA* wurde geprüft, ob signifikante Mittelwertsunterschiede zwischen den einzelnen Gruppen bestanden bzw., ob sich signifikante Veränderungen vom Prä-Test hin zum Post-Test ergaben. Nachfolgend wurden T-Tests für abhängige Stichproben im Zusammenhang mit den einzelnen Interpretationsniveaus berechnet, um zu untersuchen, ob sich signifikante Veränderungen bei den einzelnen Niveaus zeigen ließen.

### 8.1.7 Ergebnisse

Mit der folgenden Analyse wird im Rahmen des ersten Teils der ersten Forschungsfrage (Forschungsfrage 1a) geprüft, ob sich die Schlussfolgerungsfähigkeit von Kindern im Alter von fünf bis sechs Jahren fördern lässt, beantwortet. Es wurde mit Hilfe der Werte des Prä- bzw. Post-Tests bei den Einzelereignissen  $[p \mid q]$ ,  $[p \mid \neg q]$ ,  $[\neg p \mid q]$ ,  $[\neg p \mid \neg q]$  eine Differenz gebildet. Diese wurden jeweils als abhängige Variable bei der Berechnung der Mittelwerte genutzt.

Bei der Betrachtung der Differenzwerte der Einzelereignisse beim Schlussfolgerungstest im Post- und Prä-Test ergab sich in der Trainingsgruppe 2 (Scaffolding durch Modellierung und adaptives Nachfragen) beim Ereignis  $[\neg p \mid \neg q]$  ein signifikanter Zuwachs an richtigen Antworten. Sowohl in der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) als auch in der Kontrollgruppe gab es keine signifikanten Veränderungen (siehe Tabelle 30).

Tabelle 29

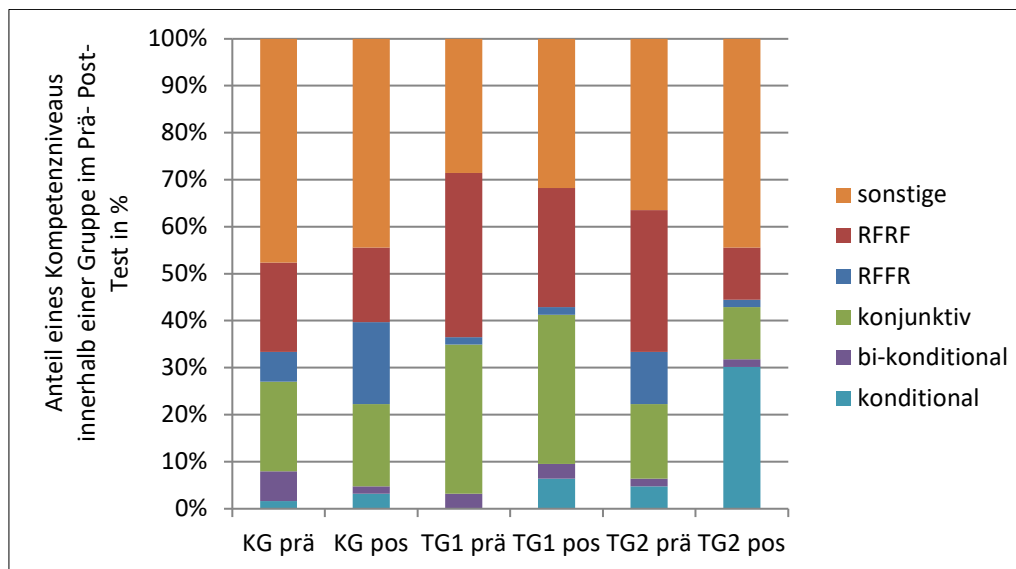
Differenzwerte Post-Test und Prä-Test (Standardabweichungen)

Ereignis	Training 1	Training 2	Kontrollgruppe
[p q]	-0.13 (0.35)	-0.27 (0.46)	-0.07 (0.59)
[p¬q]	-0.13 (1.12)	-0.13 (0.92)	-0.27 (0.92)
[¬pq]	0.13 (0.92)	0.4 (1.35)	0.13 (0.92)
[¬p¬q]	0.47 (1.13)	1.13 (1.50)*	0 (0.85)

\* MANOVA  $F(4, 39) = 3.89$ ; ns.,  $p < 0.05$   $\epsilon^2 = 0.22$

Zur Beantwortung des zweiten Teils der ersten Forschungsfrage (Forschungsfrage 1b) wurden aus den Einzelereignissen [ p q], [p¬q], [¬p q], [¬p¬q] Interpretationsniveaus gebildet (Beschreibung siehe Kapitel 7.1.4, Tabelle 10).

Die relativen Häufigkeiten in den Interpretationsniveaus in den Gruppen zeigten zwar eine Veränderung der Antwortmuster zwischen Prä- und Post-Test, diese erwiesen sich jedoch in einem Gruppenvergleich mit MANOVA mit  $F(10, 74) = 0.82$ ;  $\epsilon^2 = 1.0$  als nicht signifikant (siehe Abbildung 21).



Anmerkung. MANOVA  $F(10, 74) = 0.82$ ; ns.;  $\epsilon^2 = 1.0$

Abbildung 21. Relative Häufigkeiten beim Schlussfolgern im Prä und Post-Test bei allen drei Gruppen.



Allerdings wurde beim Betrachten der Prä-Test-Werte in beiden Trainingsgruppen offensichtlich, dass diese sich deutlich unterschieden. Beispielsweise hatte die Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) mit 32% doppelt so viele Antworten auf dem konjunktiven Antwortniveau wie die Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen); sie hatten im Prä-Test nur 16% auf diesem Niveau.

In einzelnen Vergleichen der Zuwächse in den fortgeschrittenen Antwortmustern konjunktiv, bikonditional und konditional ergaben sich bei der Trainingsgruppe 1 sowie bei der Kontrollgruppe keine signifikanten Veränderungen bei den Interpretationsniveaus. Im Gegensatz hierzu war bei der Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) in T-Tests für abhängige Stichproben ein signifikanter Zuwachs beim konjunktiven und beim bikonditionalen Interpretationsniveau festzustellen. Hier ergab sich ein signifikanter Unterschied zwischen Prä-Test und Post-Test. Der T-Test zeigte in der Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptivem Nachfragen) beim konjunktiven Muster einen Wert von  $t(14) = 2.3$ ;  $p < 0.05$  und beim bikonditionalen Muster  $t(14) = 2.8$ ;  $p < 0.05$ , in der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) beim konjunktiven Niveau  $t(14) = 0$ ;  $p > 0.05$  ns. und beim bikonditionalen Niveau  $t(14) = -1.0$ ;  $p > 0.05$  ns.

Die Ergebnisse der Vorstudie beim Schlussfolgerungstest konnten als Validierungshinweise für eine effektive Wirkung der Trainingsmaßnahmen der Gruppe durch Modellierung im Hinblick auf die Lernzuwächse gewertet werden. Insbesondere war es bei der Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptivem Nachfragen) gelungen, mithilfe der Trainingsmaßnahme ein besseres Verständnis für die schwierigen Ereignisse mit  $[\neg p]$  zu erreichen. Dies wurde bei der Analyse auf Grundlage der Einzelereignisse an den signifikanten Ergebnissen deutlich (siehe Tabelle 38). In einer multivariaten Analyse mit *MANOVA* mit  $F(10, 74) = 0.82$ ; ns.;  $\epsilon^2 = 1.0$  hatte sich beim Vergleich der Differenzwerte der Interpretationsmuster (RFFR, RFRF, konjunktiv, bikonditional, konditional) kein signifikanter Effekt gezeigt.

Die Antworten des Wissenschaftsverständnistests (dichotom, 0 = falsch; 1 = richtig) wurden nach dem zuvor beschriebenen Kodier-Schema von zwei unabhängigen Beobachtern kodiert (siehe Kapitel 8.1.6). Die Übereinstimmung wurde mithilfe des Kappa-Werts berechnet. Die innere Konsistenz des Tests war mit einem Kappa-Wert von 0.74 als akzeptabel anzusehen. Die Mittelwerte der fünf- bis sechsjährigen Kinder waren vergleichbar mit denen derselben Altersgruppe aus der Querschnittstudie (siehe Kapitel 7.2.5, Tabelle 20) und sind daher hier nicht erneut dargestellt.

Detaillierte Analysen des Wissenschaftsverständnistests der Trainingsstudie werden in Kapitel 7.2.6. im Rahmen der Hauptstudie dargestellt.

## 8.2 Hauptstudie

Die Hauptstudie wurde mit Blick auf die Trainingsmaßnahmen sowie das Prä-Post-Design in ähnlicher Form wie die Vorstudie mit zwei Trainingsphasen durchgeführt. So erfolgte die Erhebung des Prä-Tests sowie im Post-Tests im Inhaltsgebiet *Elastizität und Plastizität*. Zusätzlich zum Prä und Post-Test aus der Vorstudie wurden in der Hauptstudie neben dem Schlussfolgerungs- und Wissenschaftsverständnistest noch weitere Tests durchgeführt: ein Inhibitionstest (Früchte-Gemüse-Stroop), ein Transfer-Test im Kontext *Schwimmen und Sinken* sowie ein Argumentationstest (AGT).

Der Inhibitionstest wurde zur Parallelisierung zur Einteilung gleichstarker Gruppen genutzt. In Kapitel 7.2.2 ist die Durchführung des Tests beschrieben. Kinder mit gleichen bzw. ähnlichen Werten wurden zufällig jeweils auf eine der drei Gruppen verteilt. Bei der Trainingsstudie handelte es sich daher um eine experimentelle Studie. Der Transfer-Test im Anschluss an den Post-Test prüfte, ob die Trainingsmaßnahmen eine nachhaltige Wirkung auf das Schlussfolgern von Kindergartenkindern ausübten. Zum Abschluss der Testphase wurde ein Argumentationstest durchgeführt. Er gab Aufschluss darüber, ob Kinder im Alter von fünf bis sechs Jahren generalisierte Vorstellungen über Sprungverhalten von Gegenständen entwickelten und umfassende Aussagen treffen konnten.

Aufgrund der Erfahrungen aus der Vorstudie wurden in der Hauptstudie einige Änderungen vorgenommen, die hier kurz erläutert werden. Die Vorstudie zeigte, dass einige Kinder Schwierigkeiten hatten, bei einem Tennisball zu erkennen, ob er mit bzw. ohne Luft gefüllt war. Daher wurde dieser in der Hauptstudie durch einen herkömmlichen Spielball ersetzt. Außerdem wurden in der Hauptstudie statt der mehrfarbigen Smileys (grün, gelb, rot) gleichfarbige Smileys als Symbole für Zustimmung, Widerlegung bzw. Irrelevanz im Zusammenhang mit einer Vermutung verwendet. Die Beobachtungen der Versuchsleiter während der Erhebungsphase der Vorstudie deuteten darauf hin, dass die unterschiedlichen Farben der Smileys manche Kinder in ihrer Evaluation der Ereignisse beeinflussten. Womöglich verknüpften sie die Farbe Rot mit Verbotshinweisen, die Farbe Grün mit Geboten sowie die Farbe Gelb mit Warnhinweisen. Daher wurden in der Hauptstudie alle Smileys als Symbol in einer einheitlichen Farbe (Orange) gewählt. Stattdessen wurden unterschiedliche Smiley-Symbole (lachender, trauriger, nachdenklicher Smiley) für die Kennzeichnung der Evaluationen gewählt.

### 8.2.1 Fragestellungen und Hypothesen

Im folgenden Absatz werden die sechs Forschungsfragen mit den dazugehörigen Hypothesen der Hauptstudie vorgestellt.

Frage 1a: Lässt sich die Schlussfolgerungsfähigkeit von Kindergartenkindern durch ein Training mit adaptiver Unterstützung bzw. mit Modellierung fördern und inwiefern zeigt sich dies bei den einzelnen Evaluationen?

Hypothese 1a: Kinder der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) und der Trainingsgruppe 2 (Scaffolding durch Modellierung und adaptives Nachfragen) können deutlich besser mit den irrelevanten Ereignissen  $[\neg p \vee q]$ ;  $[\neg p \wedge \neg q]$  umgehen als Kinder der Kontrollgruppe. Bei der Analyse werden die Einzelereignisse als abhängige Variablen genutzt.

Frage 1b: Lässt sich die Schlussfolgerungsfähigkeit von Kindergartenkindern durch ein Training mit adaptiver Unterstützung bzw. mit Modellierung fördern und inwiefern zeigt sich dies im Hinblick auf die Interpretationsniveaus?

Hypothese 1b: Kinder der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) und der Trainingsgruppe 2 (Scaffolding durch Modellierung und zusätzliches Nachfragen) sind der Kontrollgruppe überlegen. Daher äußern Kinder beider Trainingsgruppen häufiger Antworten auf einem bikonditionalen oder konditionalen Niveau als Probanden der Kontrollgruppe. Bei der Analyse werden die Interpretationsniveaus als abhängige Variablen genutzt.

Frage 2: Zeigen sich Unterschiede zwischen zwei Formen des Scaffolding (Scaffolding durch adaptives Nachfragen, Unterstützung durch Modellierung und adaptives Nachfragen) beim Vergleich vom Prä-Test und dem Post-Test?

Hypothese 2: Kinder der Trainingsgruppe 2 (Scaffolding durch Modellierung und zusätzliches Nachfragen) sind Kindern der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) im Post-Test überlegen.

Frage 3: Lassen sich nachhaltige Trainingseffekte in den drei Experimentalgruppen über die Zeit darstellen?

Hypothese 3: Kinder der Trainingsgruppe 2 zeigen ein überlegenes Antwortverhalten sowohl im Vergleich zur Trainingsgruppe 1 als auch im Vergleich zur Kontrollgruppe. Die Werte der Kinder von Trainingsgruppe 1 und 2 bei den Einzelereignissen sind daher höher als die bei der Kontrollgruppe.

Frage 4: Zeigen sich durch das Training mit Evidenz Effekte beim Vergleich der Trainingsdomäne und des Transfer-Tests?

Hypothese 4: Das Antwortverhalten der Kinder beider Trainingsgruppen (Scaffolding durch adaptives Nachfragen, Scaffolding durch Modellierung und zusätzliches Nachfragen) im Transfer-Test bei den Einzelereignissen ist gleich dem Antwortverhalten der Kinder während der zweiten Trainingssitzung. Bei der Berechnung werden jeweils die Einzelereignisse als abhängige Variable genutzt.

In Frage 5 wurden die Fähigkeiten zum Argumentieren (siehe Kapitel 8.1.6) mit einem adaptierten Test, dem Argumentationstest (Beschreibung siehe Kapitel 8.1.2.4) geprüft. In diesem Zusammenhang wurde folgende Frage untersucht.

Frage 5: Zeigen sich signifikante Unterschiede zwischen den Gruppen bei den Antworten im Argumentationstest?

Hypothese 5: Probanden der Trainingsgruppe 2 (Scaffolding durch Modellierung und zusätzliches Nachfragen) können besser argumentieren als Kinder der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) sowie der Kontrollgruppe und haben daher höhere Summenwerte bei den Antworten im Argumentationstest. Als abhängige Variable wurde jeweils der Summenwert in allen drei Gruppen genutzt.

Befunde verschiedener Studien (Brown et al., 2010; Tytler, 2005) zeigten, dass Menschen mit ausgeprägten Kompetenzen im Schlussfolgern besser in der Lage waren, logisch zu argumentieren. Mithilfe des Argumentationstests sollte in der vorliegenden Studie herausgefunden werden, inwiefern ein Zusammenhang zwischen einer Schlussfolgerungskompetenz und der Fähigkeit zum Argumentieren (siehe Kapitel 5.4) bestand.

Frage 6: Gibt es einen Zusammenhang zwischen der Schlussfolgerungskompetenz und dem folgerichtigen Begründen im Argumentationstest?

Hypothese 6: Kinder, die konditional schlussfolgern, können im Argumentationstest das passende Beispiel nennen, mit dem sie eine gegebene Vermutung widerlegen können (Beispiel für eine Vermutung: „Dinge, die mit Luft gefüllt sind, springen“). Als

### 8.2.2 Methoden

Das Training erfolgte in gleicher Weise wie es vorher in der Vorstudie durchgeführt worden war (siehe Kapitel 8.2.2).

### 8.2.3 Stichprobe

Die Stichprobe der Trainingsstudie bestand aus 63 Kindern, die durchschnittlich 5.9 (SD =0.4). Jahre alt waren. Dabei handelte es sich um einsprachig aufwachsende Kinder mit der Erstsprache Deutsch. Die Kinder besuchten Kindergärten im Vordertaunus (Hessen). Wie in der Querschnittstudie wurden in der Trainingsstudie Kinder aus Mittelschichtfamilien bzw. bildungsnahe Familien gewählt. Die Kinder benötigten ein gutes Sprachverständnis, um Aufgaben im komplexen Themenbereich des Schlussfolgerns bearbeiten zu können.

### 8.2.4 Design

Die Trainingsstudie folgte einem Prä-Post-Testdesign und ist als Laborstudie angelegt, um möglichst optimale Bedingungen zur Überprüfung der Wirksamkeit der Trainingsmaßnahmen zu schaffen. Die Probanden wurden auf Grundlage des Inhibitionstestes (Früchte-Gemüse-Stroop-Test) in drei gleichgroße Gruppen eingeteilt. Der Test wurde als Maß genutzt, da er sich in der Querschnittstudie bei der logistischen Regressionsanalyse als einflussreichster Prädiktor auf die Schlussfolgerungskompetenz (siehe Kapitel 7.2.5) zeigte. Es wurden immer drei Kinder mit ähnlichen Ergebnissen im Inhibitionstest jeweils einer der drei Gruppen (Kontrollgruppe, Trainingsgruppe 1, Trainingsgruppe 2) zugeordnet. Anschließend wurden die nächsten drei Kinder mit ähnlichen Kompetenzen zufällig jeweils einer Gruppe zugeteilt. Dieser Prozess wurde wiederholt, bis jede Gruppe 21 Kinder umfasst. Auf diese Weise wurden drei niveaugleiche Gruppen gebildet. Die zufällige Verteilung der Kinder auf die Gruppen wird auch als Randomisierung bezeichnet; daher handelte es sich bei der Trainingsstudie um eine experimentelle Studie.

In Tabelle 31 ist das Testdesign der Trainingsstudie dargestellt. Die Tests fanden an vier Tagen statt. Am ersten Tag wurden der Vorwissenstest, der Schlussfolgerungstest (Prä-Test) sowie der Inhibitionstest durchgeführt. Am zweiten Tag wurden der Wissenschaftsverständnistest sowie die erste Trainingssitzung, am dritten Tag die zweite Trainingssitzung sowie am vierten Tag der Schlussfolgerungstest (Post-Test), Transfer-Test und Argumentationstest realisiert.

Tabelle 30

*Design Trainingsstudie*

Alle Gruppen (KG, TG 1, TG 2)			
Tag 1	Tag 2	Tag 3	Tag 4
Vorwissenstest Schlussfolgerungstest, Prä-Test im Kontext <i>Elastizität und Plastizität</i> , Inhibitionstest	Wissenschaftsverständnisstest Trainings Sitzung I (TG 1, TG 2)	Trainings Sitzung II (TG 1, TG2)	Schlussfolgerungstest, Post-Test im Kontext <i>Elastizität und Plastizität</i> , Transfer-Test im Kontext <i>Schwimmen und Sinken</i> , Argumentationstest

8.2.5 *Analysestrategien bei der Auswertung*

Das Vorwissen der Kinder wurde in der Untersuchung zu Beginn des ersten Testtages ermittelt. Daher werden die Analysen zum Vorwissenstest zu Beginn aufgeführt. Die Antworten der Probanden reichten auf einer vierstufigen Skala von 1 bis 4; dabei wurde eine Vermutung mit 1 kodiert für *stimmt nicht*, mit 2 für *stimmt eher nicht*, mit 3 kodiert für *stimmt eher* und mit 4 kodiert für *stimmt*. Zunächst wurde eine Häufigkeitsverteilung für jede Vermutung über alle Gruppen hinweg vorgenommen, um zu schauen, ob eine konzeptuelle Vorstellung häufiger auftrat als eine andere. Im Anschluss daran wurde eine Reliabilitätsanalyse durchgeführt. Zudem wurde eine Korrelation gerechnet, um den Zusammenhang zwischen einer Vorannahme in der Vermutung „mit Luft“ und derselben Vermutung im Training zu prüfen.

Zur Beantwortung der ersten und zweiten Forschungsfrage wurden die Antworten der Probanden im Prä-Test und Post-Test miteinander verglichen. Die erste Forschungsfrage zielte darauf ab, zu prüfen, ob sich Unterschiede zwischen den Gruppen (Scaffolding durch adaptives Nachfragen/Unterstützung durch Modellierung und adaptives Nachfragen) im Post-Test nachweisen ließen.

Mit der zweiten Forschungsfrage wurde geprüft, inwiefern sich Mittelwertunterschiede zwischen den Trainingsgruppen (Scaffolding durch adaptives Nachfragen, Unterstützung durch Modellierung und adaptives Nachfragen) ergaben. Dazu wurden Antworten vom Prä-Test und Post-Test für jedes Ereignis  $[p \ q]$ ;  $[p \neg q]$ ;  $[\neg p \ q]$ ;  $[\neg p \neg q]$  in den beiden Gruppen mit Hilfe einer ANOVA mit Messwiederholung (mit zwei Messzeitpunkten: Prä-Test, Post-Test) miteinander verglichen.

Als Kovariaten berücksichtigt wurden dabei der Inhibitionstest sowie der Wissenschaftsverständnistest, da sich beide Tests als einflussreiche Prädiktoren in der Querschnittstudie erwiesen hatten (vgl. Kapitel 7.1.6). Für den Wissenschaftsverständnistest wurde eine Raschanalyse durchgeführt, um detailliert zeigen zu können, welche Kompetenzen die Kinder in den einzelnen Aufgabenbereichen (Produktion von Evidenz, Erkennen von Evidenz, Produktion von Theorie, Interpretation von Theorie) in den drei Kontexten erreichten. Das Vorwissen wurde bei den Analysen in der Trainingsstudie nicht als Kovariate verwendet, da in der Trainingsstudie kein Zusammenhang zwischen Vorwissen und Schlussfolgerungskompetenzen nachgewiesen werden konnte.

Zur Untersuchung der dritten Frage, wie effektiv sich trainierte Schlussfolgerungskompetenzen im Anschluss an ein intensives Training im Transfer-Kontext darstellten, wurde eine ANOVA mit Messwiederholung (zwei Messzeitpunkte, 2. Trainingssitzung, Transfer-Test) für beide Trainingsgruppen vorgenommen. Hierbei galt es, herauszufinden, ob sich das Antwortverhalten der Probanden vom Training zum Transfer-Test hin deutlich veränderte. Zur Überprüfung der Wirksamkeit der beiden Trainingsmaßnahmen wurde untersucht, ob die Antworten in den beiden Trainingsgruppen sich im Transfer-Test signifikant voneinander unterschieden. Dazu wurde ein T-Test für unabhängige Stichproben durchgeführt. Zusätzlich wurde mittels T-Tests für unabhängige Stichproben geprüft, ob sich die Anzahl an Hilfen in beiden Gruppen deutlich voneinander unterschied.

### 8.3 Ergebnisse der Hauptstudie

Im Vorfeld der Analysen zu den Forschungsfragen der Trainingsstudie werden die Ergebnisse zum Vorwissenstest dargestellt, da der Test zu Beginn der Testphase am ersten Tag durchgeführt worden ist. Zunächst wurden die Häufigkeiten für jedes Konzept (leicht, rund, mit Luft) im Zusammenhang mit dem Sprungverhalten von Gegenständen betrachtet. Die konzeptuelle Vorstellung „leicht und springt“ wurde von 54% der Kinder in der Kategorie „stimmt“ gewählt, die Vorstellung „rund und springt“ von 50% der Kinder, die konzeptuelle Vorstellung „mit Luft und springt“ von 34% der Kinder gewählt. Im Anschluss daran wurde ein Summenwert gebildet (siehe Tabelle 32), es wurden Gruppenunterschiede geprüft. In der Berechnung mit ANOVA fungierte der Summenwert als abhängige Variable, dabei zeigten sich mit  $F=(2, 60)= 0.83 ns$ ;  $\epsilon^2 = 0.03$  keine signifikanten Unterschiede zwischen den Gruppen (siehe Tabelle 32). In der Reliabilitätsanalyse ergab sich mit  $\alpha = 0.13$  kein zufriedenstellender Wert für Cronbachs Alpha. Das bedeutete, dass der Test nicht zuverlässig das Vorwissen der Kinder erfasste.

Tabelle 31

*Summenwerte zum Vorwissen in den drei Gruppen*

Summenwert	Trainingsgruppe 1	Trainingsgruppe 2	Kontrollgruppe
0	0%	0%	0%
1	0%	0%	0%
2	0%	0%	0%
3	9.5%	14.3%	9.5%
4	9.5%	19.0%	14.3%
5	23.8%	9.5%	19.0%
6	19.0%	9.5%	23.8%
7	4.8%	14.3%	23.8%
8	14.3%	4.8%	9.5%
9	9.5%	23.8%	0%
10	4.8%	0%	0%
11	4.8%	4.8%	0%
12	0%	0%	0%

*Anmerkung.* ANOVA  $F = (2, 60) = 0.83$  ns;  $\epsilon^2 = 0.03$

6+Weiterhin wurde eine Korrelation zwischen einer Vorannahme des Vorwissenstests mit der Vermutung „Dinge, die mit Luft gefüllt sind, springen“ und derselben Vermutung, mit denen Kindern im Training gefördert werden, durchgeführt. Dazu wurden die Rohwerte des Vorwissenstests der Vermutung *Luft* genutzt. Es ergab sich keine signifikante Korrelation für ein Ereignis ([p q best]; [p¬q wid]; [¬p q irrel] oder [¬p¬q irrel]). Daher war anzunehmen, dass Kinder ihre Kompetenzen beim Schlussfolgern im trainierten Kontext „Luft“ (Trainings-sitzung 1) unabhängig von ihren Vorannahmen tätigten.



### *8.3.1 Ergebnisse zur ersten Forschungsfrage: Förderung der Schlussfolgerungsfähigkeit*

Um den ersten Teil der ersten Forschungsfrage (Forschungsfrage 1a) nachzugehen, ob sich die Schlussfolgerungsfähigkeit von Kindern fördern lässt, werden zunächst die Ergebnisse des Schlussfolgerungstests vom Prä-Test sowie vom Post-Test für die Einzelereignisse (=abhängige Variable) in Tabelle 33 und Tabelle 34 anhand der relativen Häufigkeiten aufgeführt. Hervorgehoben in blau sind die korrekten Evaluationen im Sinne konditionalen Schlussfolgerns. Dabei zeigte sich, dass Kinder der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) und der Kontrollgruppe keinen deutlichen Zuwachs vom Prä- zum Post-Test erreichten.

Tabelle 32

*Relative Häufigkeiten der Einzelereignisse im Prä-Test*

Ereignis	Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen)			Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen)			Kontrollgruppe		
	best	irr	wid	best	irr	wid	best	irr	wid
[p q]	97%	0%	3%	87%	10%	3%	81%	6%	13%
[p¬q]	2%	13%	86%	10%	11%	79%	17%	8%	75%
[¬p q]	56%	13%	32%	41%	21%	38%	38%	16%	46%
[¬p¬q]	5%	17%	78%	17%	14%	68%	20%	17%	63%

*Anmerkung.* Nicht-Addition zu 100 (pro Zeilenzelle) aufgrund von Rundung. best = bestätigend; irr = irrelevant; wid = widerlegend.

Bei den Ereignissen mit [¬p], die irrelevant im Zusammenhang mit einer Vermutung waren, zeigten Kinder der Trainingsgruppe 2 signifikant bessere Antworten als Kinder der beiden anderen Gruppen (Trainingsgruppe 1, Kontrollgruppe). Kinder der Trainingsgruppe 2 (Scaffolding durch Modellierung und adaptives Nachfragen) gaben bei den Ereignissen [¬p q irrel] und [¬p¬q irrel] im Post-Test zu jeweils 38% die richtige Antwort. Im Vergleich dazu erreichten Kinder der Trainingsgruppe 1 nur 27% bei [¬p q irrel]; 17% bei [¬p¬q irrel]; Kinder der Kontrollgruppe erreichten 24% beim Ereignis [¬p q irrel] und 19% bei [¬p¬q irrel].

Zur Beantwortung der Forschungsfrage 1b wurden auf Grundlage der Einzelereignisse Interpretationsniveaus zum Schlussfolgern (Barrouillet et al., 2008; Tröbst et al., 2011) gebildet. Für jedes Interpretationsniveau sind die relativen Häufigkeiten im Gruppenvergleich (Prä-Test, Post-Test) sind sowohl in Tabelle 35 als auch in Abbildung 22 dargestellt.

Kinder beider Trainingsgruppen antworteten im Post-Test häufiger auf dem konditionalen Niveau als im Prä-Test, Kinder der Kontrollgruppe antworteten nicht häufiger auf diesem Niveau. Bei Trainingsgruppe 1 stieg der Anteil von 0% auf 9%, bei Trainingsgruppe 2 von 9% auf 43%. Bei der Kontrollgruppe blieb der Wert mit 9% im Prä- und Post-Test unverändert.

Auf dem konjunktiven Interpretationsniveau zeigte sich in beiden Trainingsgruppen eine deutliche Abnahme an Antworten zwischen Prä- und Post-Test. Bei Trainingsgruppe 1 reduzierten sich die Antworten von 52% auf 43%, bei Trainingsgruppe 2 verringerten sich die Antworten von 28% auf 9%. bei der Kontrollgruppe blieb der prozentuale Anteil an Antworten bei 33%. Die Antworten in der Kategorie „Sonstige“ nahmen sowohl in der Trainingsgruppe 1 von 5% auf 14% zu; in der Trainingsgruppe 2 blieben sie gleich. In der Kontrollgruppe erhöhten sie sich leicht von 24% auf 29%.

Beim konjunktiven Interpretationsniveau zeigte sich bei der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) eine deutliche Abnahme von 52% auf 43% an Antworten zwischen Prä- und Post-Test, bei der Kontrollgruppe erhöhten sich die Antworten auf diesem Niveau nicht. Die Antworten auf dem Niveau „Sonstige“ blieben in der Trainingsgruppe 2 mit 19% gleich, in der Trainingsgruppe 1 stiegen die Antworten von 5% auf 14% an. In der Kontrollgruppe nahmen die Antworten hier auch von 24% im Prä-Test auf 29% im Post-Test zu. In einer multifaktoriellen Analyse zeigte sich ein signifikantes Ergebnis. MANOVA  $F(1, 60) = 3.85$ ;  $p = .05$ ;  $\epsilon^2 = 0.6$

Tabelle 33

Relative Häufigkeiten der Einzelereignisse im Post-Test

Ereignis	Trainingsgruppe 1			Trainingsgruppe 2			Kontrollgruppe		
	best	irr	wid	best	irr	wid	best	irr	wid
$p \ q$	95%	0%	5%	95%	5%	0%	95%	3%	2%
$p \neg q$	10%	22%	67%	28%	28%	50%	8%	14%	76%
$\neg p \ q$	22%	26%	54%	33%	47%	19%	23%	33%	38%
$\neg p \neg q$	5%	10%	85%	10%	43%	47%	34%	10%	48%

*Anmerkung.* Nicht-Addition zu 100 (pro Zeilenzelle) aufgrund von Rundung.  
best = bestätigend; irr = irrelevant; wid = widerlegend.

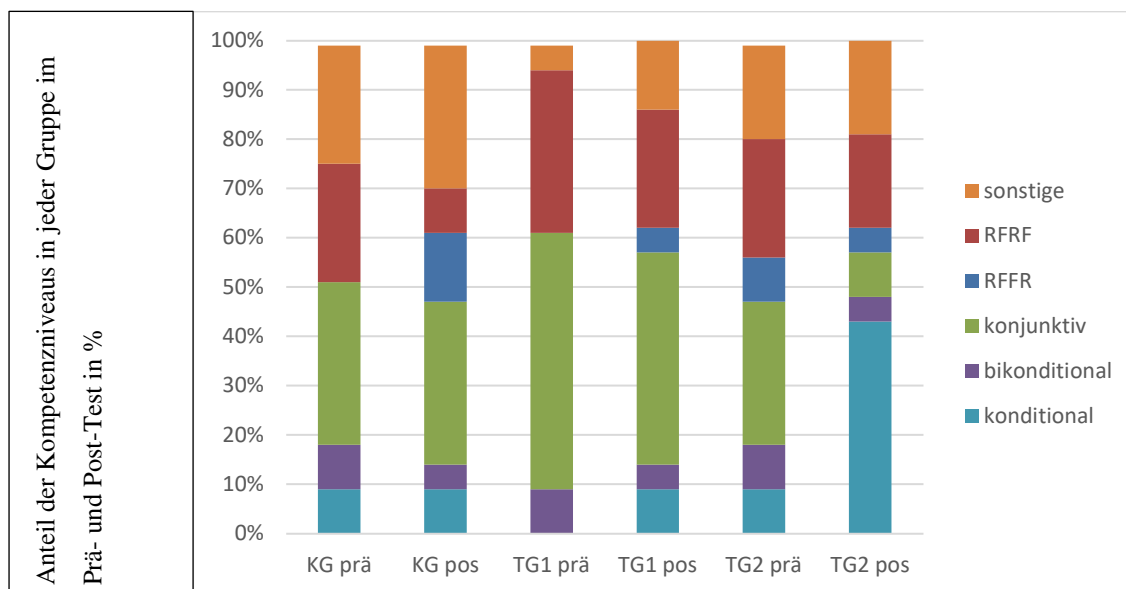
Tabelle 34

Relative Häufigkeiten Interpretationsniveaus im Prä-Post-Vergleich

Interpretationsniveaus	Trainingsgruppe 1		Trainingsgruppe 2		Kontrollgruppe	
	Prä-Test	Post-Test	Prä-Test	Post-Test	Prä-Test	Post-Test
konditional	0%	9%	9%	43%	9%	9%
bikonditional	9%	5%	9%	5%	9%	5%
konjunktiv	52%	43%	28%	9%	33%	33%
RFFR	0%	5%	9%	5%	0%	14%
RFRF	33%	24%	24%	19%	24%	9%
sonstige	5%	14%	19%	19%	24%	29%

Anmerkung: Nicht-Addition zu 100 (pro Spalte) aufgrund von Rundung.

MANOVA  $F(1, 60) = 3.85$ ;  $p = .05$ ;  $\epsilon^2 = 0.6$



Anmerkung. MANOVA  $F(1, 60) = 3.85$ ;  $p = .05$ ;  $\epsilon^2 = 0.6$

Abbildung 22. Relative Häufigkeiten im Prä- und Post-Test für die drei Gruppen

Der Wissenschaftsverständnistest wurde in die Analysen einbezogen, da die Fähigkeiten bei der Theorie-Evidenz-Koordination einen signifikanten Einfluss auf das Schlussfolgern nahmen (siehe Kapitel 7.1.6).

Im folgenden Abschnitt werden daher die Analysen zum Wissenschaftsverständnistest im Rahmen der Beantwortung der ersten Forschungsfrage dargestellt. In Tabelle 35 werden die Mittelwerte, sowie in Tabelle 37 die Item-Kennwerte des Wissenschaftsverständnistests, die durch eine eindimensionale Raschanalyse berechnet wurden, aufgeführt. Beim Betrachten der Mittelwerte (Tabelle 35) wurde deutlich, dass das Erkennen von Evidenz mit einem Wert von 1.9 im Vergleich zur Interpretation deutlich niedriger als der Mittelwert bei Interpretation von Theorie mit 2.6 lag. Wie bereits in der Querschnittstudie, zeigte sich auch in dieser Studie, dass den fünf- bis sechsjährigen Kindern die Aufgaben zur Produktion von Evidenz deutlich schwerer fielen als die Aufgaben zur Interpretation von Theorie. Insgesamt zeigten sich ähnliche Ergebnisse für diese Altersstufe wie in der Querschnittstudie (vgl. Kapitel 7.1.6).

Tabelle 35

*Mittelwerte (Standardabweichungen) Wissenschaftsverständnis*

Rubriken Wissenschaftsverständnis	Mittelwerte (Standardabweichungen)
Produktion von Evidenz	1.2 (0.9)
Erkennen von Evidenz	1.9 (1.0)
Produktion von Theorie	2.4 (0.8)
Interpretation von Theorie	2.6 (0.6)

In der nachfolgenden Tabelle (Tabelle 36) sind die Item-Kennwerte, die mit eindimensionaler Rasch-Skalierung genau wie in der Querschnittstudie durchgeführt wurden, dargestellt. Dabei wird das schwierigste Item als erstes dargestellt.

Tabelle 36

*Itemkennwerte des Wissenschaftsverständnisses*

Items	Aufgabenty p	N	Lösungs- häufigkei t (%)	Schwie rigkeit	Trenn schärf e	Infit MNS Q	T- Wert
1_1 Produktion von Evidenz Marius	Produktion von Evidenz	63	19.05	3.03	0.63	1.09	1.63
3_2 Erkennen von Evidenz Struppi	Erkennen von Evidenz	63	36.51	1.93	0.57	0.93	-0.60
3_1 Produktion von Evidenz Struppi	Produktion von Evidenz	63	41.27	1.67	0.62	0.93	1.23
2_1 Produktion von Evidenz Elena	Produktion von Evidenz	63	57.14	0.85	0.57	0.98	1.25
2_3 Produktion von Theorie Theorie Elena	Produktion von Theorie	63	58.73	0.76	0.49	1.03	1.26
2_2 Erkennen von Evidenz Elena	Erkennen von Evidenz	63	61.90	0.60	0.74	0.80	1.27
2_4 Interpretation von Theorie Elena	Interpretation von Theorie	63	73.02	-0.05	0.35	1.2	1.32

Items	Aufgabentyp	N	Lösungshäufigkeit (%)	Schwierigkeit	Trennschärfe	Infit MNS Q	T-Wert
1_3 Produktion Theorie Marius	Produktion von Theorie	63	85.71	-1.03	0.62	0.82	1.47
1_4 Interpretation Theorie Marius	Interpretation von Theorie	63	85.71	-1.03	0.39	1.06	1.47
1_2 Erkennen Evidenz Marius	Erkennen von Evidenz	63	88.89	-1.36	0.52	0.88	1.55
3_3 Produktion Theorie Struppi	Produktion von Theorie	63	92.06	-1.78	0.15	1.17	1.69
3_4 Interpretation Theorie Struppi	Interpretation von Theorie	63	98.41	-3.58	0.15	1.03	1.17

Anmerkung. Skala EAP/PV Reliabilität (expected a priori Reliabilität) = 0.79; Varianz: 3.39; Schwierigkeit: M = 0.0; SD= 1.84; MIN =-3.58; MAX = 1.93; Trennschärfe: M = 0.48, SD =0.19; MIN= 0.15 ; MAX= 0.74; Infit: M= 0.90; SD = 0.19; MIN = 0.8 ; MAX= 1.17.

Die Werte (Mean Weighted Squares) der zwölf Items für die 63 Probanden lagen zwischen 0.80 und 1.17 (M = 0.90; SD = 0.19). Nach Wilson (2005) liegen Werte für den MNSQ zwischen 0.75 und 1.33 in einem akzeptablen Bereich. Der Wert der Personenfähigkeit *Weighted Likelihood Estimates* (WLE) ist ein Schätzwert für die Reliabilität (Warm et al., 1989). Die innere Konsistenz war mit einer *EAP/PV*-Reliabilität, die ähnlich der Reliabilität von Cronbachs Alpha ist, mit einem Wert von 0.79 zufriedenstellend (siehe Tabelle 37).

Mit der *Wright Map* (siehe Abbildung 23) wurden Personenfähigkeiten sowie Itemschwierigkeiten auf einer Skala, die die Modellparameter abbildeten, dargestellt. Dabei sind auf der linken Seite die Personenfähigkeiten auf einer Skala (von -3 bis +5) und auf der rechten Seite die zwölf Items dargestellt.

Die mittlere Itemschwierigkeit wurde auf 0 Logits mit einer Standardabweichung (SD = 1.84) restringiert und bildet den Nullpunkt der Skala. Die Trennschärfe lag im Mittel bei 0.48 (SD = 0.19). An der *Wright Map* kann man erkennen, dass genügend Aufgaben mit mittlerer Schwierigkeit vorhanden waren. Sie zeigt auch, dass die Hälfte der Aufgaben (wie z.B. Item 12, Item 3) wenig Anforderungen (auf der Skala bei 0 bis -2,2) an die Probanden stellten und von fast allen Probanden (Item 12 mit 98%, Item 3 mit 92%) gelöst werden konnten.



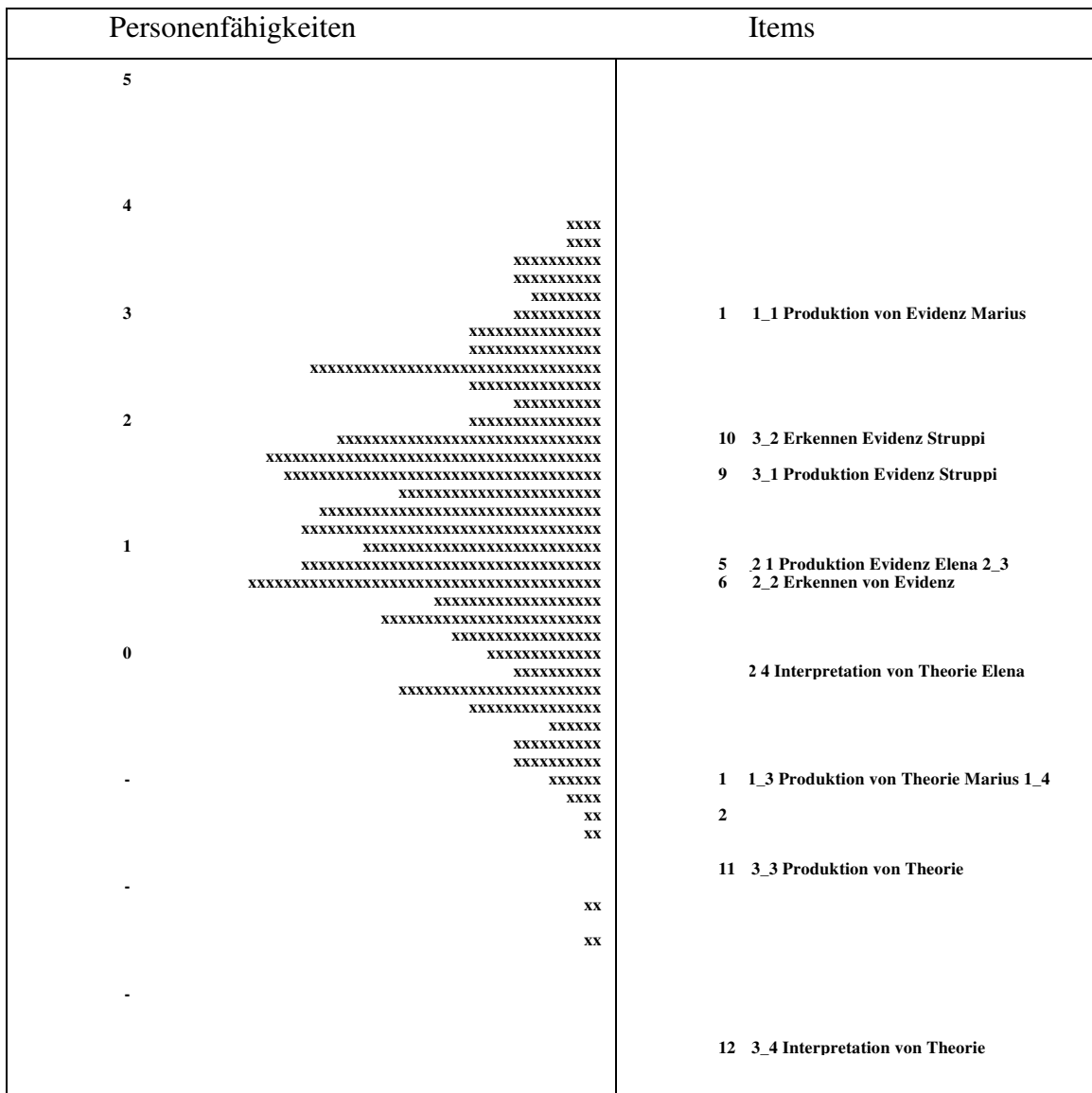


Abbildung 23: Wright Map in Trainingsstudie

### 8.3.2 Ergebnisse zur zweiten Forschungsfrage: Unterschiede zwischen den Trainingsmaßnahmen

Zur Beantwortung der zweiten Forschungsfrage wurden die beiden Trainingsgruppen (Scaffolding durch adaptives Nachfragen, Unterstützung durch Modellierung und adaptives Nachfragen) miteinander verglichen. Dabei zeigte sich ein deutlicher Unterschied zwischen den Gruppen. Probanden der Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) konnten ihre Antworten auf konditionalem Interpretationsniveau vom Prä-Test (5%) im Post-Test deutlich steigern (auf 30%). Im Vergleich dazu verbesserten Kinder der

Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) ihre Antworten vom Prä-Test hin zum Post-Test kaum.

Tabelle 37

*Mittelwerte für die Einzelereignisse (Prä-Test und Post-Test)*

Ereignis	Erläuterung	Mittelwerte Trainingsgrup pe 1	Mittelwerte Trainings- gruppe 2	Mittelwerte Kontrollgruppe
[pq best]	Ereignis bestätigt Vermutung	.24 (.57)	.23(.30)	.47 (-0.65)
[p¬q wid]	Ereignis wi- derlegt Vermutung	.13 (0.13)	.014 (.13)	-.05 (1.2)
[¬pq irrel]	Ereignis ist irrelevant für Vermutung	.05 (0.14)	.81 (0.30)	.38 (0.37)
[¬p¬q irrel]	Ereignis ist irrelevant für Vermutung	-.05 (0.63)	.86 (.00)	.05 (0.06)

\*=signifikantes Ergebnis; *MANOVA*  $F(4,57) = 3.37$   $p = .015$ ;  $\epsilon^2 = .19$

In Tabelle 38 sind die mit einer univariaten Varianzanalyse berechneten Mittelwertunterschiede als Differenzwerte (Post-Test und Prä-Test) für die drei Gruppen (Trainingsgruppe 1, Trainingsgruppe 2, Kontrollgruppe) aufgeführt. Dabei war jeweils der Differenzwert (Post-Test, Prä-Test) eines Ereignisses Grundlage für die Berechnung.

In der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) hatten sich die Mittelwerte bei den Ereignissen mit [¬p] im Vergleich vom Prä- hin zum Post-Test nicht erhöht bzw. leicht verringert. In der Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) zeigte sich beim Ereignis [¬p¬q irrel], das irrelevant im Zusammenhang mit der Vermutung war, ein statistisch signifikanter Zuwachs. Im Vergleich dazu stellte sich in der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) beim Ereignis [¬p¬q irrel] kein signifikanter Zuwachs heraus. Tendenziell konnte für das Ereignis [¬p q irrel] ein ähnlicher Unterschied festgestellt werden, dieser war jedoch statistisch nicht signifikant (siehe Tabelle 38). Bei den Berechnungen wurden die individuellen Werte der Personenfähigkeit (WLE) aus der eindimensionalen Raschanalyse sowie der Zeitfaktor beim Inhibitionstest als Kovariaten berücksichtigt, da sie sich in der Querschnittstudie als signifikante Prädiktoren für die Schlussfolgerungskompetenz herausgestellt hatten (siehe Kapitel 7.2.5).

Tabelle 38

*Differenzwerte der ANOVA aus Prä- und Post-Test*

Ereignis	F-Werte
[p q]	$F(4, 58) = 0.4; p < 0.5; \varepsilon^2 = 0.21$
[p ¬q]	$F(4, 58) = 0.05; p = \text{ns.}; \varepsilon^2 = 0.0$
[¬p q]	$F(4, 58) = 1.7; p = \text{ns.}; \varepsilon^2 = 0.1$
[¬p ¬q]	$F(4, 58) = 2.5; p = 0.05; \varepsilon^2 = 0.15$

Die Probanden der Trainingsgruppe 2 (Scaffolding mit Modellierung und adaptivem Nachfragen) können im Vergleich zur Trainingsgruppe 1 (Scaffolding mit adaptivem Nachfragen) beim Ereignis [¬p ¬q irrel] ihre Antworten steigern. Insgesamt ergibt sich bei der Berechnung mit MANOVA mit  $F(4,57) = .3.37$   $p = .015$ ;  $\varepsilon^2 = .19$  ein signifikanter Unterschied zwischen den Gruppen bei den Einzelereignissen. Dabei werden jeweils die F-Werte der Pillai-Spur gewählt, da nach Olson (1976) bei geringen Korrelationen der abhängigen Variablen sich diese Teststatistik am besten eignet. Hier liegen geringe nicht signifikante Korrelationen zwischen den einzelnen Ereignissen vor.

In Tabelle 38 sind die Differenzwerte der ANOVA aus Prä- und Post-Test für die vier Ereignisse [p q best]; [p¬q wid]; [¬p q irrel]; [¬p¬q irrel] einzeln aufgeführt. Beim Ereignis [p q best] mit  $F(4, 58) = 0.4; p < 0.05$  ergab sich ein signifikanter Wert mit einer Effektstärke von  $\varepsilon^2 = 0.02$ . Bei den beiden anderen Ereignissen [p¬q wid]; [¬p q irrel] zeigten sich keine signifikanten Werte (siehe Tabelle 39). Beim Ereignis [¬p¬q irrel] stellte sich wiederum ein signifikanter Unterschied zwischen Prä-Test und Post-Test heraus.

In Tabelle 40 sind die Differenzwerte der ANOVA aus Prä- und Post-Test für die vier Ereignisse [p q best]; [p¬q wid]; [¬p q irrel]; [¬p¬q irrel] einzeln aufgeführt. Beim Ereignis [p q best] mit  $F(4, 58) = 0.4; p < 0.05$  ergab sich ein signifikanter Wert mit einer Effektstärke von  $\varepsilon^2 = 0.02$ . Bei den drei anderen Ereignissen [p¬q wid]; [¬p q irrel] und [¬p ¬q irrel] zeigten sich keine signifikanten Werte (siehe Tabelle 41).

Die Kovariaten, die Personenfähigkeit (WLE) aus der eindimensionalen Raschanalyse ebenso wie der Inhibitionstest (Früchte-Gemüse-Stroop) sind in die Berechnungen einbezogen worden. Die Kovariaten (Inhibition, Wissenschaftsverständnis) unterschieden sich in der Berechnung mit *ANOVA* in den drei Gruppen (Trainingsgruppe 1, Trainingsgruppe 2, Kontrollgruppe) statistisch nicht (siehe Tabelle 39).

Tabelle

39

*ANOVA-Analysen der Kovariaten*

	Trainingsgruppe 1	Trainingsgruppe 2	Kontrollgruppe
Inhibitionstest (Zeit in Sekunden)	74.3 (26.6)	78.6 (22.4)	75.9 (28.3)
Inhibitionstest (ANOVA)	$F(2, 99) = 0.15; p = \text{ns.}; \varepsilon^2 = 0.005$		
Wissenschaftsverständnis	8.38(2,29)	7.42(2.50)	8.14(2.24)
ändnistest (ANOVA)	$F(2, 2) = 0.85; p = \text{ns.}; \varepsilon^2 = 0.03$		

Die Probanden der Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) konnten im Vergleich zur Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) beim Ereignis [ $\neg p \neg q$  irrel] ihre Antworten signifikant steigern. Insgesamt ergab sich bei der Berechnung mit *MANOVA* mit  $F(8, 112) = 3.64; p < 0.05; \varepsilon^2 = 0.02$  ein signifikanter Unterschied zwischen den Gruppen. Dabei wurden jeweils die F-Werte des Pillai-Spur-Tests gewählt.

In der zweiten Forschungsfrage wurden die Unterschiede beider Trainingsmaßnahmen herausgearbeitet. In diesem Zusammenhang wurde die Prüfung der Wirksamkeit der beiden Trainingsmaßnahmen (Scaffolding durch adaptives Nachfragen, Unterstützung durch Modellierung und adaptives Nachfragen) anhand der Anzahl der notwendigen Unterstützungen in beiden Trainingsgruppen miteinander verglichen. Dazu wurden die Mittelwerte und Standardabweichungen in beiden Trainingsgruppen berechnet (siehe Tabelle 40). Es zeigte sich deskriptiv, dass in Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) der Mittelwert

mit 1.62 deutlich höher lag als mit 1.09 in Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen).

Daher wurde die Anzahl der eingesetzten Unterstützungsmaßnahmen in der zweiten Sitzung mithilfe von T-Tests für unabhängige Stichproben in beiden Gruppen verglichen. Es zeigte sich, dass die Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) signifikant häufiger durch adaptives Nachfragen unterstützt werden musste als die Trainingsgruppe 2 (Scaffolding durch Modellierung und adaptives Nachfragen). Der T-Test ergab einen Wert von  $T(40) = 2.29$ ;  $p < 0.05$  (siehe Tabelle 41).

Tabelle 40

*Mittelwerte (Standardabweichungen) der Trainingsmaßnahmen*

Trainingsgruppe 1	Trainingsgruppe 2
1.62 (0.75)	1.09 (0.89)
$T(40) = 2.29$ ; $p < 0.05$	

Durch die Analysen (Prä-Post-Test-Vergleich der Einzelereignisse in beiden Trainingsgruppen, Vergleich der Unterstützungsmaßnahmen durch T-Test) zeigten sich signifikante Gruppenunterschiede zwischen den beiden Trainingsgruppen.

*8.3.3 Ergebnisse zur dritten Forschungsfrage:  
Langfristige Trainingseffekte*

Die dritte Forschungsfrage untersuchte, ob sich nachhaltige Trainingseffekte in den drei Experimentalgruppen über die Zeit hinweg zeigen lassen. Daher wurde geprüft, inwiefern sich die beiden Trainingsgruppen im Vergleich zur Kontrollgruppe über die verschiedenen Messzeitpunkte (Prä-Test, Trainingssitzung I, Trainingssitzung II, Post-Test, Transfer-Test) unterschieden. Erwartet wurde, dass die Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) sowohl der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) als auch der Kontrollgruppe über die Zeit hinweg deutlich überlegen war. Dazu wurden die Gruppen im Zusammenhang mit den erwartungsgemäßen Antworten der Kinder aller drei Gruppen (Trainingsgruppe 1 und 2, Kontrollgruppe) in Bezug auf die vier Ereignisse [p q best]; [p¬q wid]; [¬p q: irrel]; [¬p¬q irrel] über fünf Messzeitpunkte (Prä-Test,

1. Trainings-sitzung, 2. Trainings-sitzung, Post-Test sowie Transfer-Test) dargestellt. Die Antworten im Prä- sowie im Post-Test wurden in beiden Trainingsgruppen weder durch adaptives Nachfragen noch durch Modellierung unterstützt. Die Antworten in den beiden Trainingsphasen wurden in der Trainingsgruppe 1 durch Scaffolding mit adaptivem Nachfragen unterstützt (siehe Tabelle 43 bis Tabelle 46). Bei der Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) legte der Versuchsleiter in der ersten Trainings-sitzung während der Durchführung der vier Ereignisse seinen Gedankengang vollständig offen. Das bedeutete, dass er den Probanden den richtigen Weg bis hin zur Evaluation zeigte und dabei die Lösungen nannte (siehe Tabelle 42 bis Tabelle 45). Bei der zweiten Trainings-sitzung wurde die Modellierung nur beim Ereignis der Form  $[p \rightarrow q \text{ best}]$  durchgeführt. Danach wurde auch in Trainingsgruppe 2 nur durch adaptives Nachfragen unterstützt (diese Werte sind ebenfalls lila gekennzeichnet). Daher waren beide Trainingsgruppen in Bezug auf drei Ereignisse ( $[p \rightarrow \neg q \text{ wid}]$ ;  $[\neg p \rightarrow q \text{ irrel}]$ ;  $[\neg p \rightarrow \neg q \text{ irrel}]$ ) durch identische Form der Hilfe (Scaffolding durch adaptives Nachfragen) gekennzeichnet und konnten miteinander verglichen werden. Im Transfer-Test wurden Kinder in beiden Trainingsgruppen jeweils mit Scaffolding durch adaptives Nachfragen unterstützt. Die Kontrollgruppe erhielt im Transfer-Test keinerlei Unterstützung.

In Tabelle 41 bis Tabelle 44 sind die Analysen über die fünf Messzeitpunkte mit ihren Prozentwerten und direkt im Anschluss jeweils die graphische Darstellung (Abbildung 24 bis Abbildung 27) aufgeführt. In jeder Tabelle sind die Einträge der Kontrollgruppe bei der ersten und zweiten Trainings-sitzung mit einem Strich gekennzeichnet, da diese Gruppe keine Unterstützung erhält.

Das erste Ereignis  $[p \rightarrow q \text{ best}]$  war am leichtesten zu bewerten, da hier die Hypothese bestätigt wurde (siehe Tabelle 41). Dieses Ereignis konnten Kinder aller drei Gruppen bereits häufig im Prä-Test richtig beantworten (97%, 81%, 67%).

In der Trainingsgruppe 1 gab es einen Abfall vom Prä-Test (97%) hin zum Post-Test (67%), während bei der Trainingsgruppe 2 ein leichter Anstieg an richtigen Antworten (von 81% im Prä-Test auf 86% im Post-Test) festzustellen war. Im Transfer-Test konnten die Kinder beider Trainingsgruppen bessere Lösungs-raten (100%, 95%) erzielen als im Post-Test. In der Kontrollgruppe ergaben sich 86% richtige Antworten im Post-Test, aber nur 52% im Transfer-Test.

Tabelle 41

*Prozessanalyse Ereignis [p q best]*

	Prä-Test	Training I	Training II	Post-Test	Transfer-Test
TG 1	95%	95%	100%	67%	100%
TG 2	81%	100%	100%	86%	95%
KG	67%	—	—	86%	52%

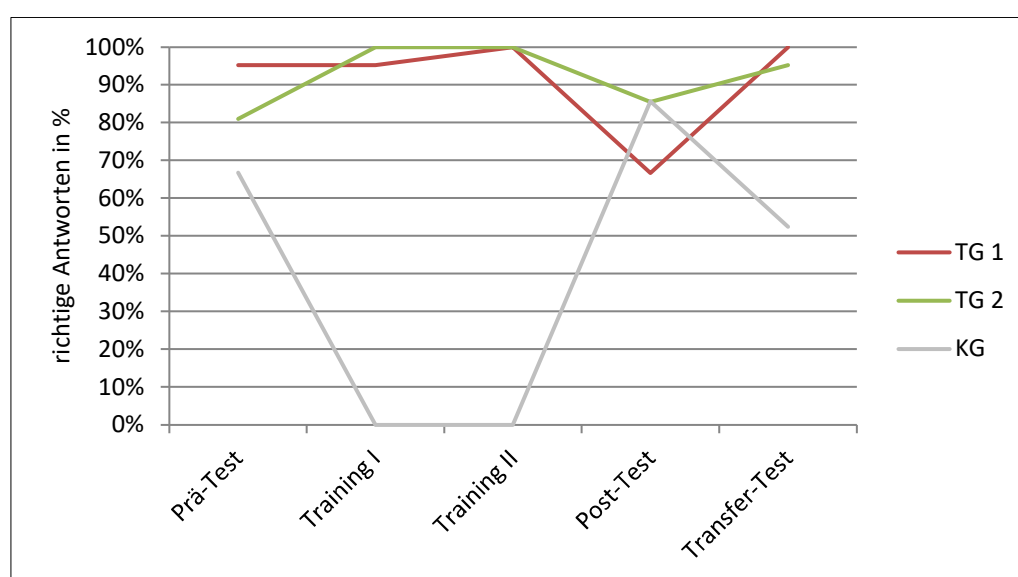


Abbildung 24. Prozessanalyse für das Ereignis [p q best].

Ein ähnlicher Verlauf vom Prä-Test über den Post-Test hin zum Transfer-Test zeigte sich beim Ereignis [p-q wid], mit dem eine genannte Hypothese widerlegt wurde (siehe Tabelle 42). Dabei stellte sich heraus, dass die Anzahl richtiger Antworten zu den drei Zeitpunkten (Prä-Test, Post-Test, Transfer-Test) deutlich niedriger ausfielen als beim ersten Ereignis der Form [p q best]. In Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) war ein leichter Abfall vom Prä-Test (71%) hin zum Post-Test (67%) festzustellen, in der Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) war ebenso eine Abnahme richtiger Antworten vom Prä-Test (62%) hin zum Post-Test zu erkennen (52%). Im Transfer-Test nahmen die Lösungsraten wieder deutlich zu und stiegen in beiden Trainingsgruppen jeweils auf 86% an. Das heißt, dass Kinder der Trainingsgruppe 1 im trainierten Kontext, Training II sowie im Transfer-Test nahezu genauso gut antworten konnten

wie Kinder der Trainingsgruppe 2. In der Kontrollgruppe fielen die richtigen Antworten der Kinder im Post-Test auf 48% ab.

Tabelle 42

*Prozessanalyse für das Ereignis [p¬q wid]*

	Prä-Test	Training I	Training II	Post-Test	Transfer-Test
TG 1	71%	95%	95%	67%	86%
TG 2	62%	100%	100%	52%	86%
KG	62%	—	—	57%	48%

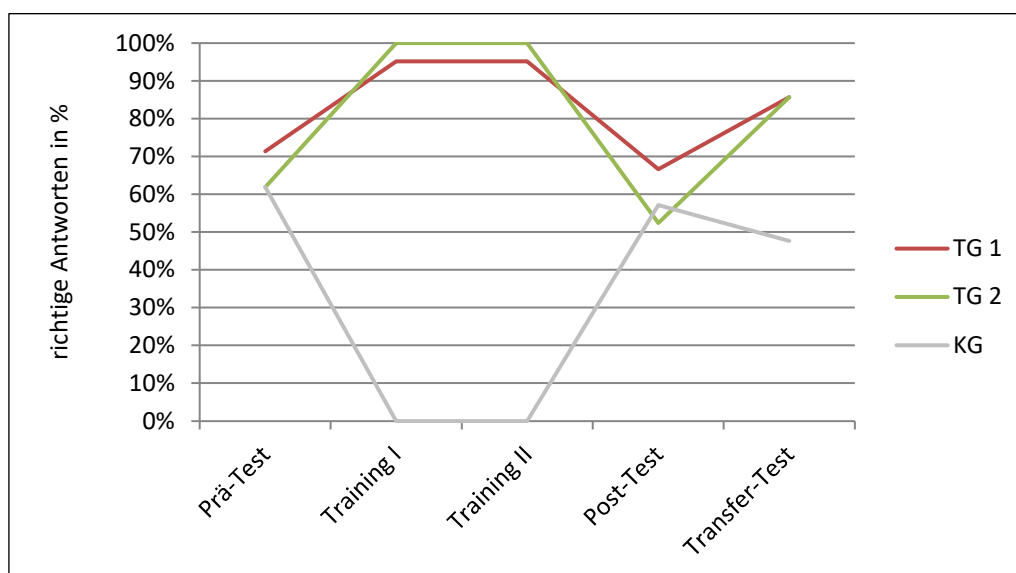


Abbildung 25. Prozessanalyse für das Ereignis [p¬q wid].

Betrachtet man die beiden Ereignisse [¬p q irrel]; [¬p¬q irrel], die in Tabelle 43 und Tabelle 44 dargestellt sind, so zeigte sich in der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) beim irrelevanten Ereignis [¬p q irrel] ein Abfall von 10% im Prä-Test auf 5% im Post-Test. Im Vergleich dazu nahmen die richtigen Antworten in der Trainingsgruppe 2 (Unterstützung durch Modellierung und durch adaptives Nachfragen) deutlich zu. Hier stiegen die Antworten von 5% auf 29%. Die Kinder der Kontrollgruppe erreichten beim Ereignis [¬p q irrel] mit 5% im Post-Test eine gleich hohe Lösungsrate wie Kinder der Trainingsgruppe 1 im Post-Test. Demzufolge zeigte sich in Trainingsgruppe 1 sowie in der Kontrollgruppe kein Lernzuwachs.



Vergleicht man die Ergebnisse des Transfer-Tests in beiden Trainingsgruppen, so wichen die Lösungsraten beider Gruppen beim Ereignis der Form  $[\neg p \ q \ \text{irrel}]$  deutlich voneinander ab. In der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) konnten 52% der Kinder und in der Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) 95% der Kinder richtig antworten. Die Lösungsrate im Transfer-Test lag bei Kindern der Kontrollgruppe mit 5% deutlich niedriger als die beider Trainingsgruppen (siehe Tabelle 43). Vergleicht man die Entwicklung der Lösungsraten für das Ereignis  $[\neg p \ q \ \text{irrel}]$  innerhalb einer Gruppe vom Post-Test zum Transfer-Test, so zeigte sich ein deutlicher Zuwachs. Kinder der Trainingsgruppe 1 erhöhten ihre richtigen Antworten nahezu um das Zehnfache vom Post-Test mit 5% hin zum Transfer-Test mit 52%. In der Trainingsgruppe 2 nahmen die richtigen Antworten von 29% im Post-Test auf 95% im Transfer-Test stark zu. In der Kontrollgruppe blieben mit 5% die richtigen Antworten im Post-Test genauso niedrig wie im Transfer-Test (Tabelle 43).

Tabelle 43

*Prozessanalyse für das Ereignis  $[\neg p \ q \ \text{irrel}]$*

	Prä-Test	Training I	Training II	Post-Test	Transfer-Test
TG 1	10%	19%	38%	5%	52%
TG 2	5%	100%	67%	29%	95%
KG	5%	—	—	5%	5%

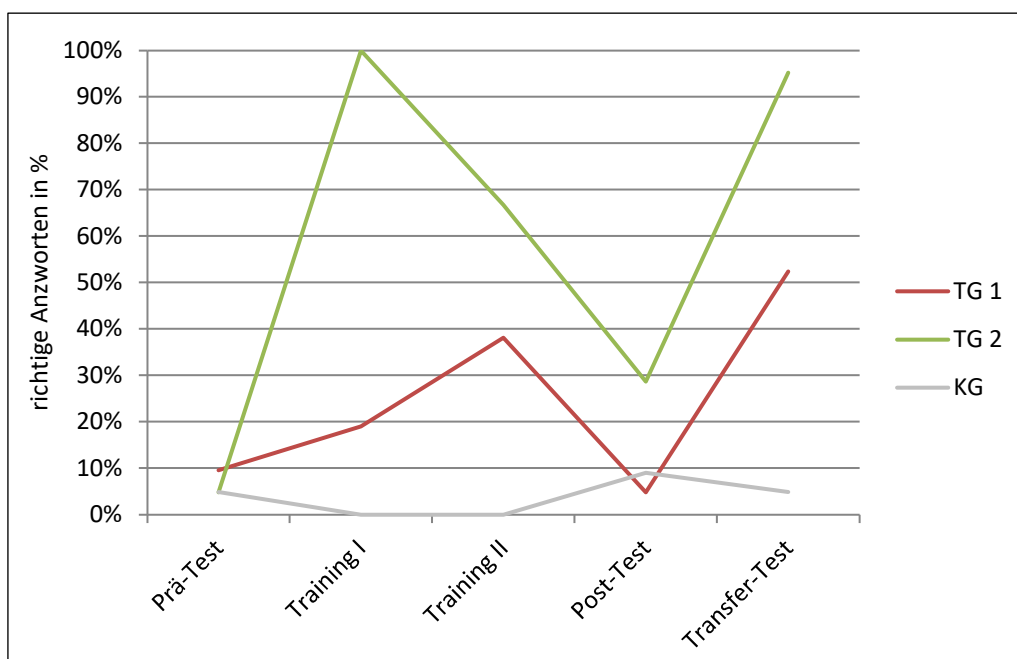


Abbildung 26. Prozessanalyse für das Ereignis [¬ p q irrel].

Blickt man auf die Ergebnisse beim Ereignis der Form [¬ p¬q irrel] (siehe Tabelle 44), so zeigte sich eine ähnliche Entwicklung wie beim Ereignis [¬ p q irrel]. In der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) stiegen die korrekten Antworten von 5% im Prä-Test auf 10% im Post-Test deutlich an (siehe Tabelle 45). Ebenso stellte sich auch in der Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) ein deutlicher Anstieg richtiger Antworten von 10% im Prä-Test auf 29% im Post-Test heraus. In der Kontrollgruppe gab es keinen Anstieg, die Antworten blieben mit 5% im Prä-Test unverändert im Post-Test.

Tabelle 44

Prozessanalyse für das Ereignis [¬p¬q irrel]

	Prä-Test	Training 1	Training II	Post-Test	Transfer-Test
TG 1	5%	33%	33%	10%	24%
TG 2	10%	100%	57%	29%	52%
KG	5%	—	—	5%	0%

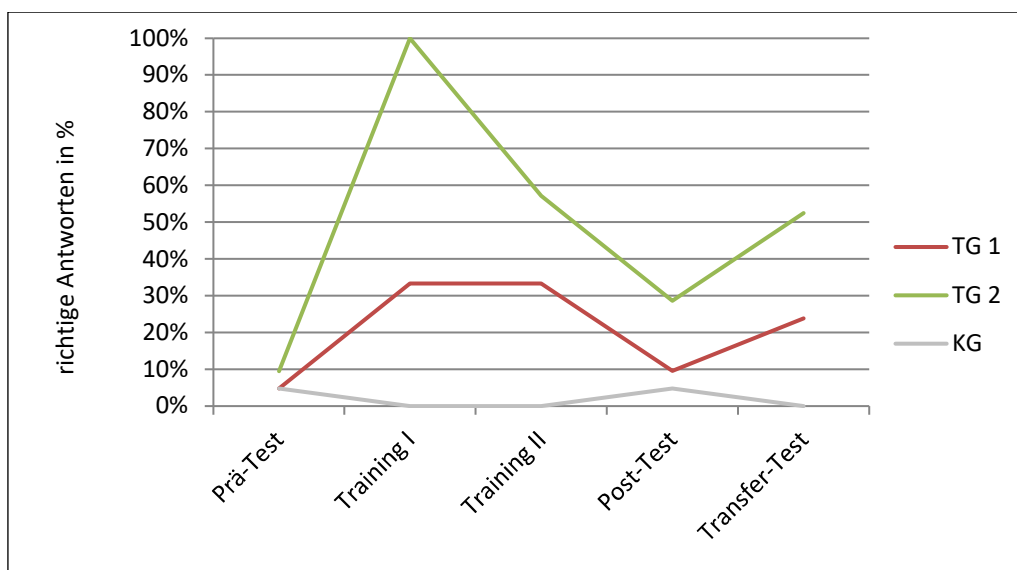


Abbildung 27: Prozessanalyse für das Ereignis  $[\neg p \neg q \text{ irrel}]$

Die Antworten der Probanden beider Trainingsgruppen in der zweiten Trainingssitzung  $[p \text{ q best}]$ ;  $[\neg p \text{ q irrel}]$ ;  $[\neg p \neg q \text{ irrel}]$  waren direkt miteinander vergleichbar, da bei diesen Ereignissen auch in Trainingsgruppe 2 nur durch adaptives Nachfragen unterstützt wurde. Falls ein Kind einen Fehler machte, lenkte der Versuchsleiter die Aufmerksamkeit des Kindes nochmals auf die zu beobachtende Eigenschaft, und das Kind konnte erneut antworten.

Vergleicht man die Antworten der Kinder von Post-Test und Transfer-Test für das Ereignis  $[\neg p \neg q \text{ irrel}]$ , so nahmen die Antworten in Trainingsgruppe 1 von 10% auf 24% und in Trainingsgruppe 2 von 29% auf 52% zu, während sie in der Kontrollgruppe von 5% auf 0% sanken (Tabelle 46). Die Kinder beider Trainingsgruppen lernten beim gruppeninternen Vergleich von Prä-Test und Transfer-Test deutlich hinzu. Die Werte beider Trainingsgruppen im Transfer-Test unterschieden sich erkennbar von der Kontrollgruppe. Insgesamt zeigte die Entwicklung der Antworten über alle Ereignisse im Verlauf über die Zeit deutlich, dass die Lösungsraten in beiden Trainingsgruppen jeweils anstiegen. In der Trainingsgruppe 1 waren höhere Werte bei den Einzelereignissen nur während der Trainingsphasen (Training I, Training II, Transfer-Test) sichtbar. Bei den Einzelereignissen war bei der Betrachtung von Prä-Test und Post-Test ein Abfall an richtigen Antworten zu erkennen. In der Kontrollgruppe fielen die Werte von Prä-Test zum Post-Test bzw. hin zum Transfer-Test bei den Einzelereignissen ab.

### 8.3.4 Ergebnisse zur vierten Forschungsfrage: Vergleich der Effekte von Trainingsdomäne und Transfer-Test

Im Rahmen der vierten Forschungsfrage wurde untersucht, wie nachhaltig der Umgang mit Evidenz von der Trainingsdomäne im Inhaltsgebiet *Elastizität und Plastizität* im Transfer-Test im Inhaltsgebiet *Schwimmen und Sinken* gewesen ist. Die Hypothese hierzu lautete, dass sich das Antwortverhalten der Probanden in ähnlicher Weise wie im trainierten Kontext zeigte. Erwartet wurde ein deutlicher Unterschied zwischen den Teilnehmern der beiden Trainingsgruppen (Scaffolding durch adaptives Nachfragen, Unterstützung durch Modellierung und adaptives Nachfragen). Dazu wurden die Antworten bei allen vier Ereignissen zu beiden Messzeitpunkten (Training II und Transfer-Test) in den beiden Trainingsgruppen miteinander verglichen. Hierbei wurde ausschließlich das zweite Training mit drei Ereignissen berücksichtigt, da in dieser Phase die Trainingsbedingungen in beiden Gruppen gleich waren.

Zur Beantwortung der Frage wurde eine *ANOVA* mit Messwiederholung für die beiden Messzeitpunkte (Training II und Transfer-Test) durchgeführt. Es wurde geprüft, ob sich signifikante Unterschiede über die Zeit zwischen beiden Gruppen vom zweiten Training zum Transfer-Test hinzeigen ließen. Dazu wurde für jedes Ereignis ein Differenzwert gebildet, der als abhängige Variable in einer *ANOVA* genutzt wurde. In Tabelle 46 sind die deskriptiven Ergebnisse aufgeführt. Betrachtet man die beiden Trainingsgruppen, so zeigte sich für das bestätigende Ereignis [p q best] sowie für das irrelevante Ereignis [ $\neg$ p q irrel] ein signifikanter Unterschied über die Zeit. Für das Ereignis [ $\neg$ p $\neg$ q irrel] ergaben sich keine signifikanten Unterschiede zwischen den beiden Messzeitpunkten von Training und Transfer-Test zwischen beiden Gruppen (siehe Tabelle 46). Das widerlegende Ereignis [p $\neg$ q wid] wurde beim Vergleich der beiden Trainingsgruppen ausgelassen, da dies in der zweiten Trainingssitzung in Trainingsgruppe 2 noch durch Modellierung unterstützt worden ist.

Tabelle 45

*ANOVA mit Messwiederholung (Training II und Transfer-Test)*

Ereignis	ANOVA mit Messwiederholung (Training II und Transfer-Test)
[p q]	$F(1, 38) = 52.45; p < .05; \epsilon^2 = 0.58$
[¬p q]	$F(1, 38) = 8.57; p < .05; \epsilon^2 = 0.18$
[¬p¬q]	$F(1, 38) = 0.24; p > .05; \epsilon^2 = 0.001$

Zur Veranschaulichung sind die Ergebnisse von Training II und Transfer-Test (Abbildung 28 bis Abbildung 30) dargestellt.

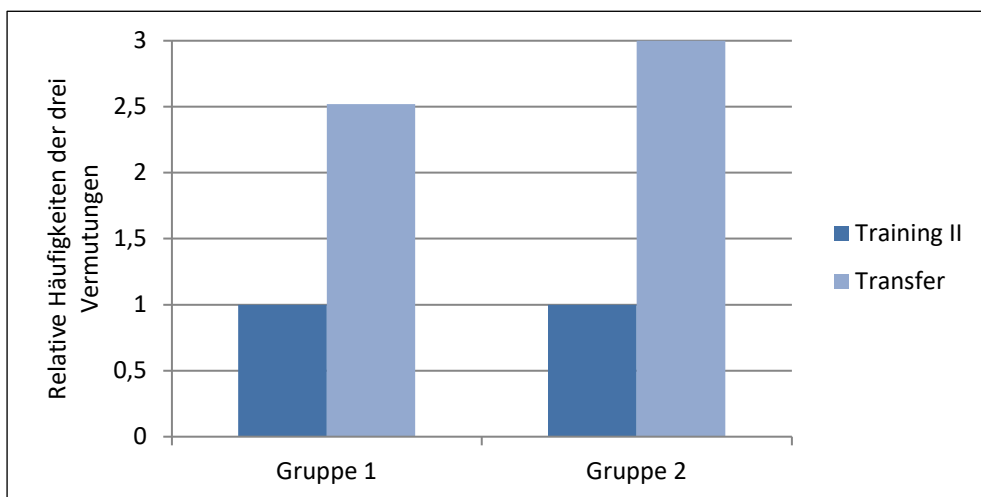


Abbildung 28. Ereignis [p q best] (Training II und Transfer-Test).

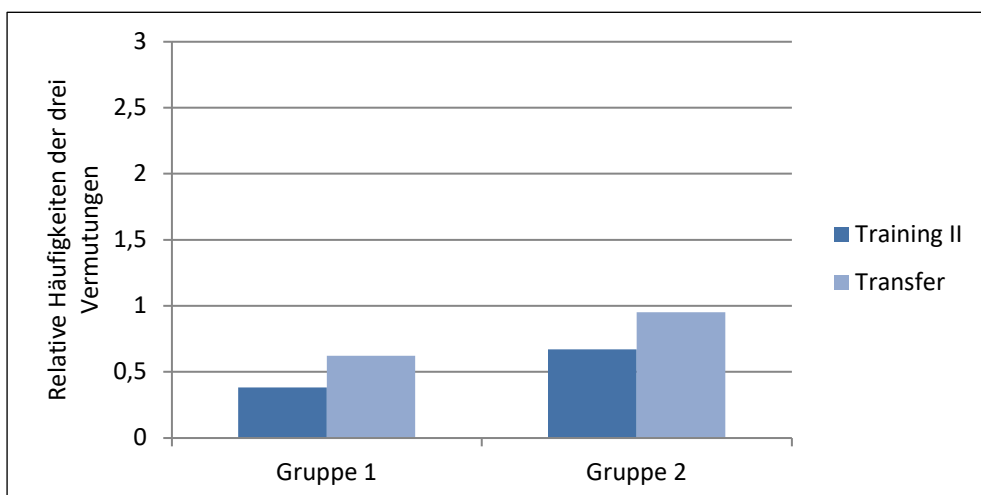
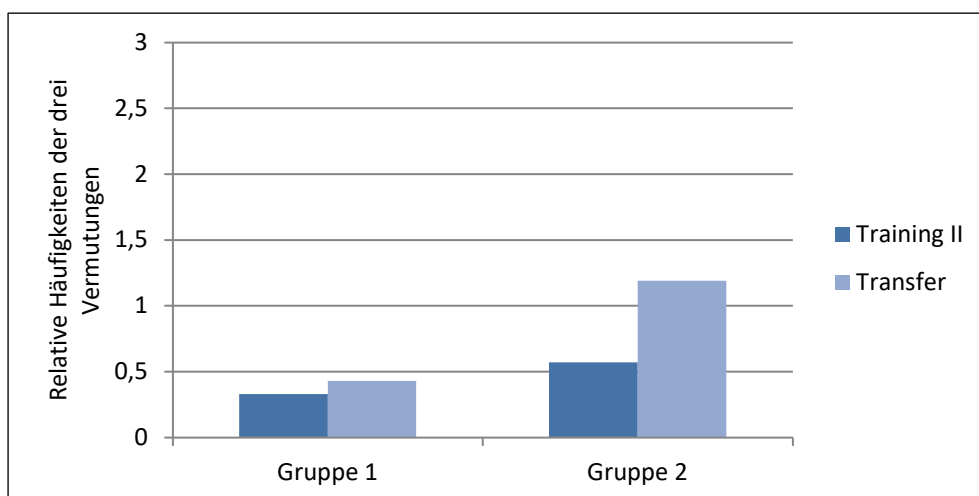


Abbildung 29. Ereignis  $[\neg p \ q \ \text{irrel}]$  (Training II und Transfer-Test).Abbildung 30.  $[\neg p \neg q \ \text{irrel}]$  (Training II und Transfer-Test).

Im Zusammenhang mit der vierten Forschungsfrage, in der untersucht wurde, wie nachhaltig die Effekte der Trainingssitzungen bei beiden Trainingsgruppen waren, werden die Ergebnisse des Transfer-Tests im Inhaltsgebiet *Schwimmen und Sinken* vorgestellt. Der Test wurde in allen drei Gruppen (Trainingsgruppe 1, Trainingsgruppe 2, Kontrollgruppe) durchgeführt. Der Vergleich ist nur bei Training II zwischen den Trainingsgruppen möglich, da diese in gleicher Weise Scaffolding-Maßnahmen erhalten haben. Daher sind nur die Ergebnisse für beide Trainingsgruppen aufgeführt. Beim Vergleich der relativen Häufigkeiten (Tabelle 47) zeigt sich, dass die beiden Trainingsgruppen bei den beiden ersten Ereignissen  $[p \ q \ \text{best}]$  und  $[p \neg q \ \text{wid}]$  fast identische Lösungsraten aufweisen (100% und 95% beim Ereignis  $[p \ q \ \text{best}]$ ; jeweils 86% beim Ereignis  $[p \neg q \ \text{wid}]$ ). Beim dritten Ereignis  $[\neg p \ q \ \text{irrel}]$  zeigten die Kinder der Trainingsgruppe 2 eine deutlich höhere Lösungsrate (95%) als die Trainingsgruppe 1 (52%); beim vierten Ereignis  $[\neg p \neg q \ \text{irrel}]$  wurde ein ähnlicher Unterschied zwischen beiden Gruppen offensichtlich: In der Trainingsgruppe 2 lagen die richtigen Ergebnisse mit 52% mehr als doppelt so hoch wie in Trainingsgruppe 1 mit 24%. Der Unterschied in Bezug auf das Antwortverhalten war in beiden Gruppen anhand der deskriptiven Ergebnisse (Abbildung 28-30) sehr deutlich. Statistisch signifikant war er beim Ereignis  $[pq \ \text{best}]$  sowie beim Ereignis  $[\neg p \ q \ \text{irrel}]$  (siehe Tabelle 43).

Tabelle 46

*Relative Häufigkeiten im Transfer-Test Schwimmen und Sinken*

Ereignis	Trainingsgruppe 1			Trainingsgruppe 2		
	best	irr	wid	best	irr	wid
[p q]	100%	0%	0%	95%	1%	0%
[p¬q]	2%	1%	86%	2%	1%	86%
[¬p q]	14%	52%	29%	11%	95%	10%
[¬p¬q]	6%	24%	32%	11%	52%	10%

*MANOVA*  $F(4, 55) 129,43; p < 0.05; \varepsilon^2 = 0.09$   
*Anmerkung.* best = bestätigend; irr = irrelevant; wid = widerlegend.

Die Mittelwerte in Tabelle 48 bestätigten das zuvor dargestellte Bild. Die Gruppenunterschiede zwischen beiden Trainingsgruppen zeigten sich bei den irrelevanten Ereignissen [¬p q irrel]; [¬p¬q irrel] besonders deutlich; beim Ereignis [¬p¬q irrel] wurde ein signifikanter Unterschied zwischen beiden Gruppen in einer Analyse mit *ANOVA* deutlich (siehe Tabelle 48).

In einer Berechnung mit *MANOVA*, in der alle vier Ereignisse als abhängige Variablen fungierten, zeigte sich mit  $F(4, 35) = 386.01; p < 0.05; \varepsilon^2 = 0.97$  ein signifikanter Unterschied zwischen den beiden Gruppen (Trainingsgruppe 1, Trainingsgruppe 2) beim Ereignis [¬p¬q irrel]. In der Berechnung wurden sowohl der Inhibitionstest (Früchte-Gemüse-Stroop) als auch die Personenfähigkeit des Wissenschaftsverständnisses (WLE) als Kovariaten berücksichtigt.

Tabelle 47

*Mittelwerte (Standardabweichungen) im Transfer-Test*

Ereignis	Trainingsgruppe 1	Trainingsgruppe 2
[p q]	3.0 (0.0)	3.0 (0.2)
[p¬q]	2.9 (0.4)	2.9 (0.4)
[¬p q]	0.1 (1.1)	2.1 (0.1)
[¬p¬q]	1.1 (1.2)	2.0 (1.2) *

\**MANOVA*  $F(4, 35) = 386.01$ ;  $p < 0.05$ ;  $\epsilon^2 = 0.97$

Im nächsten Schritt wurde die Anzahl an eingesetzten Unterstützungsmaßnahmen im Transfer-Test in beiden Trainingsgruppen in einer Analyse mit einem T-Test für unabhängige Stichproben verglichen (Tabelle 48). Es ergab sich mit  $T(39) = 1.6$ ;  $p = ns.$  kein signifikanter Unterschied zwischen den beiden Gruppen im Hinblick auf die Anzahl der Unterstützungsmaßnahmen.

Tabelle 48

*Mittelwerte (Standardabweichungen) der Unterstützungsmaßnahmen beim Transfer-Test*

Trainingsgruppe 1	Trainingsgruppe 2
5.24 (2.02)	3.95 (2.95)

$T(39) = 1.6$ ; ns.

### 8.3.5 Ergebnisse zur fünften Forschungsfrage: Argumentationstest

Zur Beantwortung der fünften Forschungsfrage, die untersuchte, ob sich Unterschiede zwischen den drei Gruppen im Argumentationstest zeigten, wurde die Verteilung der relativen Häufigkeiten in den drei Gruppen dargestellt. Hierbei wurde die Kontrollgruppe als



Vergleichsgruppe wieder einbezogen, da alle drei Gruppen beim Argumentationstest keinerlei Unterstützung erhalten hatten. Im Argumentationstest konnten maximal vier Punkte erreicht werden (für eine genaue Beschreibung siehe Kapitel. 8.1.2).

An der Verteilung der Häufigkeiten (Tabelle 51) ist zu erkennen, dass 0 und 1 Punkt von etwa 25% bis 40% von Kindern in allen drei Gruppen erreicht wurden. Zwei und drei Punkte erlangten in jeder Gruppe über 50% der Probanden, während vier Punkte in allen drei Gruppen von weniger als 20% der Kinder erreicht wurden. Mittels einer ANOVA wurde geprüft, ob Mittelwertsunterschiede zwischen den drei Gruppen (Trainingsgruppe 1, Trainingsgruppe 2, Kontrollgruppe) vorlagen.

Tabelle 49

*Häufigkeitsverteilung Argumentationstest*

Punkte (0-4)	Trainingsgruppe 1		Trainingsgruppe 2		Kontrollgruppe	
	Häufigkeit	Prozent	Häufigkeit	Prozent	Häufigkeit	Prozent
0	4	19%	0	0%	1	5%
1	5	24%	2	43%	4	19%
2	7	33%	9	43%	7	33%
3	1	5%	9	43%	7	33%
4	4	19%	1	5%	2	10%

*ANOVA*  $F(2, 60) = 1.24$  ns.;  $\epsilon^2 = 0.04$

Die Übersicht in Tabelle 51 zeigt die Verteilung der Antworthäufigkeiten in den drei Experimentalgruppen im Argumentationstest. Aufgrund der vorliegenden Verteilung (siehe Tabelle 51) wurden zwei Wertebereiche gebildet (0–1 Punkt = 0; 2–4 Punkte = 1). Knapp 40% der Probanden (38%) lagen mit ihren Antworten im Wertebereich 1 (zwei, drei oder vier richtige Antworten).

Tabelle 50

*Häufigkeiten im Argumentationstest (nach Wertebereichen)*

Wertebereich	Häufigkeit	Prozent
0	39	62%
1	24	38%

*ANOVA*  $F(2, 60) = 1.07$ ;  $p > 0.05$  ns.;  $\epsilon^2 = 0.03$

Anhand der Tabelle zeigte sich, dass über die Hälfte der Kinder (62%) im Wertebereich 0 (keine oder eine richtige Antwort) antworteten. Die Reliabilität des Argumentationstests war mit Cronbachs Alpha von 0.77 akzeptabel.

### 8.3.6 Ergebnisse zur sechsten Forschungsfrage:

#### *Schlussfolgern, Argumentieren*

Auf Grundlage von empirischen Befunden (Chinn & Brewer, 2001; Lawson, 2005; Tytler, 2005) wurde der Zusammenhang zwischen Schlussfolgerungskompetenzen und Argumentieren geprüft. Dazu wurden die Antworten aus dem Post-Test beim Schlussfolgern in den fortgeschrittenen Interpretationsniveaus konditional und bikonditional aufsummiert. Insgesamt beinhaltete die Stichprobe 63 Antworten, die nach ihrer Qualität zwei Kategorien zugeordnet worden sind. Anschließend wurde innerhalb der Stichprobe überprüft, welche Ergebnisse im Argumentationstest erzielt wurden (siehe Tabelle 52).

Tabelle 51

*Zusammenhang zwischen Schlussfolgern und Argumentieren*

Wertebereiche des Post-Tests Schlussfolgern *	Argumentationstest *: Anzahl der Antworten im Wertebereich 0	Argumentationstest **: Anzahl der Antworten im Wertebereich 1	Post-Test Schlussfolgern: Anzahl der Antworten auf dem konditionalen, bikonditionalen Niveau
0	31	14	45
1	8	10	18
Gesamt	39	24	63

\*Schlussfolgern: 0 = keine oder eine Antwort bikonditional, konditional;

1 = zwei oder drei Antworten bikonditional, konditional;

\*\*AGT: zwei Wertebereiche (0 = keine bzw. eine richtige Antwort;

1 = zwei, drei, vier richtige Antworten)

$\chi^2 = (1, 63) 3.26; ns.$

Der Zusammenhang von Schlussfolgerungskompetenz und Fähigkeiten beim Argumentieren wurde mithilfe eines Chi-Quadrat-Tests geprüft. Das Ergebnis ( $\chi^2 = (1, 63) 3,26; ns.$ ) zeigte, dass kein statistisch signifikanter Zusammenhang zwischen ausgeprägten Kompetenzen beim Schlussfolgern und guten Fähigkeiten beim Argumentieren bei der vorliegenden Stichprobe bestand.

### 8.3.7 Zusammenfassung und Fazit

In diesem Kapitel wurde die Wirkungsweise der beiden Trainingsmaßnahmen (Scaffolding durch adaptives Nachfragen, Unterstützung durch Modellierung und adaptives Nachfragen) verglichen. Dazu wurden die Ergebnisse des Schlussfolgerungstests im Kontext *Elastizität und Plastizität* des Prä-Tests sowie des Post-Tests in allen Gruppen miteinander verglichen. Dabei hatten sich signifikante Unterschiede zwischen Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) und Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) besonders beim Ereignis  $[\neg p \rightarrow q \text{ irrel}]$  im Post-Test gezeigt. Betrachtete man die Interpretationsniveaus (Abbildung 22), so zeigte sich, dass die Teilnehmer in Trainingsgruppe 2 besonders viel hinzugelernt hatten. Knapp 30% der Probanden aus Trainingsgruppe 2 hatten im Post-Test auf dem konditionalen Niveau geantwortet, und damit ihre

Schlussfolgerungskompetenzen um 25% steigern können (siehe Abbildung 22). In der Berechnung mit *MANOVA* ergab sich nur ein signifikanter Unterschied von Prä-Test und Post-Test beim Ereignis [ $\neg p \neg q$  irrel]. Zusätzlich zu den erläuterten Analysen wurden Differenzwerte von Prä-Test und Post-Test gebildet. In Berechnungen mit *ANOVA* diente jeweils ein Differenzwert eines Ereignisses als abhängige Variable. Signifikante Unterschiede von Prä-Test- und Post-Testwerten stellten sich für die Ereignisse [ $p q$  best] und [ $\neg p \neg q$  irrel] heraus.

Weiterhin wurde zur Beantwortung der zweiten Forschungsfrage die Effektivität beider Trainingsmaßnahmen anhand der Anzahl der eingesetzten Unterstützungsmaßnahmen (Scaffolding durch adaptives Nachfragen, Unterstützung durch Modellierung und adaptives Nachfragen) in Trainingssitzung II für drei Ereignisse [ $p q$  best]; [ $\neg p q$  irrel]; [ $\neg p \neg q$  irrel] anhand von T-Tests miteinander verglichen. Dabei stellte sich bei einem Vergleich der Mittelwerte heraus, dass Probanden der Trainingsgruppe 2 signifikant weniger Unterstützungsmaßnahme benötigten als Probanden der Trainingsgruppe 1 (siehe Tabelle 40).

Zusätzlich wurde eine Prozessanalyse über fünf Messzeitpunkte (Prä-Test, Training I, Training II, Post-Test, Transfer-Test) durchgeführt, um die Entwicklung der Schlussfolgerungskompetenzen über einen Zeitraum von drei Wochen zu ermitteln. In einer weiteren Analyse wurde mit Hilfe einer *ANOVA* mit Messwiederholung geprüft, wie nachhaltig der trainierte Umgang mit Evidenz im Transfer-Kontext *Schwimmen und Sinken* sich erwiesen hatte. Die Kinder beider Trainingsgruppen zeigten deutliche Unterschiede bei Antworten mit irrelevanten Ereignissen [ $\neg p q$  irrel] und [ $\neg p \neg q$  irrel]. Kinder der Trainingsgruppe 2 hatten statistisch nachweisbar höhere Werte beim Ereignis [ $\neg p q$  irrel] gezeigt (siehe Tabelle 45). Die Förderung durch die Trainingsmaßnahmen gelang demnach in Trainingsgruppe 2 bei diesem irrelevanten Ereignis besonders gut. Die Kinder in Trainingsgruppe 1 hatten auch beim Training profitiert; dies zeigte sich in den deskriptiven Werten von den beiden Trainingsphasen sowie im Transfer-Test (Tabellen 43–46). Statistisch nachweisen ließ sich der Lerngewinn beim Vergleich von Training II und Transfer-Test allerdings nur in Trainingsgruppe 2. Ein statistisch messbarer Unterschied zwischen beiden Trainingsgruppen konnte im Transfer-Test bei den Ereignissen [ $p q$  best] und [ $\neg p q$  irrel] nachgewiesen werden (siehe Tabelle 48). Weiterhin wurde im Zusammenhang mit der fünften Forschungsfrage ein Argumentationstest durchgeführt (siehe Tabelle 51) Dabei zeigten sich keine messbaren Unterschiede zwischen beiden Trainingsgruppen und der Kontrollgruppe. Im Rahmen der sechsten Forschungsfrage wurden Zusammenhänge zwischen Schlussfolgern auf fortgeschrittenem Niveau und Argumentieren geprüft. Dabei zeigte sich kein statistisch nachweisbarer Zusammenhang zwischen Schlussfolgern und Argumentieren. Die empirischen Befunde (Chinn & Brewer,

2001; Lawson, 2005; Tytler, 2005) konnten daher in der vorliegenden Stichprobe nicht bestätigt werden.

#### 8.4 Diskussion

In der Diskussion werden die Ergebnisse gemäß den sechs Forschungsfragen zusammengefasst und diskutiert. Die erste Forschungsfrage zielte darauf ab, herauszufinden, ob sich Schlussfolgerungskompetenzen durch ein Training fördern lassen. Die relativen Häufigkeiten in allen drei Gruppen (Trainingsgruppe 1, Trainingsgruppe 2, Kontrollgruppe) haben keine Zunahme bei den Ereignissen mit  $p$  ( $[p \rightarrow q \text{ best}]$ ;  $[p \rightarrow q \text{ wid}]$ ) aufgewiesen. Dies ist darauf zurückzuführen, dass die Erfolgsraten im Prä-Test hier in allen drei Experimentalgruppen mit über 75% bereits hoch waren. Bei den Ereignissen mit  $\neg p$ , ( $[\neg p \rightarrow q \text{ irrel}]$  und  $[\neg p \rightarrow \neg q \text{ irrel}]$ ), gab es keine Erhöhung der korrekten Antworten in der Kontrollgruppe sowie in Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen). Bei Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) stieg die Rate der richtigen Antworten bei den irrelevanten Ereignissen mit  $\neg p$  deutlich an. Auf Basis der Einzelereignisse wurden Interpretationsniveaus gebildet. In einer Analyse mit *MANOVA* ergab sich vom Prä-Test hin zum Post-Test eine signifikante Steigerung über alle Gruppen hinweg (siehe Abbildung 22). Im Hinblick auf die Förderung der Schlussfolgerungskompetenz für die Ereignisse mit  $\neg p$  konnte sie besonders für Trainingsgruppe 2 bestätigt werden. Für Trainingsgruppe I stellte sich eine signifikante Steigerung vom Prä-Test hin zum Post-Test heraus. In einer Berechnung mit *MANOVA* zeigte sich folgendes signifikante Ergebnis.  $F(1, 60) = 3.85$ ;  $p = .05$ ;  $\epsilon^2 = 0.6$  Insgesamt konnte daher die erste Forschungsfrage bejaht werden.

Die zweite Forschungsfrage untersuchte, ob es Unterschiede in der Wirkung beider Trainingsmaßnahmen (Scaffolding durch adaptives Nachfragen, Unterstützung durch Modellierung und adaptives Nachfragen) im Hinblick auf die Schlussfolgerungskompetenzen gab. Bei einem Vergleich von Prä-Test und Post-Test im Hinblick auf Einzelereignisse (verglichen durch Differenzwerte) wurden Unterschiede zwischen den Gruppen deutlich. Bei Betrachtung der Interpretationsniveaus (konditional, bikonditional, konjunktiv, RFRF, RFFR, Sonstige) zeigten sich deutliche Differenzen zwischen den Trainingsgruppen. Beim konditionalen, dem bestmöglichen Interpretationsniveau, konnten die Probanden in der Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) im Post-Test einen hohen Zuwachs an richtigen Antworten erreichen. Probanden der Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) hatten beim konditionalen Interpretationsniveau

einen deutlich geringeren Zuwachs. Probanden der Kontrollgruppe antworteten im Post-Test nicht häufiger auf einem der fortgeschrittenen Interpretationsniveaus (konditional, bikonditional) als im Prä-Test.

Auf dem bikonditionalen Niveau gab es in allen drei Gruppen nicht häufiger richtige Antworten im Post-Test als im Prä-Test. In beiden Trainingsgruppen nahmen Antworten auf dem RFFR-Interpretationsniveau ab. Ähnliche Ergebnisse zeigten sich in den Trainingsgruppen beim RFRF-Interpretationsniveau (siehe Tabelle 33). In einer Analyse mit *MANOVA* hatte sich nur für das Ereignis [ $\neg p \neg q$  irrel] ein signifikanter Unterschied herausgestellt (siehe Tabelle 34). Im Hinblick auf die zweite Frage wurde die Anzahl notwendiger Trainingsmaßnahmen in Training II in beiden Trainingsgruppen im Rahmen eines T-Tests verglichen. Dabei stellten sich signifikante Unterschiede zwischen Trainingsgruppe 1 (Scaffolding durch adaptives Nachfragen) und Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) heraus. Kinder der Trainingsgruppe 2 benötigten signifikant weniger Unterstützungsmaßnahmen (siehe Tabelle 40). Die zweite Forschungsfrage konnte auf Grundlage der erläuterten Ergebnisse nur teilweise bejaht werden.

Um die dritte Frage nach den Unterschieden zwischen den drei Experimentalgruppen in Bezug auf eine nachhaltige Trainingswirkung zu beleuchten, wurde zusätzlich eine Prozessanalyse durchgeführt. Daher soll im folgenden Absatz dargelegt werden, welche Effekte die Trainingsmaßnahmen im Verlauf der Studie über alle fünf Messzeitpunkte hinweg (Prä-Test, Training I, Training II, Post-Test, Transfer-Test) ergaben (siehe Tabellen 41 - 44). In den Fällen mit [p] gab es in den Gruppen teilweise leichte Erhöhungen der richtigen Antworten vom Prä-Test hin zum Post-Test. Beim Ereignis [ $p \neg q$  wid] konnte in allen drei Gruppen eine leichte Abnahme richtiger Antworten vom Prä-Test zum Post-Test festgestellt werden. In Trainingsgruppe 2 ebenso wie in der Kontrollgruppe sanken die Werte vom Prä-Test hin zum Post-Test. Bei Trainingsgruppe 1 gab es vom Prä-Test hin zum Post-Test einen leichten Anstieg.

Beim Ereignis [ $\neg p \neg q$  irrel] zeigten sich signifikante Unterschiede zwischen Trainingsgruppe 1 und Trainingsgruppe 2. Zwischen der Kontrollgruppe und der Trainingsgruppe 1 stellte sich kein signifikanter Unterschied heraus. Insgesamt machten die deskriptiven sowie die inferenzstatistischen Ergebnisse deutlich, dass die Förderungsmaßnahmen in den Trainingsgruppen differenziert zu betrachten sind. Insbesondere in der Trainingsgruppe 2 lernten die Kinder mit Ereignissen ([ $\neg p q$  irrel] [ $\neg p \neg q$  irrel]) umzugehen, die im Zusammenhang mit einer Vermutung irrelevant waren. Zusammenfassend kann festgehalten werden, dass sowohl die Vergleiche der Mittelwerte (Differenzwerte Post-

Test, Prä-Test) als auch die inferenzstatistischen Analysen sowie die Prozessanalyse eine deutliche Zunahme korrekter Antworten im Post-Test in der Trainingsgruppe 2 beim Ereignis mit  $[\neg p \neg q \text{ irrel}]$  aufgewiesen haben. Daher konnte die dritte Forschungsfrage bejaht werden. Kinder der Trainingsgruppe 2 zeigten nachhaltig sowohl im Post-Test als auch im Transfer-Test bessere Antworten als Kinder der Trainingsgruppe 1 oder der Kontrollgruppe.

Zur Beantwortung der vierten Forschungsfrage, ob sich nachhaltige Effekte des Trainings im Transfer-Test zeigen ließen, wurden die Werte beider Trainingsgruppen zu zwei Messzeitpunkten mit Hilfe einer *ANOVA* mit Messwiederholung miteinander verglichen. Die Werte der Kontrollgruppe wurden nicht hinzugezogen, da die Probanden dieser Gruppe beim Transfer-Test keine Unterstützung erhalten hatten. Daher konnten nicht alle drei Experimentalgruppen miteinander verglichen werden. Bei der Berechnung mit einer *ANOVA* mit Messwiederholung zeigten sich im Transfer-Test signifikante Unterschiede bei den Ereignissen  $[p q \text{ best}]$  und  $[\neg p q \text{ irrel}]$  zwischen Trainingsgruppe 1 und Trainingsgruppe 2. Für das Ereignis  $[\neg p \neg q \text{ irrel}]$  stellte sich kein signifikanter Unterschied zwischen den beiden Trainingsgruppen heraus (siehe Tabelle 44). Zusätzlich wurde die Anzahl der eingesetzten Unterstützungsmaßnahmen in beiden Trainingsgruppen im Transfer-Test mit Hilfe eines *T-Tests* für unabhängige Stichproben überprüft (siehe Tabelle 49). Dabei wurde kein signifikanter Unterschied zwischen den beiden Gruppen im Hinblick auf die Anzahl der verwendeten Unterstützungsmaßnahmen deutlich. Weiterhin wurde mit den Daten des Transfer-Tests eine *MANOVA* berechnet, in der alle vier Ereignisse als abhängige Variable fungierten. Dabei ergaben sich signifikante Unterschiede beim Ereignis  $[\neg p \neg q \text{ irrel}]$  zwischen den beiden Trainingsgruppen (siehe Tabelle 44). Nach einer differenzierten Betrachtung der Lösungsraten beider Trainingsgruppen im Transfer-Test sowie nach einem Vergleich von Trainingssitzung II und Transfer-Test konnte die vierte Hypothese, es gebe signifikante Unterschiede zwischen beiden Trainingsgruppen nur teilweise bestätigt werden, da sich nur bei zwei von vier Ereignissen signifikante Unterschiede zeigten (siehe Tabelle 48).

Die fünfte Forschungsfrage untersuchte, ob zwischen den drei Gruppen (Trainingsgruppe I, Trainingsgruppe 2, Kontrollgruppe) deutliche Unterschiede im Argumentationstest bestanden. In einer Analyse mit *ANOVA* konnte kein signifikanter Unterschied zwischen den drei Experimentalgruppen nachgewiesen werden.

Die sechste Forschungsfrage prüfte, ob Schlussfolgerungskompetenzen mit der Kompetenz zum Argumentieren verknüpft sind. Empirischen Befunden zufolge besteht ein Zusammenhang zwischen beiden Fähigkeiten (Chinn & Brewer, 2001; Lawson, 2003; Toulmin, 2003). Kinder mit fortgeschrittenen Kompetenzen im Schlussfolgern können auch gut

argumentieren. Insgesamt zeigten achtzehn von dreiundsechzig Kindern der gesamten Stichprobe in der vorliegenden Studie im Post-Test gute Schlussfolgerungskompetenzen. Sie antworteten auf bikonditionalem bzw. konditionalem Interpretationsniveau. Von den achtzehn Kindern schlussfolgerten zehn Kinder auf fortgeschrittenem Niveau und erreichten gleichzeitig hohe Testwerte im Argumentationstest. Statistisch nachweisbar war der Zusammenhang zwischen Schlussfolgerungskompetenzen und dem Argumentieren nicht (siehe Tabelle 50). Somit konnte die sechste Hypothese nicht bestätigt werden. Dieses Ergebnis ließ sich damit erklären, dass nur zehn Kinder ausgeprägte Schlussfolgerungskompetenzen aufwiesen und ebenso im Argumentationstest gut abschnitten. Es ist zu erwarten, dass bei einer größeren Stichprobe mit einer größeren Anzahl von Kindern mit hochwertigen Schlussfolgerungskompetenzen (auf bikonditionalem bzw. konditionalem Antwortniveau) sich ein Zusammenhang mit guten Argumentationsfähigkeiten nachweisen ließe. Insgesamt konnten die empirischen Befunde (Chinn & Brewer, 2001; Lawson, 2003; Toulmin, 2003) mit den vorliegenden Ergebnissen nicht bestätigt werden können.

Zusammenfassend kann aus den Ergebnissen der Schluss gezogen werden, dass die untersuchten Trainingsmaßnahmen einen angemessenen Umgang mit Evidenz über einen bestimmten Zeitraum fördern. Die Trainingsmaßnahmen waren den empirischen Befunden zufolge besonders in Trainingsgruppe 2 (Unterstützung durch Modellierung und adaptives Nachfragen) wirksam über eine Zeitspanne von zwei Wochen. Dies konnte mit den zuvor beschriebenen Ergebnissen mit Hilfe des Gruppenvergleiches (Trainingsgruppe 1, Trainingsgruppe 2) gezeigt werden (siehe Kapitel 8.3.3).

## 9 Gesamtdiskussion

In dieser Arbeit wurden zwei Studien, Querschnitt-, und Trainingsstudie vorgestellt und ihre Ergebnisse auf Grundlage der Forschungsfragen dargestellt. Die Gesamtdiskussion dient der Zusammenfassung und abschließenden Bewertung der Studien im Hinblick auf die Forschungsziele der vorliegenden Arbeit. Weiterhin werden Grenzen der Studien sowie Möglichkeiten der Umsetzung in die Praxis aufgezeigt. Auf Grundlage dieser Aspekte werden Forschungsperspektiven dargestellt.

Aufbauend auf der Darstellung der Forschungslage und Forschungslücken (siehe Kapitel 6) wurden in der Arbeit zwei Studien entwickelt. Die Querschnittstudie untersuchte die Kompetenzen zur Koordination von Theorie und Evidenz sowie Schlussfolgerungskompetenzen bei Kindern in vier Altersstufen. Zudem wurde der Einfluss



von exekutiven Funktionen sowie von Kompetenzen bei der Koordination von Theorie und Evidenz auf Schlussfolgerungskompetenzen geprüft.

In der Querschnittstudie ging es darum, die genaue Erforschung des Zusammenhangs von schlussfolgerndem Denken und exekutiven Funktionen sowie die Verknüpfung schlussfolgernden Denkens mit der Koordination von Theorie-Evidenz bei Kindern zwischen vier und zehn Jahren zu untersuchen. Es hat sich ein Zusammenhang zwischen schlussfolgerndem Denken und exekutiven Funktionen herausgestellt. Weiterhin wurde mit dem Wissenschaftsverständnistest und dem Schlussfolgerungstest der Zusammenhang zwischen schlussfolgerndem Denken und der Koordination von Theorie und Evidenz geprüft. Dabei zeigte sich ein Zusammenhang zwischen der Theorie-Evidenz-Koordination und dem schlussfolgernden Denken. Diese Verknüpfung beider Komponenten nimmt mit zunehmendem Alter der Kinder ab.

Im ersten Teil der Arbeit wurden die Kompetenzen der Probanden in den vier Altersstufen beim Schlussfolgern untersucht. Der Test in Anlehnung an die *Truth Testing Task* (Wason, 1966; Barrouillet et al., 2008; Tröbst et al., 2011) machte deutlich, dass die Kompetenzen beim Schlussfolgern mit dem Alter ansteigen. So konnten die vierjährigen Kinder in über der Hälfte aller Fälle die bestätigenden Ereignisse richtig einschätzen (vgl. Ruffman et al., 1993); doch gelang es ihnen noch nicht, Ereignisse zu erkennen, die eine Hypothese widerlegen oder für eine Hypothese irrelevante Ereignisse zu identifizieren. Die empirischen Befunde zeigten, dass die Gruppe der fünf- bis sechsjährigen Probanden signifikant häufiger als die Gruppe der vierjährigen Probanden Ereignisse erkennen konnte, mit denen eine vorgegebene Hypothese widerlegt wird (siehe Kapitel 7.2.7). Der deutliche Unterschied zwischen beiden Altersgruppen im Hinblick auf einen Zuwachs an Kompetenzen hatte sich ebenso beim Inhibitionstest herausgestellt. Der Kompetenzzuwachs war beim Vergleich von benachbarten Altersgruppen offenkundig (z.B. beim Vergleich von vierjährigen mit fünf- bis sechsjährigen Probanden). In der vorliegenden Querschnittstudie konnte gezeigt werden, dass die Kompetenzen zum Schlussfolgern mit dem Alter ansteigen. Insgesamt bestätigen die Ergebnisse des Schlussfolgerungstests die bisherigen Forschungsbefunde (Barrouillet et al., 2008; Tröbst et al., 2011).

Die Ergebnisse der Querschnittstudie (siehe Kapitel 7.2.5, 7.2.6, 7.2.7) belegten zudem den Einfluss exekutiver Funktionen auf das Schlussfolgern und bestätigen die Ergebnisse von Gropen et al. (2011). Es konnte gezeigt werden, dass exekutive Funktionen mit dem Alter ansteigen. Die entwicklungspsychologische Forschung zeigt, dass exekutive Funktionen ebenso eine große Rolle im Zusammenhang mit der Entwicklung von Schulfähigkeit spielen

(Roethlisberger et al, 2010; Fitzpatrick, Mc Kinnen, Blair & Willoughby, 2014). Dabei wird hervorgehoben (Fitzpatrick et al., 2014; Gropen et al., 2011), dass Kinder mit ausgeprägten exekutiven Funktionen gut Informationen sowie Instruktionen während der Lösung einer Aufgabe präsent halten und gleichzeitig persönliche sowie äußere, aufgabenunabhängige Faktoren ignorieren können.

Weiterhin wurde in der vorliegenden Arbeit ein Wissenschaftsverständnistest zur Überprüfung der Kompetenzen bei der Theorie-Evidenz-Koordination in den verschiedenen Altersstufen der vier- bis zehnjährigen Probanden entwickelt. Die Ergebnisse des Wissenschaftsverständnistests in den beiden Studien (Querschnitt-, Trainingsstudie) zeigten, dass bereits Kinder im Elementarbereich über grundlegende Fähigkeiten bei der Koordination von Theorie und Evidenz verfügen (siehe. Kapitel 7.2.6). Die Altersgruppe der Vierjährigen konnte den schwersten Aufgabenbereich, die Aufgaben zur Produktion von Evidenz, noch nicht lösen. Daher wurde ein Bodeneffekt für diese Altersgruppe deutlich. Im Vergleich dazu konnten die fünf- bis sechsjährigen Kinder die Aufgaben zur Produktion von Evidenz deutlich besser beantworten. Sie konnten beide Komponenten, Theorie und Evidenz, häufiger koordinieren. Im Test ergab sich ein signifikanter Unterschied zwischen beiden Altersgruppen.

Weiterhin stand im Rahmen der Trainingsstudie die Entwicklung geeigneter Maßnahmen zur Förderung von Denkprozessen bei der Koordination von Theorie und Evidenz im Zentrum der vorliegenden Arbeit. Dazu dienten die beiden Trainingsmaßnahmen (Scaffolding durch adaptives Nachfragen, Unterstützung durch Modellierung und adaptives Nachfragen). Daher erhielten fünf- bis sechsjährige Probanden im Elementarbereich zwei unterschiedliche Trainingsmaßnahmen, die darauf abzielten, die Koordination von Theorie und Evidenz zu fördern und auf diese Weise die Kompetenzen der Kinder beim Schlussfolgern zu steigern. Frühere Interventionsstudien (Chen & Klahr, 1999; Klahr & Nigam, 2004) zeigten, dass Kinder in ihren Fähigkeiten zum schlussfolgernden Denken gefördert werden können. Die Ergebnisse des Post-Tests für Trainingsgruppe 2, die durch Modellierung und adaptives Nachfragen gefördert wurde, bestätigten die Befunde von Klahr & Nigam (2004). Von den beiden Trainingsmaßnahmen der vorliegenden Studie hatte sich die Maßnahme *Unterstützung durch Modellierung und adaptives Nachfragen* im Vergleich zur Maßnahme *Scaffolding durch adaptives Nachfragen* als besonders wirksam herausgestellt. Wenn eine zusätzliche Modellierung durchgeführt wurde, lernten die Kinder wesentlich besser mit schwierigen, irrelevanten Ereignissen umzugehen. Eine Erklärung für die deutlichen Lernerfolge bei der stärker gesteuerten Maßnahme mit Modellierung liegt vermutlich im Alter der Probanden.

Kinder im Alter von fünf Jahren verfügen noch über geringe Kompetenzen bei der Koordination von Theorie und Evidenz.

Dies legt den Schluss nahe, dass sie ihren Denkprozess durch gezielte Vorgaben einzelner Denkschritte besser entwickeln können als im Rahmen einer Fördermaßnahme mit weniger Steuerung wie in Trainingsgruppe 1 (Scaffolding durch adaptive Nachfragen). Aber auch Kinder in Trainingsgruppe 1 konnten ihre Kompetenzen erweitern, da sie sowohl während des Trainings als auch im Post-Test häufiger richtige Antworten äußerten als im Prä-Test. Beim Vergleich von Prä-Test und Post-Test ließ sich aber kein signifikanter Unterschied zwischen Trainingsgruppe 1 und Kontrollgruppe feststellen.

Im folgenden Absatz werden Reflexionen in Bezug auf den Verlauf der Erhebungen in der Trainingsstudie dargestellt; zudem werden einige alternative Möglichkeiten zur Förderung von Kindern bei der Koordination von Theorie und Evidenz aufgezeigt. Insgesamt hat die praktische Umsetzung gut funktioniert. Die Kinder sollten sich über einen Zeitraum von mindestens 30 Minuten im Rahmen einer Trainingsphase konzentrieren. Nun stellt sich die Frage, warum hat die Arbeit mit den Kindern so gut funktioniert? Verschiedene Gründe tragen dazu bei, dass die Erhebungen mit den Kindern erfolgreich waren und allen beteiligten Personen in der Regel Freude bereiteten. Zum einen könnte es daran liegen, dass die Kinder immer wieder selbst tätig waren und Rückmeldung durch die Versuchsleiterin im Gespräch erhielten. Zudem waren die Probanden nach dem ersten Testtag mit dem Thema vertraut und konnten sich gut in die Vermutungen anderer Kinder hineinversetzen, da es sich um typische Vermutungen aus der Lebenswelt von Kindern handelte. Dabei haben sie überlegt, ob das beschriebene Kind mit seiner Vermutung richtig lag oder nicht und mit welcher Evidenz diese bestätigt oder widerlegt werden konnte. Weiterhin nutzte die Versuchsleiterin motivierende Materialien (Fotos, Bilder, sowie Spielzeug), so dass die Kinder quasi im Spiel zum Denken angeregt wurden.

Die in der Untersuchung verwendeten Trainingsmaßnahmen, besonders die Modellierung, waren effektiv. Dennoch sind auch andere Maßnahmen zur Förderung der Kompetenzen bei der Koordination von Theorie-Evidenz vorstellbar. Bei der Überprüfung der Vermutung „Dinge, die mit Luft gefüllt sind, springen“ wäre es auch möglich gewesen, dass Kinder zwei Luftballons jeweils mit unterschiedlich viel Luft aufblasen und anschließend direkt ausprobieren, ob sie springen oder nicht. Dabei ist die Überlegung zentral, ob die Vermutung bestätigt oder widerlegt wird. Der Luftballon mit viel Luft bestätigt die Vermutung, der andere mit wenig Luft widerlegt diese. In der sprachlichen Umsetzung im Rahmen der Erhebung könnte die Frage lauten: „Zeigt der Luftballon, bei dem du viel Luft reinpustest, dass Mia Recht hat, oder zeigt er, dass sie nicht Recht hat?“ (vgl. Anhang J\_2; vgl. Anhang K\_2). In gleicher

Weise könnte die Frage bei dem Luftballon mit wenig Luft gestellt werden. Im nächsten Schritt könnte mit dem Kind gemeinsam überlegt werden, wie es mit Gegenständen aussieht, die ohne Luft gefüllt sind. Hierzu könnte ein Gummiball sowie ein aufgeschnittener Luftballon verwendet werden. Beim Gummiball könnte der Versuchsleiter beispielsweise erklären, dass der Ball vollständig aus Gummi und vollkommen anders gefüllt ist als der Luftballon. Hier könnte die Frage gestellt werden: „Kann ich mit dem Gummiball prüfen, ob Mia Recht hat? Der Gummiball ist anders als der Luftballon, weil er ohne Luft ist. Glaubst du, ich kann damit Mias Vermutung prüfen?“ Durch adaptives Nachfragen könnte das Kind zum Denken angeregt werden.

Eine weitere Trainingsmöglichkeit könnte darin bestehen, ein kurzes Video mit den einzelnen Denk- und Handlungsschritten dem Probanden vorzuspielen. Im Anschluss daran agiert der Proband gemäß dargestellten Aktivitäten und spricht dabei die Denkschritte mit. Je nach Ausmaß der Merkfähigkeit der Kinder müsste das Video in Abschnitte eingeteilt werden. Nach jedem Abschnitt könnte das Kind die einzelnen Schritte nachahmen. Im zweiten Evaluationskontext könnte die Modellierung reduziert und durch adaptives Nachfragen durch den Versuchsleiter ersetzt werden.

Denkbar wäre es auch gewesen, Kinder eigene Vermutungen aufstellen und überprüfen zu lassen. Hier müssten Kinder stärker ihre eigenen Ideen einbringen. Dazu wäre ein umfangreiches Materialangebot zur Überprüfung einzelner Vermutungen erforderlich. Im Rahmen einer Förderung müsste der Versuchsleiter ein Kind bei der Auswahl der Materialien viel unterstützen, damit das Kind jeweils auf die relevanten Merkmale zur Überprüfung einer Vermutung achtet. Die Gegenstände könnten entsprechend einer Vermutung nach Material und Sprungverhalten geordnet werden. Im Anschluss daran könnten Probanden zusammen mit dem Erwachsenen eine Vermutung überprüfen.

Im Vorfeld der Untersuchung hatten wir uns gegen diese Art der Gestaltung der Lernumgebung entschieden, da die erläuterte Aufgabe eine weitaus höhere Anforderung an die Probanden gestellt hätte. Für die Altersgruppe der Fünf- bis Sechsjährigen im Elementarbereich ist das eigene Aufstellen und Überprüfen von Vermutungen sehr anspruchsvoll. Probanden müssen überlegen, welche Objekte sich eignen, um eine Vermutung zu bestätigen oder zu widerlegen.

Sinnvoll erscheint die zuvor beschriebene Form der Umsetzung im Anschluss an eine Phase der Förderung, da Kinder dann bereits mit dem Thema vertraut sind. Sie wissen dann, was eine Vermutung ist und wie man sie überprüfen kann. Bei einer Experimentierphase, wie sie gerade dargestellt worden ist, lernt ein Kind mit der Kontrollvariablenstrategie umzugehen.

An den beiden vorliegenden Studien haben ausschließlich Kinder mit guten bzw. sehr guten Sprachkenntnissen teilgenommen. Naturwissenschaftliches Lernen, Konzepterwerb und Erwerb relevanter Fachbegriffe sind eng miteinander verknüpft (Hardy, Mannel & Sauer, 2015). Nun stellt sich die Frage, wie die Trainingsmaßnahmen adaptiert werden müssten, damit auch Kinder mit Migrationshintergrund und wenigen Deutschkenntnissen eine effektive Förderung erhalten können. Im Gespräch mit dem Kind sollte der Erwachsene im Sinne einer Modellierung einen umfangreichen sprachlichen Input dem Kind bieten. Sinnvoll ist es, dass dieser an das Lernniveau des Kindes anknüpft und eine Stufe über dem Lernniveau des Kindes liegt. Wichtig dabei sind zudem die Anregung zu Vergleichen sowie eine Einbettung in einen inhaltlichen Kontext, damit unbekannte, neue Begriffe besser verständlich werden (Hardy, Mannel & Sauer, 2015).

Im Vorfeld einer Erhebung und einer Trainingsmaßnahme müssten Kinder die zentralen Begriffe im Kontext *Elastizität und Plastizität* wie beispielsweise „springen, nicht springen“, „hart, weich“ und „Vermutung überprüfen“, „bestätigen“ etc. anhand von Bildern oder Tätigkeiten aufgezeigt bekommen.

In spielerischer Weise sollten die Begriffe mehrfach wiederholt werden, damit sie den Kindern verständlich werden. Die sprachlichen Sequenzen sollten kürzer als bei den vorliegenden Trainingsmaßnahmen sein, damit ein Kind mit wenig Sprachverständnis die Aufgaben dennoch lösen kann. Sowohl bei der Förderung durch Modellierung (siehe Anhang J\_2; siehe Anhang K\_2) als auch durch adaptives Nachfragen (siehe Anhang H\_1, Anhang H\_2) müssten die sprachlichen Anteile reduziert und durch Mimik und Gestik bzw. durch Symbolkarten ersetzt werden. Bei der Überlegung, ob ein Gegenstand Luft enthält, könnte beispielsweise zusätzlich zum Objekt ein Bild mit entsprechendem Symbol genutzt werden. Zusätzlich könnten Eigenschaften mit Mimik bzw. Gestik dargestellt werden (z.B. durch pusten). Ebenso könnte das Sprungverhalten des Gegenstandes durch die eigene Körperbewegung (z.B. springen bzw. bewusstes stehenbleiben) sowie durch ein Bild mit passendem Symbol dargestellt werden. Zusätzlich wäre im Laufe eines Evaluationsschrittes das Aufzeigen der Denkschritte mit Hilfe von Symbolen sinnvoll. Diese sollten während der gesamten Phase für das Kind sichtbar sein und passend zum Denkschritt jeweils fokussiert angeschaut werden. Zusammenfassend kann festgehalten werden, dass die Maßnahmen zur Förderung der Koordination von Theorie und Evidenz für Kinder mit Migrationshintergrund in sprachlicher Hinsicht verändert werden müssten.

*Grenzen der Studien*

In den hier untersuchten Studien wurde der Schlussfolgerungstest im Inhaltsgebiet *Elastizität und Plastizität* in Anlehnung an die Studien von Barrouillet et al. (2008) sowie von Tröbst et al. (2011) durchgeführt. Darin wurde der Schlussfolgerungstest mit Hilfe von Interviewleitfäden realisiert. Dies hat den Vorteil, dass ein weitgehend standardisiertes Vorgehen auch bei jungen Kindern möglich ist. So bestehen vergleichbare Bedingungen bei der Testdurchführung und der Versuchsleiter kann sich darauf konzentrieren, dass der jeweilige Proband die Aufgabe versteht. Dabei kann der Versuchsleiter in der Testphase jedoch nicht auf mögliche Fragen oder Ideen der Kinder eingehen und also auch keinen dynamischen Prozess innerhalb eines Gesprächs zwischen sich und dem Probanden initiieren. Möglicherweise üben unbeantwortete Fragen auf Seiten der Kinder einen negativen Einfluss auf ihre Motivation und damit auf ihr Antwortverhalten aus.

Der Zeitpunkt der Durchführung war aus organisatorischen Gründen in beiden Studien nicht bei allen Probanden gleich; manche Teilnehmer beantworteten die Fragen des Interviews am Vormittag, andere Teilnehmer am Nachmittag. Die Tageszeit beeinflusst möglicherweise die Konzentrationsfähigkeit, die in der Studie nicht kontrolliert werden konnte.

Die Probanden kamen aus bildungsnahen Familien der Mittel- bzw. oberen Mittelschicht und verfügten alle über gute Deutschkenntnisse. Bei der Zusammensetzung der Stichproben wurde in beiden Studien (Querschnitt- und Training) darauf geachtet, ausschließlich Kinder ohne Migrationshintergrund zu wählen. Das Ziel war möglichst optimale Voraussetzungen für das Bearbeiten der Aufgaben zu schaffen. Empirische Befunde belegen, dass ein Zusammenhang zwischen exekutiven Funktionen und sozioökonomischem Status (Fitzpatrick et al., 2014; Noble et al., 2005) besteht. Die erläuterten besonderen Testbedingungen (siehe Kapitel 7.2.2; 8.1.5; 8.2.2) sowie die spezifische Zusammensetzung der beiden Stichproben verdeutlichen, dass die Studien unter speziellen Voraussetzungen durchgeführt wurden. Daher lassen sich die Trainingsbedingungen sowie die Ergebnisse nicht einfach auf Kinder mit Migrationshintergrund übertragen.

Weiterhin wurden im Anschluss an die Trainingsphasen ein Post-Test sowie ein Transfer-Test durchgeführt, um die Nachhaltigkeit der Trainingsmaßnahmen zu überprüfen. Dabei hatten Kinder der beiden Trainingsgruppen gleiche Testbedingungen; Probanden beider Gruppen wurden im Transfer-Test mit Scaffolding durch adaptives Nachfragen unterstützt. Falls sie fehlerhaft antworteten, fragte der Testleiter nach und die Kinder hatten die Möglichkeit den Gegenstand in Bezug auf die Vermutung erneut zu bewerten. Kinder der Kontrollgruppe erhielten keinerlei Unterstützung. Aufgrund der unterschiedlichen Bedingungen ist ein Vergleich mit der Kontrollgruppe nicht möglich. Sowohl die Interviews in der

Querschnittstudie als die Trainingssitzungen wurden als Einzelinterviews durchgeführt, um möglichst genau feststellen zu können wie jedes Kind antwortet. In der Trainingsstudie ging es darum durch möglichst gute Voraussetzungen im Trainingssetting zu etablieren, um erkennen zu können, welche Trainingsmaßnahme lernwirksam ist.

Als Fazit kann festgehalten werden, dass der Einsatz von Interviewleitfäden vergleichbare Ergebnisse sowohl in der Querschnitt-, als auch in der Trainingsstudie ermöglichte. Allerdings kann bei dieser Art der Durchführung kein dynamischer Gedankenaustausch zwischen Kind und Erwachsenem ermöglicht werden. Weiterhin kann abschließend gesagt werden, dass mit der Querschnittstudie die Schlussfolgerungsfähigkeit in den Altersstufen bzw. auf den einzelnen Kompetenzniveaus erfasst werden konnte. Zudem wurde der Einfluss exekutiver Funktionen sowie der Zusammenhang von Wissenschaftsverständnis und Schlussfolgerungsfähigkeit ermittelt.

Die Trainingsmaßnahmen in der zweiten Studie waren altersentsprechend gewählt. Das Training im Rahmen der Modellierung in Kombination mit adaptivem Nachfragen war sehr effektiv. Zudem war die Zusammensetzung der vorliegenden Stichprobe zentral für die Passung von Kompetenzniveau und dem Maß an Unterstützung; in der Studie war ein angemessenes Zusammenspiel der genannten Aspekte gegeben. Daher konnte eine effektive Förderung der kindlichen Denkprozesse in der Trainingsstudie gelingen.

#### *Ausblick für Forschungsperspektiven*

Im folgenden Absatz werden Forschungsperspektiven aufgezeigt, um auf weitere sinnvolle Analysen hinzuweisen. Diese werden vor allem für die Trainingsstudie skizziert, da auf diese Weise eine noch differenziertere Wirkungsweise einer Trainingsmaßnahme herausgearbeitet werden könnte als es mit der Stichprobe von 63 Kindern möglich war. Zweckgemäß wäre es eine weitere Trainingsgruppe in die Stichprobe aufzunehmen, in der ausschließlich mit Modellierung unterstützt wird, um die isolierte Wirkung von dieser Trainingsform zu identifizieren. Zusätzlich wäre es ratsam jede Gruppe auf 40 Probanden zu erhöhen. Folglich hätte man dann vier Experimentalgruppen mit einem 4 x 4 Design und insgesamt 160 Probanden. Mit den Daten einer wesentlich größeren Stichprobe könnten Mehrebenenregressionen angefertigt werden. Damit könnte die Wirkung der einzelnen Trainingsmaßnahmen zu den verschiedenen Zeitpunkten (Training I, Training II, Transfer-Test) bei den Probanden differenzierter dargestellt werden als in den Analysen der vorliegenden Studie. Zudem könnte sich die Wirksamkeit von den erläuterten Trainingsmaßnahmen

möglicherweise noch klarer zeigen, wenn mindestens drei Sitzungen (anstatt zwei Sitzungen wie in dieser Studie) aufeinander folgen und damit der Trainingszeitraum länger wäre.

In der vorliegenden Studie konnte kein Zusammenhang zwischen Schlussfolgern und Argumentieren festgestellt werden (siehe Kapitel. 8.3.6). Empirische Befunde (z.B. Lawson et al., 2010) weisen einen statistisch signifikanten Zusammenhang zwischen beiden Fähigkeiten nach. Auch in dieser Hinsicht wäre es sinnvoll in einer Folgestudie, die Anzahl der Probanden deutlich zu erhöhen (z.B.  $n > 40$ ), um so in jeder Gruppe mehr Kinder im Post-Test mit guten Schlussfolgerungsfähigkeiten (z.B. auf bikonditionalem bzw. konditionalem Niveau) zu erhalten. In einer größeren Stichprobe können Zusammenhänge leichter ermittelt werden. Der Zusammenhang zwischen Schlussfolgerungsfähigkeit und Argumentationstest müsste erneut geprüft werden. Künftige Forschung sollte die erläuterten Erkenntnisse dieser Studie zur Förderung schlussfolgernden Denkens aufgreifen und in weiteren altersgerechten Kontexten mit anschaulichen Materialien erproben.

Eine Umsetzung der erprobten Fördermaßnahmen in die Praxis ist generell möglich. Denkbar wäre eine Förderung mit einer Gruppe von vier Kindern, wobei jeweils zwei Kinder im Rahmen einer Partnerarbeit eine Vermutung mit Hilfe von vier Gegenständen überprüfen und diese dabei auf die Unterlagen entsprechend der Vermutung (z.B. harte, weiche Unterlage) und des Sprungverhaltens (springt, springt nicht) zuordnen könnten. Die Aufgabe des Erwachsenen bestünde dann darin, schrittweise durch das Lernangebot zu führen und im Sinne einer Fokussierung den Blick der Kinder immer wieder auf das relevante Merkmal (z.B. Eigenschaft, Sprungverhalten) im Zusammenhang mit der vorliegenden Vermutung zu richten. Bei einer Förderung in einer Gruppe von vier Kindern ist zu erwarten, dass reale Bedingungen im Rahmen eines Kindergartenalltages die Durchführung maßgeblich beeinflussen. Das bedeutet, eine Realisierung in Form eines Einzelinterviews mit Laborbedingungen, wäre nicht möglich. Vielmehr ginge es dann darum, einen Austausch zwischen pädagogischer Fachkraft bzw. Testleiter und Kindern möglichst zielorientiert zu planen und durchzuführen. Ein Kind sollte nach dem Training im Inhaltsgebiet *Elastizität und Plastizität* verstehen, dass es für das Überprüfen einer Vermutung entsprechend der vorgegebenen Formulierung sein Augenmerk erst auf die Eigenschaft und anschließend auf das Sprungverhalten richten muss. Dazu ist eine entsprechende Gesprächslenkung durch die pädagogische Fachkraft bzw. den Testleiter im Sinne von *Sustained Shared Thinking* (Siraj-Blatchford et al., 2002) relevant.

In der vorliegenden Studie dienten Laborbedingungen dazu, die Wirksamkeit der beiden Trainingsmaßnahmen im Vergleich zur Kontrollgruppe zu überprüfen. Bei einer Erprobung der Scaffolding-Maßnahmen in der Praxis im Rahmen einer Kleingruppe müsste die



Zusammensetzung der Gruppen im Hinblick auf relevante Aspekte (kognitive Voraussetzungen, Alter, Entwicklung sprachliche Fähigkeiten) im Vorfeld der Erhebungen berücksichtigt werden.

Zusammenfassend kann festgehalten werden, dass die Durchführung der Trainingsmaßnahmen im Rahmen der Trainingsstudie unter Laborbedingungen erfolgreich war. Sie hat gezeigt, dass vor allem jüngere Kinder mit wenig Vorwissen in Bezug auf das Überprüfen von Vermutungen die Modellierung durch den Erwachsenen benötigen, umso hilfreiche Denkanstöße zu erhalten. Mit zunehmendem Verstehen der einzelnen Denkschritte kann die intensive Unterstützung durch Modellierung zurückgenommen werden und mit Scaffolding durch adaptives Nachfragen ersetzt werden.

## 10 Literaturverzeichnis

Adams (2002). Scaling PISA cognitive data. In: R. Adams, M. Wu (Hrsg.). PISA 2000 technical report. Paris: OECD, S. 99-108.

Alloway, T. P., Gathercole, S. E., Willis, C., Adams, A. M. (2004). A structural analysis of working memory and related cognitive skills in young children. *Journal of Experimental Child Psychology*, 87, S. 85-106.

Anders, Y., Hardy, I., Pauen, S., Steffensky, M. (2013). Zieldimensionen früher naturwissenschaftlicher Bildung im Kita-Alter und ihre Messung. In: Stiftung „Haus der kleinen Forscher“ (Hrsg.). Wissenschaftliche Untersuchung zur Arbeit der Stiftung „Haus der kleinen Forscher“ (Bd. 5). Schaffhausen: Schubi-Lernmedien AG.

Anders, Y., Roßbach, H.-G., Weinert, S., Ebert, S., Kuger, S., Lehrl, S., Maurice, J. v. (2012). Home and preschool learning environments and their relations to the development of early numeracy skills. *Early Childhood Research Quarterly*, 27, (2), S. 231-244.

Archibald, S.J., Kerns, K. A. (1999). Identification and Description of New Tests of Executive Functioning. *Child Neuropsychology*, 5, S. 115-129.

Artelt, C., Demmrich, A., Baumert, J. (2001). Selbstreguliertes Lernen. In: PISA 2000. Springer, S. 271-298.

- Aufschnaiter, C. von, Riemann, T. (2005). Experimente im naturwissenschaftlichen Unterricht, *Lernchancen*, 47, S. 6-10.
- Aufschnaiter, C., Sibel, E., Osborne, J., Simon, S. (2008). Arguing to learn and learning to argue. Case studies of how students` argumentation relates to their scientific knowledge. *Journal of Research in Science Teaching*, 45 (1). S. 101-131.
- Baddeley, A. D. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, 4, S. 417-423.
- Baddeley, A. D., Allen, R., Hitch, G. J. (2011). Binding in visual working memory: The role of the episodic buffer. *Neuropsychologia*, 49, S. 1393-1400.
- Baddeley, A. D., Hitch, G. (1974). Working Memory. In G.A. Bower (Hrsg.). *The psychology of learning and motivation*. New York: Academic Press.
- Barrouillet, P., Gauffroy, C., Lecas, J.-F. (2008). Mental models and the suppositional account of conditionals. *Psychological Review*, 115 (3), S. 760–771.
- Barrouillet, P., Grosset, N., Lecas, J. F. (2000). Conditional reasoning by mental models: Chronometric and developmental evidence. *Cognition*, 75, S. 237–266.
- Barrouillet, P., Lecas, J.-F. (1999). Mental Models in Conditional Reasoning and Working Memory. *Thinking & Reasoning*, 5 (4), S. 289–302.
- Barrouillet, P., Lecas, J.F. (2002). Content an context effects in children`s and adults conditional reasoning. *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, 55 (3), S. 839-854.
- Barrouillet, P., Lecas, J. F. (2008). Mental Models and the Suppositional Account of Conditionals. *Psychological Review*, 115 (3), S. 760-772.
- Bassok, D., French, D., Fuller, B., Sharon, L.,K. (2008). Do child care centers benefit poor children after entry? *Journal of Early Childhood Research*, 3, S. 211-231.
- BMBF (2004) OECD-Veröffentlichung „Bildung auf einen Blick“ Wesentliche Aussagen in der Ausgabe 200 BMFSJ (2003)  
<http://www.bmfsfj.de/BMFSFJ/Service/publikationen,did=15934.html>
- Bodrova, E., Leong, D. (2007). *Tools of The Mind*. 2<sup>nd</sup>. Edition. Columbus: Prentice Hall.

- Bond, T., G., Fox, C., M. (2007). Applying the Rasch Model. New Jersey: Routledge.
- Bos, W., Bensen, M., Baumert, J., Prenzel, M., Selter, C., Walther, G. (Hrsg.) (2008). TIMSS 2007. Mathematische und naturwissenschaftliche Kompetenzen von Grundschulkindern in Deutschland im internationalen Vergleich. [https://www.phil-fak.uni-duesdorf.de/fileadmin/Redaktion/Institute/Erziehungswissenschaften/Abteilungen/Bildungsforschung/Lehrveranstaltungen/2009\\_WS/VL\\_\\_Sozialwissenschaftliche\\_BF/TIMSS\\_2007\\_Pressemappe.pdf](https://www.phil-fak.uni-duesdorf.de/fileadmin/Redaktion/Institute/Erziehungswissenschaften/Abteilungen/Bildungsforschung/Lehrveranstaltungen/2009_WS/VL__Sozialwissenschaftliche_BF/TIMSS_2007_Pressemappe.pdf) Zugriff am 03.10.2015.
- Bowman, B. T., Donovan, M. S., Burns, M. S. (2001). Eager to learn: Educating our preschoolers. Washington, DC: National Academic Press.
- Breddermann, T. (1983). Effects of Activity-Based elementary Science on Students Outcomes: A quantitative synthesis. In Review of Educational Research, 53 (4), S. 499-518
- Brophy, J. (1999). Toward a model of the value aspects of motivation in Education. *Educational Psychologist*, 34 (2), S. 75–85.
- Brown, A. L. (1989). Analogical learning and transfer: What develops? In: S. Vosniadou (Hrsg.). Similarity and analogical reasoning. S. 369-412. Cambridge: Cambridge University Press.
- Brown, N., Nagashima, S., Fu, A., Timms, M., Wilson, M. (2010). A Framework for Analyzing Scientific Reasoning in Assessments. *Educational Assessment*, 15, S. 142–174,
- Bullock, M., Sodian, B., Koerber, S. (2009) Doing experiments and Understanding Science. Development of Scientific Reasoning from childhood to Adulthood. In: Schneider, W. Bullock (Eds.) Human Development from Early Childhood to Early Adulthood. New York: Psychology Press, S. 173-197.
- Bullock, M., Ziegler, A. (1999). Scientific Reasoning. Developmental and individual differences. In: F. E. Weinert E. & W. Schneider W. (Hrsg.). Entwicklung des wissenschaftlichen Denkens von 3-12. New York: Cambridge University Press., S. 38–54.
- Bundesministerium für Bildung und Forschung (Hrsg.). (2012). PISA-Programme for International Student Assessment. <https://www.bmbf.de/de/pisa-programme-for-international-student-assessment-81.html> Zugriff 14.1.2014.

- Burger, K. (2010). How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. *Early Childhood Research Quarterly*, 25, S. 140-165.
- Butts, D., Hofman, H.M., Anderson, M. (1994) Is direct experience enough? A study of young children`s views of sounds. *Journal of Elementary Science Education*, 6(1), S. 1-16.
- Bybee, R. W. (1997). Toward an understanding of scientific literacy. In. W. Gräber und C. Bolte (Hrsg.). *Scientific literacy*. Kiel: IPN, S. 37–68.
- Bybee, R. W., McCrae, B., Laurie, R. (2009). PISA 2006. An assessment of scientific literacy. *Journal of Research in Science Teaching*, 46 (8), S. 865–883.
- Carey, S., Evans, R., Honda M., Unger J. E. (1989). An experiment is when you try it and see if it works. A study junior high school students understanding of the construction of scientific knowledge. *International Journal of Science Education*, 11, S. 514–529.
- Carle, U., Welzel, (2007). *Vorschulische Bildung im Kindergarten*. Verlag für Sozialwissenschaften: Wiesbaden.
- Carlson, S. M. (2005). Developmentally sensitive measures of executive function in preschool children. *Developmental Neuropsychology*, 28, S. 595-616.
- Chen, Z., Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development*, 70 (5), S. 1098-1120.
- Chi, M.T., Feltovich, P., J., Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive science*. 4, S. 3-45.
- Chinn, C. Brewer (2001). Models of Data: A theory of How People Evaluate Data. *Cognition and Instruction*, S. 323-393.
- Clark, K., Graves, M. (2005). Scaffolding students` comprehension of text. *Reading Teacher*, 58, S. 570 - 580.
- Colberg-Schrader, H. (1998). Kindergarten-für Kinderleben und Treffpunkt für Eltern. Zur Qualität von Kindergärten. In. Fhtenakis, W.F., Textor, M.R. (Hrsg.). *Qualität von Kinderbetreuung*. Weinheim: Beltz, S. 86-97.

- Collins, A., Brown, J.S., Newman, S.E. (1989). Cognitive Apprenticeship: Teaching the Crafts of Reading, Writing and Mathematics. In. L. Resnick. Knowing, Learning, and instruction. New York: Lawrence Erlbaum.
- Daniel, D., B., Klaczyński, P. A. (2006). Developmental and Individual Differences in Conditional Reasoning: Effects of Logic Instructions and Alternative Antecedents. *Child Development*, 77, S. 339-354.
- Decristan, J., Hondrich, L., Büttner, G., Hertel, S., Klieme, E., Kunter, M., Lühken, A., Adl-Amini, K., Djakovic, S.- K., Mannel, S., Naumann, A., Hardy, I. (2015). Impact of Additional Guidance in Science Education on Primary Students` Conceptual Understanding. *The Journal of Educational Research*, S. 1-13.
- De Jong, T. (2006). Scaffolds for computer simulation based scientific discovery learning. In. J. Elen, R.E. Clark (Hrsg.). Dealing with Complexity in Learning Environments London: Elsevier Science Publishers. S. 107-128.
- Diamond, A. (2002). Normal development of prefrontal cortex from birth to young adulthood: Cognitive functions, anatomy and biochemistry. In. D. Stuss; R. Knight (Hrsg.). Principles of frontal lobe function. New York: Oxford University Press, S. 466-503.
- Diamond, A., Prevor, M.,B., Callender, G., Druin, P. (1997). Prefrontal Cortex Cognitive Deficits in Children Treated Early and Continuously For PKU. Monographs of the Society for Research. *Child Development*, 252 (62), No 4, S. 1-208.
- Diamond, A., Carlson, S.M., Beck, D.M. (2005). Preschool children`s performance in task switching on the dimensional change card sort task: Separating the dimensions aids the ability to switch. *Developmental Neuropsychology*, 28, S. 689-729.
- Dohmen, D. (2010). Die ökonomischen Folgen der Bildungsarmut. Verlag für Sozialwissenschaften: Wiesbaden.
- Dowsett, S., Livesey, D. (2000). The development of inhibitory control in preschool children: Effects of executive skills training. *Developmental Psychology*, 36, S. 161-174.
- Duit, R. Treagust, D. F. (2008). Conceptual change: A discussion of theoretical, methodological and practical challenges for science education. *Cultural Studies of Science Education*, 3. S. 297-328.

- Driver, R., Leach, J., Millar, R. (1996). *Young People`s images of science*. Open University Press.
- ECCE-Study Group (1997). *European Child Care and Education Study: Cross National Analyses of the quality and effects of early childhood programmes on children`s development*: Free University Berlin.
- Einsiedler, W., Hardy, I. (2010). Kognitive Strukturierung im Unterricht-Einführung und Begriffsbestimmungen. *Unterrichtswissenschaft*, 38, (3), S. 194–209.
- Erziehungs- und Bildungsplan (2007). Hessisches Sozialministerium Hessisches Kultusministerium. mww. druck und so. GmbH: Mainz-Kastell.
- Evans, J., St. B. T. (2008). Dual-processing accounts of reasoning, judgement and social cognition. *Annual Review of Psychology*, 59, S. 255-378
- Evans, J. St. B. T., Over, D. E., Handley, S. H. (2003). A theory of hypothetical thinking. In D. Hardman & L. Maachi (Hrsg.). *Thinking: Psychological perspectives on reasoning, judgement and decision making*. Chichester: Wiley. S. 3–22.
- French L. (2004). Science as the center of a coherent, integrated early childhood curriculum. *Early Childhood Research Quarterly* 19, Bd. 19, S. 138–149.
- Field, A. (2009). *Discovering Statistics Using SPSS*. London, Thousand Oaks, New Delhi, Singapore: Sage Publications Ltd.
- Fitzpatrick, D., Mc Kinnen, R., D., Blair, C., Willoughby, M. T. (2014). Do preschool executive function skills explain the school readiness gap between advantaged and disadvantaged children? *Learning and Instruction* 30, S. 25 – 31.
- Friedman, N., Miyake, A. (2004). The Relations among Inhibition and Interference Control Functions: A Latent-Variable Analysis. *Journal of Experimental Psychology*, 133 (1), S. 101-135.
- Fröhlich-Gildhoff, K., Nentwig-Gesemann, I., Pietsch, S. (2011). *Kompetenzorientierung in der Qualifizierung fröhpädagogischer Fachkräfte*. Weiterbildungsinitiative Fröhpädagogische Fachkräfte-Herausforderungen. München.

- Furtak, E.M., Hardy, I., Beinbrech, C. Shavelson, R., J., & Shemwell, J.,T., (2010). A Framework For Analyzing Evidence-Based Reasoning in Science Classroom Discourse. *Educational Assessment, 15*, S. 175-196.
- Garon , N., Bryson, S., Smith, I. M. (2008). Executive Function in Preschoolers: A Review Using an Integrative Framework. *Psychological Bulletin, 134*. 1, S. 31–60.
- Gathercole, S. E., Pickering, S. J., Ambridge, B., Wearing, H. (2004). The Structure of Working Memory from 4 to 15 Years of Age. *Developmental Psychology, 40*, S.177-190.
- Gauffroy, C., Barrouillet, P. (2009). Heuristic and analytical processes in mental models for conditionals: An integrative developmental theory. *Developmental Review, 29*, S. 249-282.
- Gauffroy, C., Barrouillet, P. (2011). The Primacy of Thinking about Possibilities in the Development of Reasoning. *Developmental Psychology, 47*, (4), S. 1000-1011.
- Gelman, R., Brennemann, K. (2004). Science learning pathways for young children, *Early Childhood Research Quarterly, 19*, S. 150–158.
- Gelman, R., Bullock, S., Meck, D. (1980). Preschoolers` Understanding of Simple Object Transformations. *Child Development, 51*, S. 691-699.
- Gelman, S. A., Markman, E. M. (1986). Categories and inductions in young children. *Cognitive Science, 23*, S.183-209.
- Gerstenmaier, J., Mandl , H. (1995). Wissenserwerb unter konstruktivistischer Perspektive. *Zeitschrift für Pädagogik , 41* (6), S. 867–888.
- GDSU (2002). Gesellschaft für Didaktik des Sachunterrichts. Perspektivrahmen Sachunterricht [http://www.gdsu.de/wb/media/upload/pr\\_gdsu\\_2002.pdf](http://www.gdsu.de/wb/media/upload/pr_gdsu_2002.pdf). Zugriff am 29.5.2015.
- Gold, A., Dubowy, M. (2013). Frühe Bildung. Lernförderung im Elementarbereich. Stuttgart: Kohlhammer.
- Goswami, U. (2001). So denken Kinder. Bern: Huber.
- Gropen J., Clark-Ciarelli N. Hoisington, C., Stacy, B., Ehrlich, S.B. (2011). The Importance of Executive Function in Early Science Education. *Child Development Perspectives, 5* (4), S. 298–304.

- Grygier, P. (2008). Wissenschaftsverständnis von Grundschulern im Sachunterricht. Bad Heilbrunn: Klinkhardt.
- Hardy, I. (2012). Kognitive Strukturierung-Empirische Zugänge zu einem heterogenen Konstrukt der Unterrichtsforschung. In F. Hellmich, S. Förster, F. Hoya, Jahrbuch der Grundschulforschung, S.238-241. Paderborn: Springer.
- Hardy, I., Jonen, A., Möller, K., Stern, E. (2006). Effects of instructional support within constructivist learning environments for elementary school students` understanding of Floating and Sinking. *Journal of Educational Psychology*, 98 (2), S. 307-326.
- Hardy, I., Kleickmann, T., Koerber, S.; Mayer, D., Möller, K.; Pollmeier, J.; Schwippert, K.; Sodian, B. (2010 a). Die Modellierung naturwissenschaftlicher Kompetenz im Grundschulalter. Projekt Science-P. *Zeitschrift für Pädagogik*, 56, Beiheft 56, S. 115-125.
- Hardy, I., Kloetzer, Möller, K., Sodian, B. (2010 b). The Analysis of Classroom Discourse: Elementary School Science Curricula Advancing Reasoning With *Evidence*. *Educational Assessment*, 15 (3-4), S. 1997-221.
- Hardy, I., Mannel, S., Sauer, S. (2015) Inklusive sprachliche Bildung im Kindergarten: Gestaltungsmöglichkeiten im Kontext der Naturwissenschaften. In C. Huf, I. Schnell (Hrsg.). *Inklusive Bildung im Elementar-und Primarbereich*. Stuttgart: Kohlhammer, S. 163-179.
- Hardy, I., Steffensky, M. (2014). Prozessqualität im Kindergarten: Eine domänenspezifische Perspektive. *Unterrichtswissenschaft. Zeitschrift für Lernforschung* ,42(2), S. 101-116.
- Hardy, I., Stern, E. (2011). Visuelle Repräsentationen der Dichte. Auswirkungen auf die konzeptuelle Umstrukturierung bei Grundschulkindern. *Unterrichtswissenschaft*, 39 (1), S. 35-48.
- Harlen, W., Qualter, A. (2009). *The Teaching of Science in Primary Schools*. London: Routledge.
- Hasselhorn, M., Seidler, B., Brandler, U.; Körner, K. (2000). Ist das „Nachsprechen von Kunstwörtern“ für die Entwicklungsdiagnostik des phonologischen Gedächtnisses geeignet? In M. Hasselhorn, W. Schneider, H. Marx (Hrsg.). *Diagnostik von Leserechtschreibschwierigkeiten*. Jahrbuch der pädagogisch-psychologischen Diagnostik. Tests und Trends. Göttingen: Hogrefe. S. 119-133.



- Hessisches Sozialministerium, Hessisches Kultusministerium (2007). Bildung von Anfang an. Bildungs- und Erziehungsplan für Kinder von 0-10 Jahren in Hessen. Zugriff am 22.01.2014. [https://hsm.hessen.de/sites/default/files/HSM/2012-08-00\\_bildungs-und-erziehungsplan.pdf](https://hsm.hessen.de/sites/default/files/HSM/2012-08-00_bildungs-und-erziehungsplan.pdf).
- Hofstein, A., Lunetta, V., N. (2004). The Laboratory in Science Education: Foundations for the Twenty-First Century. *Science Education*, 88, 1, S. 28-54.
- Hogan K., Pressley, M. (1997). Scaffolding Student Learning. Cambridge, Massachusetts: Brookline Books.
- Holyoak, K., Thagard, P., (1989). Analogical Mapping by constraint satisfaction. *Cognitive Science*, 13, S. 295-355.
- Hopf, M. (2011). Sustained Shared Thinking im frühen naturwissenschaftlich-technischen Lernen. Münster: Waxmann.
- Hughes, C. (1998). Executive function in preschoolers: Link with theory of mind and verbal ability. *British Journal of Developmental Psychology*, 16, S. 233-253.
- Inhelder, B., Piaget, J. (1958). The growth of logical thinking from childhood to adolescence. New York: Basic.
- Janssen, J., Laatz, W. (2013). Statistische Datenanalyse. Berlin: Springer.
- Jansen, H., Mannhaupt, G., Marx, H., Skowronek, H. (1999). Bielefelder Screening zur Früherkennung zu Lese-Rechtschreibschwierigkeiten. Göttingen: Hogrefe.
- Johnson-Laird, P. N., Byrne, R., M.J. (1991). Deduction. Hillsdale, NJ: Erlbaum.
- Johnson-Laird, P. N., Byrne, R. M.J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, 109, S. 646-678.
- Kauertz, A. (2012). Naturwissenschaftliches Denken. In: Kucharz, D. Elementarbildung. Weinheim: Beltz, S. 86-121.
- Kempert, S., Hardy, I. (2012). Effekte von früher Zweisprachigkeit auf das deduktive Schließen im Grundschulalter, *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*. 44 (1), S. 27-39.

- Klahr, D., Dunbar, K. (1988). Dual Space Search During Scientific Reasoning. *Cognitive Science*, 12, S. 1-48.
- Klahr, D. (2000). Exploring Science. The cognition and development of discovery processes. Cambridge: MIT Press.
- Klahr, D., Chen, Z. (2003). Overcoming the positive-capture strategy in Young Children: learning about indeterminacy. *Child Development*, 74, 5, S. 1255–1277.
- Klahr, D., Nigam, M. (2004). The Equivalence of Learning Paths in Early Science. *Psychological Science*, 15, 10, S. 661–667.
- Kleickmann, T, Tröbst, S, Jonen, A., Vehmeyer, J., Möller, K. (2016). The effects of Expert Scaffolding in Elementary Science Professional Development on Teachers` Beliefs and Motivations, Instructional Practices and Student Achievement. *Journal of Educational Psychology*, Vol. 108 (1), S.21-42.
- Klein, E. R., Hammrich, Penny L., Bloom, S., Ragin's, A. (2000). Language development and science inquiry. The Head Start on Science and Communication program. *Early Childhood Research & Practice*. S.63-78.
- Klieme, E., Leutner, D. (2006). Kompetenzmodelle zur Erfassung individueller Lernergebnisse und zur Bilanzierung von Bildungsprozessen. Beschreibung eines neu eingerichteten Schwerpunktprogramms der DFG. *Zeitschrift für Pädagogik*, 52 (6), S. 876-903.
- Koerber, S., Kropf, N., Mayer, D., Sodian, B. & Schwippert, K. (2009). Entwicklung naturwissenschaftlicher Kompetenz in der Grundschule (Science-P): Wissen über Naturwissenschaften. In. C. Röhner, M. Hopf, C. Henrichwark (Hrsg.). *Europäisierung der Bildung. Konsequenzen und Herausforderungen für die Grundschulpädagogik*, Jahrbuch Grundschulforschung Band 13 (S. 194-198). Wiesbaden: Verlag für Sozialwissenschaften.
- Koerber, S., Sodian, B.; Kropf, N., Mayer, D.; Schwippert, K. (2011). Die Entwicklung des wissenschaftlichen Denkens im Grundschulalter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 43, 1, S. 16–21.
- Koerber, S., Sodian, B., Thoermer C., Nett, U. (2005). Scientific Reasoning in Young Children: Preschoolers Ability to Evaluate Covariation Evidence. *Swiss Journal of Psychology*, 64, (3), S. 141–152.

- Kolen, M. J., Brennan, R.L. (2004). *Test Equating, Scaling and Linking. Methods and Practices*. New York: Springer.
- Koslowski, B., Marasia, J., Chelenza, M., Dublin, R. (2008). Information becomes evidence when an explanation can incorporate it into a causal framework. *Cognitive Development*, 23(4), S. 472-487.
- Krammer, K. (2010). Individuelle Unterstützung im Unterricht mit 4-bis 8-jährigen Kindern. In: M. Leuchter (Hrsg.). *Didaktik für die ersten Bildungsjahre*. S. 112–127. Zug: Kallmeyer, Klett.
- Kucharz, D., Mackowiak, K., Zirolì, S., Kauertz, A., Mathgreb-Schnierer, M. (2014). *Professionelles Handeln im Elementarbereich (Primel). Eine deutsch-schweizer Videostudie*. Göttingen: Waxmann.
- Kuger, S.; Kluczniok, K. (2008). Prozessqualität im Kindergarten-Konzept in Institutionen. *Zeitschrift für Erziehungswissenschaft*, Sonderheft 11, S. 159-178.
- Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review*, 96, S. 674–689.
- Kuhn, D. (2002). What is scientific thinking and how does it develop? In: U. Goswami (Hrsg.) *Handbook of Childhood cognitive development*. Oxford: Blackwall. S. 371-393.
- Kuhn, D. (2010). What is Scientific Thinking and How Does it Develop? In: U. Goswami (Hrsg.) *Handbook of Childhood Cognitive Development*, S. 2-23. Malden: Blackwell.
- Kuhn, D. , Amsel, E., O` Laughlin, M. (1988). *The development of scientific thinking skills*. Orlando: FL, Academic Press.
- Kuhn, D., Dean, D. (2004). Metacognition: A Bridge between Cognitive Psychology and Educational Practice. *Developmental Psychology*, 43(4), S. 269-273.
- Kuhn, D., Franklin, S. (2006). The second decade. What develops (and how). In: D. Kuhn, R.S. Siegler (Hrsg.). *Handbook of child psychology ( 2). Cognition, perception and language*, Hoboken, NJ: Wiley, S. 953- 993.
- Kuhn, D., Pearsall, S. (2000). Developmental Origins of Scientific Thinking. *Journal of Cognition and Development*, 1, S. 113–129.

- Kuhn, D., Phelps, E. (1982). The Development of Problem-Solving Strategies child Development, *Advances in Child Development and Behavior*, 17, S. 1–44.
- Kultusministerkonferenz (2004, 2009) <http://www.kmk.org/bildung-schule/qualitaetssicherung-in-schulen/bildungsstandards/>. Zugriff am 14.01.2014.
- Lawson, A., E., (2003). The nature and development of hypothetico-predictive argumentation with implications for science teaching. *International Journal of Science Education*, 25(11), S. 1387-1408.
- Lawson, A., E., (2005). What is the Role of Induction and Deduction in Reasoning and Scientific Inquiry? *Journal of Research in Science Teaching*, 42 (6), S. 716-740.
- Lawson A., E. (2009). Basic Inferences of Scientific Reasoning, Argumentation, and Discovery. *Science Education*, 94, S. 336–364.
- Lawson, A., E. (2010). Basic Inferences of scientific reasoning, argumentation and discovery. *Science Education*. 94(2), S. 336-364.
- Leevers, J., Harris, J. (1999). Persisting Effects of Instruction on Young Children's Syllogistic Reasoning with Incongruent and Abstract Premises. *Thinking and Reasoning*, 5(2), S.145-173.
- Lehto, J. E., Juujärvi, P., Kooistra, L. Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, 21, 59–80.
- Leuchter, M., Saalbach, H., Hardy, I. (2011). Förderung naturwissenschaftlichen Verständnisses von Kindern in der Schuleingangsstufe. Empirische Forschung zur Qualität des (naturwissenschaftlichen) Lernens und Lehrens in der Schuleingangsstufe. In: F. Vogt, M. Leuchter, A. Tettenborn, U. Hottinger, M. Jäger, E. Wannack. *Entwicklung und Lernen junger Kinder*. Münster: Waxman.
- Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: a critical Appraisal. *Learning and Instruction*, 11(4), S. 357-380.
- Markovits, H. (2006). Making conditional inferences: the interplay between knowledge and logic. In L. Smith and J. Voneche (Hrsg.). *Norms in Human Development*, Cambridge: Cambridge University Press.

- Markovits, H., Barrouillet, P. (2002). The Development of Conditional Reasoning: A Mental Model Account. *Developmental Review* 22, S. 5-36.
- Mayer, R. E. (2004). Should there be a three -strikes rule against pure discovery learning. *American Psychologist*. 59, S. 14–19.
- Mayer, D., Sodian, B., Koerber, S., Schwippert, K. (2014). Scientific reasoning in elementary school children: Assessment and relation with cognitive abilities. *Learning and Instruction*, 29 (2), S. 43-55
- Mc Neill, K. L., Lizotte, D. (2006). Supporting students` construction of scientific explanations. *The international journal of higher education in the social sciences*. In: *The Journal of the Learning Sciences*, S. Online-Ressource. Online verfügbar unter <http://journals.berghahnbooks.com/ltss/>Zugriff am 07.01.2014; <http://www.bibliothek.uni-regensburg.de/ezeit/?2413663>.
- Melhuish, E. (2013). Research on Early Childhood Education in the UK. In. M Stamm und D. Edelmann (Hrsg.). *Handbuch frühkindliche Bildungsforschung*. Wiesbaden: Springer, S. 617–632.
- Mercer, N., Dawes, L., Wegerif, R. (2004). Reasoning as a scientist: ways of helping children to use language to learn science. *British Educational Research Journal*, 30b(3), S. 359-377.
- Miyake, A., Friedman, N., P., Emerson, M., J., Witzki, A.,H., Howerter, A. (2000). The Unity and Diversity of Executive Functions and their contributions to complex “Frontal Lobe” Tasks: A Latent Variable Analysis *Cognitive Psychology*, 41, S. 49-100.
- Miyake, A., Friedman, N. P., Rettinger, D. A., Priti, S., Hegarty, M. (2001). How are visuospatial working memory, executive functioning and spatial abilities related? A latent variable analysis. *Journal of Experimental Psychology*, 130 (4), S. 621-654.
- Möller, K. (2001). Lernern im Vorfeld der Naturwissenschaften- Zielsetzungen und Forschungsergebnisse. In. W. Köhnlein, H. Schreier. *Innovation Sachunterricht. Befragung der Anfänge nach zukunftsfähigen Beständen*. S. 275-294. Bad Heilbrunn: Klinkhardt, 4.
- Möller, K. (2004). Naturwissenschaftliches Lernen in der Grundschule. Welche Kompetenzen brauchen Grundschullehrkräfte? In. H. Merkens (Hrsg.). *Lehrerbildung: IGLU und die Folgen*. Schriften der deutschen Gesellschaft. Opladen: Leske+Budrich, S. 65-84.

- Möller, K. (2009). Was lernen Kinder über Naturwissenschaften im Elementar- und Primarbereich? – Einige kritische Bemerkungen. In L. Lauterbach, H. Giest, B. Marquardt-Mau (Hrsg.). *Lernen und kindliche Entwicklung. Elementarbildung und Sachunterricht*. S. 165-172. Bad Heilbrunn: Klinkhardt.
- Möller, K., Steffensky, M. (2010). Naturwissenschaftliches Lernen im Unterricht mit 4- bis 8-jährigen Kindern. In M. Leuchter. *Didaktik der ersten Bildungsjahre*, S.163-178. Zug: Klett und Balmer Verlag.
- Moosbrugger, H., Kevala, A. (2007). *Testtheorie und Fragebogenkonstruktion*. Heidelberg: Springer.
- Murphy, N., Messer, D. (2000). Differential Benefits from Scaffolding and Children Working Alone. *Educational Psychology*, 20 (2), S. 17–31.
- Noble, K., G., Norman, F., M., Farah, J., F. (2005). Neurocognitive Correlates of socioeconomic status in kindergarten children. *Developmental Science*, 8, (1), S. 74-87.
- NICHD Early Child Care Research Network (2003). Does Quality of Child Care Affect Child Outcomes at Age of 4? *Developmental Psychology*, 39(3), S. 451–469.
- NICHD Early Child Care Research Network ECCRN (2005). A day in a third grade. Classroom quality, teacher and student behaviours. *Elementary School Journal*, 105 (3), S. 305-323.
- Norris, S. P., Philipps, L. M. (2003). How literacy in its fundamental sense is central to scientific literacy. *Science Education*, 87 (2), S.224–240.
- Oberhuemer, P. (2004). *Bildungskonzepte für die frühe Kindheit in internationaler Perspektive* In. W. E. Fhtenakis, P. Oberhuemer (Hrsg.). *Frühpädagogik international*. S. 359-383. Wiesbaden: Verlag für Sozialwissenschaften.
- OECD (Hrsg.). (2001). *Lernen für das Leben. Erste Ergebnisse von PISA 2000*. Paris: OECD. Zugriff am 03.10.2015. <http://www.oecd.org/pisa/>
- OECD (2004). *Starting Strong I*. Zugriff am 14.01.2014
- Pea, R. (2004). The Social and Technological Dimensions of Scaffolding and Related Theoretical Concepts for Learning, Education, and Human Activity. *The Journal of the Learning Sciences*. 13(3), S. 423-451.

- Perner, J. (1991). *Understanding the representation of the mind*. Harvard: MIT Press.
- Perner, J., Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences*, 3 (9), S. 337-344.
- Pennington, B., Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry*, 37 (1), S. 51-87.
- Prenzel, M., Artelt, C. (2007). *Die Ergebnisse der dritten internationalen Vergleichsstudie*. Münster: New York, München: Waxmann.
- Puntambekar, S., Hübscher, R. (2005). Tools for Scaffolding Students in a Complex Learning Environment. What have we gained and What have we missed? *Educational Psychologist*, 40, (1), S. 1–12.
- Rasch, G. (1960). *Studies in mathematical psychology: I. Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danmarks Pädagogisches Institut.
- Reiser, B. J. (2004). Scaffolding complex learning: The mechanisms of structuring and problematizing student work. *The Journal of the Learning Sciences*, 13 (3), S. 273–304.
- Robisch, C., Tröbst, S., Möller, K. (2012). Inhaltliches Vorwissen und inhaltsbezogene Schlussfolgerungen im naturwissenschaftlichen Sachunterricht. In F. Hellmich, S. Förster, F. Hoya (Hrsg.). *Bedingungen des Lehrens und Lernens in der Grundschule*. S. 245-248, *Jahrbuch Grundschulforschung*, 16. Wiesbaden: Springer Verlag.
- Robisch, C., Tröbst, S., Möller, K. (2014). Hypothesenbezogene Schlussfolgerungen im Grundschulalter fördern. *Zeitschrift für Grundschulforschung*, 7. S. 88 - 101.
- Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. New York, Oxford: Oxford University Press.
- Rost, J. (2004). *Testtheorie-Testkonstruktion*. Bern, Göttingen, Toronto, Seattle: Verlag Hans Huber.
- Roßbach, H. G., Weinert, S. (2008). Kindliche Kompetenzen im Elementarbereich. In *Bildungsforschung*, 24.
- Röthlisberger, M., Neuenschwander, R., Michel, E., Roebers, C.M. (2010). Exekutive Funktionen: Zugrundeliegende kognitive Prozesse und deren Korrelate bei Kindern im späten

- Vorschulalter. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 42 (2), S. 91–110.
- Ruffman, T., Perner, J., Olson, D. R., Doherty, M. (1993). Reflecting on scientific thinking. Children's understanding of the hypothesis-evidence relation. *Child Development*, 64, S. 1617–1636.
- Saalbach, H., Grabner, H. R., Stern, E. (2013). Lernen als kritischer Mechanismus geistiger Entwicklung: Kognitionspsychologische und neurowissenschaftliche Grundlagen in frühkindlicher Bildung. In: M. Stamm, D. Edelmann (Hrsg.) *Handbuch frühkindlicher Entwicklung*. Wiesbaden: Springer Verlag.
- Samarapungavan, A., Panayota, N., Helen, P. (2008). Learning science through Inquiry in Kindergarten, *Science Education*, 1, S. 1-41.
- Sammons, P., Sylva, K., Melhuish, E. C., Siraj-Blatchford, I., Taggart, B. and Elliot, K. (2004). The Effective Provision of Pre-School Education. EPPE-Project: Technical Paper 8a. Measuring the Impact of Pre-School on Children's Cognitive Progress over the Pre-School Period. London: DfES / Institute of Education, University of London.
- Schauble, L. (1996). The development of scientific reasoning in knowledge-rich contexts. *Developmental Psychology*, 32, S. 102 - 119.
- Schmid, C., Zoelch, C., Roebers, M. (2008). Das Arbeitsgedächtnis von 4-bis 5-jährigen Kindern. Theoretische und empirische Analyse seiner Funktionen. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 40 (1), S. 2-12.
- Schwarz, C., Reiser, B., Acher, A., Kenyon, L., Fortus, L. (2012) MoDeLS. In: A. Alonzo, A., W., Gotwals (Hrsg.). *Learning Progressions in Science*. Rotterdam: Sense Publishers, S. 101-138.
- Siegler, R. (1995). How does change occur: A Microgenetic Study of Number Conversation. *Cognitive Psychology*, 28, S. 225-273.
- Siegler, R. (2006). Microgenetic Analysis of Learning. In: W. Damon; R., M. Lerner (Hrsg.) *Handbook of Child Psychology* (2. Bd.). S. 465-505. Cognition, perception, and language. New York: Wiley.
- Siegler, R.S., Crowley, K. (1991). The Microgenetic Method. *American Psychologist*, 46 (6), S. 606-620.



- Siegler, R., S., Stern, E. (1998). Conscious and Unconscious Strategy Discoveries: A Microgenetic Analysis. *Journal of Experimental Psychology*, 1127, (4), S. 377-397.
- Simons, K., D., Klein, J., D. (2007). The impact of scaffolding and student achievement levels in problem-based learning environment. *Instructional Science*, 35, S. 41-72.
- Siraj-Blatchford, I., S., Manni, L. (2008). Would you like to tidy up now? An analysis of adult questioning in the English Foundation Stage. *Early Years*, 28.(1). S. 1- 22.
- Siraj-Blatchford, I., S., Moriaty, V. (2004). Pädagogische Wirksamkeit in der Erziehung. In.. W. Fhtenakis, P., Oberhuemer. *Frühpädagogik International*. Verlag für Sozialwissenschaften. S. 87-104.
- Siraj-Blatchford, I., S. Sylva, K. (2012). Researching Pedagogy in English Pre-Schools. *Early Childhood Education*, 2, S. 145–163.
- Siraj-Blatchford, I., Sylva, K., Muttock, S., Gilden, R., Bell. D. (2002). *Researching Effective Pedagogy in the Early Years*. Norwich: DfES.
- Sodian, B. (1998). Entwicklung bereichsspezifischen Wissens. In: R. Oerter, L. Montada. *Entwicklungspsychologie*, Weinheim: Beltz, S. 622–650.
- Sodian, B., Hülsken, C. (2005). The developmental Relation of Theory of Mind and Executive Functions: A Study of Advanced Theory of Mind Abilities in Children With Attention Deficit Hyperactivity Disorder. In. W. Schneider, R. Schumann-Hengsteler, B. Sodian. *Young children`s cognitive development. Interrelationships among executive functioning, working memory, verbal ability and theory of mind*. S. 175-184. Mahwah, NJ: Lawrence Erlbaum Associates.
- Sodian, B., Jonen, A, Thoermer, C., Kircher, E. (2006). Die Natur der Naturwissenschaften verstehen. Implementierung wissenschaftstheoretischen Unterrichts in der Grundschule. In. M. Prenzel, L.A. Allolio-Näcke (Hrsg.). *Untersuchungen zur Bildungsqualität von Schule*, Münster: Waxmann, S. 147–193.
- Sodian, B.; Koerber, S. (2011). Hypothesenprüfung und Evidenzevaluation im Grundschulalter. *Unterrichtswissenschaft, Zeitschrift für Lernforschung*. 39 (1), S. 21-34.

- Sodian, B.; Mayer, D. (2013). Entwicklung des wissenschaftlichen Denkens im Vor- und Grundschulalter. In: M. Stamm und D. Edelman (Hrsg.). Handbuch frühkindliche Bildungsforschung. Wiesbaden: Springer, S. 617–632.
- Sodian, B.; Zaitchik, D., Carey, S. (1991). Young Children`s Differentiation of Hypothetical Beliefs from Evidence. *Child Development*, 62, S. 753–762.
- Spreckelsen, K. (2007). Anschlussfähiges Wissen und Können grundlegen. Handbuch Didaktik des Sachunterrichts. Bad Heilbrunn: Klinkhardt, S. 123-127.
- Stamm, M. (2010). Frühkindliche Bildung, Betreuung und Erziehung. Bern, Stuttgart, Wien: Haupt Verlag.
- Strand-Cary, M., Klahr, D. (2008). Developing elementary scientific skills. Instructional effectiveness and path independence. *Cognitive Development*, 23, S. 488-511.
- Steffensky, M. (2008). Einen naturwissenschaftlichen Blick entwickeln: Naturwissenschaftliches Lernen im Kindergarten. Naturwissenschaftliches Lernen im Kindergarten. In: F. Hellmich und M. Köster (Hrsg.). Vorschulische Bildungsprozess in Mathematik und Naturwissenschaften, S. 179–194. Bad Heilbrunn: Klinkhardt..
- Steffensky, M. (2012). Alltagssituationen und Experimente: Was sind geeignete naturwissenschaftliche Lerngelegenheiten für Kindergartenkinder? *Zeitschrift für Erziehungswissenschaft*, 1, S. 37–54.
- Steffensky, M. Lankes, E. M. (2011). Naturwissenschaftliches Lernen im Übergang vom Kindergarten zur Grundschule Handreichungen des Programms SINUS an Grundschulen, Kiel. S. 3-19
- Stephan-Gramberg, S., Hardy, I. (2012a). Die Koordination von Theorie und Evidenz bei Vorschulkindern: Ergebnisse einer Vorstudie zu Scaffolding-Maßnahmen. In: Jahrbuch Grundschulforschung. Wiesbaden: Verlag für Sozialwissenschaften, 17, S. 158-161.
- Stephan-Gramberg, S., Robisch, C., Löhner, K., Hardy, I., Tröbst, S., Möller, K. (2012b). Zusammenhänge zwischen Wissenschaftsverständnis und dem Umgang mit Evidenz bei naturwissenschaftlichen Inhalten in Kindergarten und Grundschule. *Vortrag auf der 21. Jahrestagung der Gesellschaft für Didaktik des Sachunterrichts*, Berlin.

- Sylva, K., Roy, D., Painter, M. (1980). Child watching at playgroup and nursery school. Blackwell: Oxford.
- Sylva, K., Melhuish, E., Sammons, P., Siraj-Blatchford, I., Taggart, B. (2004). The effective provision of pre-school Education (EPPE) project: <http://eprints.ioe.ac.uk/5309/1/sylva2004EPPEfinal.pdf>. Zugriff am 14.01.2014.
- Sylva, K., Siraj-Blatchford, I. Taggart, B., Sammons, P., Melhuish, E., Elliot, K., Totsika, V. (2006). Capturing quality in early childhood through environmental rating scales. *Early Childhood Research Quarterly*, 21(1), S. 76-92.
- Sylva, K.; Melhuish, E., Sammons, P., Siraj-Blatchford, I. (2012). Preschool quality and educational outcomes at age 11: Low quality has little benefit. *Journal of Early Childhood Research*, 9 (2), S. 109–124.
- Thiel, S. (1987). Wie springt ein Ball? *Grundschule*, 1, S. 18-23.
- Thoermer, C., Sodian, B. (2002). Science undergraduates' and graduates' epistemologies of science: The notion of interpretive frameworks. *New Ideas in Psychology*, 26, S. 263–283.
- Tietze, W. (1998). Wie gut sind unsere Kindergärten? Eine Untersuchung über die Qualität in deutschen Kindergärten. Neuwied: Luchterhand.
- Tietze, W. (2008). Qualitätssicherung im Elementarbereich. *Zeitschrift für Pädagogik, Beiheft 53*, S. 16-35.
- Toulmin, S. E. (1958). The uses of argument. Cambridge, UK: Cambridge University Press.
- Toulmin, S. E. (2003). The uses of argument. Cambridge, UK: Cambridge University Press. (Original work published 1958).
- Tröbst, S.; Hardy, I., Möller, K.(2011). Die Förderung deduktiver Schlussfolgerungen bei Grundschulkindern in naturwissenschaftlichen Kontexten. *Unterrichtswissenschaft. Zeitschrift für Lernforschung* , 39.Jg. 1Vj, S. 7-20.
- Tröbst, S., Robisch, C., Hardy, I., Möller, K., Stephan-Gramberg, S. (2012a). DFG-Bericht. (Manuskript).
- Tröbst, S., Robisch, C., Stephan-Gramberg S., Hardy, I., Möller K. (2012b). Förderung schlussfolgernden Denkens im Kindergarten und in der Grundschule bei

- naturwissenschaftlichen Inhalten. In: S. Förster & F. Hoya F. Hellmich. Jahrbuch Grundschulforschung. Bedingungen des Lehrens und Lernens in der GU. S. 237–240. Bilanz und Perspektiven, 16, Verlag für Sozialwissenschaften: Wiesbaden.
- Tytler, R., Peterson, S. (2005). A Longitudinal Study of Children`s Developing Knowledge and Reasoning in Science. *Science Education*, 35, S. 63-98.
- Tytler, R., Peterson, S. (2004). From “Try It and See” to strategic exploration: Characterizing young children's scientific reasoning. *Journal of Research in Science Teaching*, 41 (1), S. 94–118.
- Van de Pol, J., Volman, M., Beishuizen, J. (2010). Scaffolding in Teacher-Student Interaction: A Decade of Research. *Educational Psychological Review*, 22 (3), S. 271-297.
- Vergauwe, E., Gauffroy, C., Morsanyi, K. (2013). Chronometric evidence for the dual process mental model theory of conditional. *Journal of Cognitive Psychology*, 25, S. 174-182.
- von Maurice, J., Artelt, C., Blossfeld, H-P. Faust, G., Roßbach, H.-G. Weinert, S. (2007). Bildungsprozesse, Kompetenzentwicklung und Formation von Selektionsentscheidungen im Vor- und Grundschulalter: Überblick über die Erhebungen in den Längsschnitten BiKS 3-8 und BiKS 8-12 in den ersten beiden Projektjahren. <http://psydok.sulb.uni-saarland.de/volltexte/2007/1008/> – Zugriff am 30.05.2015.
- Völkel, P. (2002). Bildungsgeschichten aus der Kindertageseinrichtung. Wie Kinder miteinander und voneinander lernen. In: H., J., Laewen, B. Andres (Hrsg.). Forscher, Künstler, Konstrukteure. S. 70-87. Werkstattbuch zum Bildungsauftrag von Kindertageseinrichtungen. Neuwied: Luchterhand Verlag.
- Vosniadou, S. (2007). Conceptual Change and Education. *Human Development*. 50, S.47-54.
- Vygotsky L.S. (1978). Mind in society: the development of higher psychological process. Cambridge: Harvard University Press.
- Wannack, E. (2010). Bildung von 4- bis 8-jährigen Kindern: Grundlagen und Konzepte im Wandel. In: M. Leuchter (Hrsg.). Didaktik für die ersten Bildungsjahre. Unterricht mit 4- bis 8-jährigen Kindern. Zug: Klett und Balmer Verlag.
- Warm, T. (1989) Weighted Likelihood Estimation of Ability in Item Response Theory. *Psychometrika*, 54, 3, S.427-450.

- Wason, P. C. (1966). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, 20, S. 273-281.
- Wason, P.C., Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, 23, S. 63-71.
- Weiland, C., Yoshikawa, H. (2013). Impacts of a Prekindergarten Program on Children's Mathematics, Language Literacy, Executive Function, and Emotional Skills. *Child Development*, 84, 6, S. 2112-2130.
- Welzel-Breuer, M; Zimmermann, M. (2007). NFFK- ein Verfahren zur Erfassung und Förderung von naturwissenschaftlicher Frühförderkompetenz. in: Perspektiven zur pädagogischen Professionalisierung.73. Elementarpädagogik, S.15-30.
- Wilkening, F., Sodian, B. (2005). Special Issue on Scientific Reasoning in Young Children *Swiss Journal Psychology*, 64 (3), S. 137–139.
- Wilson, M. (2005). *Constructing Measures: An Item Response Modeling Approach*. Mahwah, NJ: Lawrence Erlbaum.
- Wimmer, H., Perner, J. (1983). Beliefs about beliefs: Representations and constraining function of wrong beliefs in young children's understanding of deception. *Cognition and Instruction*, Bd. 13, S. 103–128.
- Wood, D., Bruner, J. S., Ross, G. (1976). The Role of Tutoring in Problem Solving. *Child Psychology*, 17, S. 89-100.
- Wu, M. L., Adams, R. J (2007). *Applying the Rasch model to psycho-social measurement: A practical approach*. Melbourne: Educational Measurement Solutions.
- Wu, M. L., Adams, R. J., Wilson, M. R. (1998). *ACER ConQuest: Generalized Item response Modelling Software*. Hill Road, Camperwell, Melbourne: Australian Council for Educational Research.
- Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S. (2003). The development of executive function in early childhood. Monographs of the society for Research. *Childhood Development*, 68, S. 1-137.

Ziegler, T., Hardy, I. (2015) die Erfassung naturwissenschaftlicher Kompetenz im Vorschulalter: Ergebnisse einer Pilotierungsstudie. *Jahrbuch der Grundschulforschung*. S.211-215.

Zimmermann C. (2007). The development of scientific thinking skills in elementary and middle school. *Science Direct* 27, S. 172–223.

Zoelch, C., Seitz, K., Schumann-Hengsteler, R. (2005). From rag (bags) to riches:measuring the developing central executive. In. W. Schneider, R. Schumann-Hengsteler, B. Sodian. Young children`s cognitive development. Interrelationships among executive functioning, working memory, verbal ability and theory of mind. Mahwah, NJ: Lawrence Erlbaum Associates.

## 11 Tabellen- und Abbildungsverzeichnis

## Tabellen

Tabelle 1 <i>Studien zur Lerner-Unterstützung</i> .....	26
Tabelle 2 <i>Studien zur Koordination von Theorie und Evidenz</i> .....	54
Tabelle 3 <i>Wahrheitstabelle</i> .....	67
Tabelle 4 <i>Vorhersage der Schlussweisen</i> .....	70
Tabelle 5 <i>Übersicht über Studien zum schlussfolgernden Denken</i> .....	71
Tabelle 6 <i>Interpretationsniveaus beim Schlussfolgern</i> .....	74
Tabelle 7 <i>Erweiterte Interpretationsniveaus</i> .....	87
Tabelle 8 <i>Vier Vermutungen</i> .....	89
Tabelle 9 <i>Zuordnung der Antworten zu Symbolen</i> .....	90
Tabelle 10 <i>Kodierung nach Interpretationsniveau</i> .....	91
Tabelle 11 <i>Beispiel der Kodierung zu Frage 1 Kontext Marius</i> .....	95
Tabelle 12 <i>Mittelwerte (SD) im Wissenschaftsverständnis</i> .....	103
Tabelle 13 <i>Weitere Vermutungen für sieben bis zehnjährige Kinder</i> .....	105
Tabelle 14 <i>Stichprobe Querschnittstudie</i> .....	106
Tabelle 15 <i>Testdesign Querschnittstudie</i> .....	107
Tabelle 16 <i>Relative Verteilung der Evaluationen der präsentierten Ereignisse</i> .....	108
Tabelle 17 <i>Mittelwerte (Standardabweichung) der Tests im Überblick</i> .....	109
Tabelle 18 <i>Relative Häufigkeiten der Interpretationsniveaus</i> .....	111
Tabelle 19 <i>Mittelwerte (Standardabweichungen) der Interpretationsniveaus</i> .....	112
Tabelle 20 <i>Mittelwerte (Standardabweichungen) im Wissenschaftsverständnis</i> .....	113
Tabelle 21 <i>Item-Kennwerte Wissenschaftsverständnis</i> .....	115
Tabelle 22 <i>Weighted Likelihood Estimation (=WLE)</i> .....	118

Tabelle 23 <i>Regression zum Schlussfolgern/Wissenschaftsverständnis</i> .....	119
Tabelle 24 <i>Übersicht über Vorgehen bei Durchführung der Fördermaßnahmen</i> .....	127
Tabelle 25 <i>Evaluation von Ereignissen im Zusammenhang mit einer Vermutung</i> .....	130
Tabelle 26 <i>Vermutungen des Transfer-Tests</i> .....	131
Tabelle 27 <i>Übersicht über die Eigenschaften der Gegenstände</i> .....	132
Tabelle 28 <i>Kodierung der Antworten</i> .....	133
Tabelle 29 <i>Design der Vorstudie</i> .....	134
Tabelle 30 <i>Differenzwerte Post-Test und Prä-Test (Standardabweichungen)</i> .....	136
Tabelle 31 <i>Design Trainingsstudie</i> .....	142
Tabelle 32 <i>Summenwerte zum Vorwissen in den drei Gruppen</i> .....	144
Tabelle 33 <i>Relative Häufigkeiten der Einzelereignisse im Prä-Test</i> .....	146
Tabelle 34 <i>Relative Häufigkeiten der Einzelereignisse im Post-Test</i> .....	151
Tabelle 35 <i>Relative Häufigkeiten Interpretationsniveaus im Prä-Post-Vergleich</i> .....	153
Tabelle 36 <i>Mittelwerte (Standardabweichungen) Wissenschaftsverständnis</i> .....	151
Tabelle 37 <i>Itemkennwerte des Wissenschaftsverständnisses</i> .....	151
Tabelle 38 <i>Mittelwerte für die Einzelereignisse (Prä-Test und Post-Test)</i> .....	154
Tabelle 39 <i>Differenzwerte der ANOVA aus Prä- und Post-Test</i> .....	158
Tabelle 40 <i>Anova-Analysen der Kovariaten</i> .....	159
Tabelle 41 <i>Mittelwerte (Standardabweichungen) der Trainingsmaßnahmen</i> .....	157
Tabelle 42 <i>Prozessanalyse Ereignis [p q best]</i> .....	159
Tabelle 43 <i>Prozessanalyse für das Ereignis [p¬q wid]</i> .....	160
Tabelle 44 <i>Prozessanalyse für das Ereignis [¬p q irrel]</i> .....	161
Tabelle 45 <i>Prozessanalyse für das Ereignis [¬p ¬q irrel]</i> .....	162
Tabelle 46 <i>ANOVA mit Messwiederholung ( Training II und Transfer-Test)</i> .....	165
Tabelle 47 <i>Relative Häufigkeiten im Transfer-Test Schwimmen und Sinken</i> .....	166



Tabelle 48 <i>Mittelwerte (Standardabweichungen) im Transfer-Test</i> .....	168
Tabelle 49 <i>Mittelwerte (Standardabweichungen) der Unterstützungsmaßnahmen beim Transfer-Test</i> .....	168
Tabelle 50 <i>Häufigkeitsverteilung Argumentationstest</i> .....	169
Tabelle 51 <i>Häufigkeiten im Argumentationstest (nach Wertebereichen)</i> .....	170
Tabelle 52 <i>Zusammenhang zwischen Schlussfolgern und Argumentieren</i> .....	171

## Abbildungen

<i>Abbildung 1.</i> Modell Prozessqualität nach Fröhlich-Gildhoff et al (2011). .....	17
<i>Abbildung 2.</i> Schema zu <i>Scaffolding</i> nach Van de Pol (2010). .....	28
<i>Abbildung 3.</i> Erzieher-Kind-Dialoge I. ....	32
<i>Abbildung 4.</i> Erzieher-Kind-Dialoge II. ....	33
<i>Abbildung 5.</i> Scientific Dual Discovery Problem Solving (SDDS). .....	47
<i>Abbildung 6.</i> Die Relation von Theorie, Hypothese und Evidenz. ....	49
<i>Abbildung 7.</i> Vier Schlussweisen. ....	66
<i>Abbildung 8.</i> Vier Karten der „Wason Selection Task“. ....	72
<i>Abbildung 9.</i> Modell der Zentralen Exekutiven nach Baddeley & Hitch (1974). ....	77
<i>Abbildung 10.</i> Modell der Zentralen Exekutive nach Baddeley (2000). ....	78
<i>Abbildung 11.</i> Produktion von Evidenz. ....	93
<i>Abbildung 12.</i> Erkennen von Evidenz. ....	94
<i>Abbildung 13.</i> Interpretation von Theorie. ....	95
<i>Abbildung 14.</i> Interpretationsniveaus zum Schlussfolgern. ....	101
<i>Abbildung 15.</i> Relative Häufigkeiten Interpretationsniveaus. ....	110
<i>Abbildung 16.</i> Mittlere Punktzahlen beim Wissenschaftsverständnistest. ....	113
<i>Abbildung 17.</i> Wright Map des eindimensionalen Rasch-Modells. ....	117
<i>Abbildung 18.</i> Vorwissenstest, Symbole für Einschätzung zu Vermutungen. ....	126
<i>Abbildung 19.</i> Trainingsprozess in drei Schritten. ....	129
<i>Abbildung 20.</i> Visualisierung bei der Förderung (zweiter Schritt). ....	129
<i>Abbildung 21.</i> Relative Häufigkeiten beim Schlussfolgern im Prä und Post-Test. ....	136
<i>Abbildung 22.</i> Relative Häufigkeiten im Prä- und Post-Test. ....	152
<i>Abbildung 23.</i> Wright Map Trainingsstudie .....	153

*Abbildung 24.* Prozessanalyse für das Ereignis  $[p \ q \ \text{best}]$ ..... 159

*Abbildung 25.* Prozessanalyse für das Ereignis  $[p \neg q \ \text{wid}]$ ..... 160

*Abbildung 26.* Prozessanalyse für das Ereignis  $[\neg p \ q \ \text{irrel}]$ . .... 162

*Abbildung 27:* Prozessanalyse für das Ereignis  $[\neg p \neg q \ \text{irrel}]$ ..... 163

*Abbildung 28.* Ereignis  $[p \ q \ \text{best}]$  (Training II und Transfer-Test). .... 165

*Abbildung 29.* Ereignis  $[\neg p \ q \ \text{irrel}]$  (Training II und Transfer-Test)..... 166

*Abbildung 30.*  $[\neg p \neg q \ \text{irrel}]$  (Training II und Transfer-Test)..... 166

## 12 Anhang

Naive Theorie	Formen der Evidenz		Antw	Reihenfolge
Beispiel „Luft“	Der Tennisball ist mit Luft und springt. [pq]			
„Dinge, die mit Luft gefüllt sind, springen.“ [p q]; Bestantwort: konditionales Antwortmuster]	Der Spielball ist mit Luft gefüllt und springt. [p q]	☺ ☺ ☹	richtig	1
Spielball	Der hohle Fensterkittball ist mit Luft gefüllt und springt nicht. [p -q]	☺ ☹ ☹	falsch	4
aufg. Luftballon	Der Flummi ist ohne Luft gefüllt und springt. [-p q]	☺ ☺ ☹	irrelevant	3
Flummi	Der aufgeschnittene Luftballon ist ohne Luft gefüllt und springt nicht. [-p -q]	☺ ☹ ☹	irrelevant	2
hohler Fensterkittball				
„Dinge, die weich sind, springen nicht.“ [p q]; Bestantwort: konditionales Antwortmuster]	Das Kissen mit Körnern ist weich und springt nicht. [p q]	☺ ☹ ☹	Richtig	3
Filzball	Der Filzball ist weich und springt. [p -q]	☺ ☹ ☹	Falsch	1
tr. Fensterkittball	Der trockene Fensterkittball ist hart und springt nicht. [-p q]	☺ ☹ ☹	Irrelevant	2
Körnerkissen	Die Holzkugel ist hart und springt. [-p -q]	☺ ☹	Irrelevant	4

Holzku <span>g</span> el		☹		
„Dinge, die leicht sind springen.“ [p q]; Bestantwort: konditionales Antwortmuster]	Die Styroporkugel ist leicht und springt. [p q]	☺ ☹ ☹	Richtig	2
Luftballon m. Sand	Der Wattebausch ist leicht und springt nicht. [p ¬q]	☺ ☹ ☹	Falsch	3
Styroporkugel	Die Bocciakugel ist schwer und springt. [¬p q]	☺ ☹ ☹	Irrelevant	4
Wattebausch	Der Luftballon mit Sand ist schwer und springt nicht. [¬p ¬q]	☺ ☹ ☹	Irrelevant	1
Bocciakugel				
„Dinge, die hart sind, springen.“ [p q]; Bestantwort: konditionales Antwortmuster]	Der Murmel ist hart und springt. [p q]	☺ ☹ ☹	richtig	3
Murmel	Der trockene Knetklumpen ist hart und springt nicht. [p ¬q]	☺ ☹ ☹	falsch	1
Trockener Knetklumpen	Der Softball ist nicht hart und springt. [¬p q]	☺ ☹ ☹	irrelevant	2
Softball	Der Jonglierball ist nicht hart und springt nicht. [¬p ¬q]	☺ ☹ ☹	irrelevant	4
Jonglierball				
		☹ ☹		

<p>„ Dinge, die ohne Luft gefüllt sind, springen nicht.“</p> <p><i>[p q];</i></p> <p>Bestantwort: konditionales Antwortmuster]</p> <p>Knetball</p> <p>Tischtennisball</p> <p>Luftpolsterkissen</p> <p>Würfel</p>	<p>Der Knetball ist ohne Luft und springt nicht.</p> <p><i>[p q]</i></p>	<p>☺</p> <p>☹</p> <p>☹</p>	<p>richtig</p>	<p>1</p>
	<p>Der Würfel ist ohne Luft und springt.</p> <p><i>[p ¬q]</i></p>	<p>☺</p> <p>☹</p> <p>☹</p>	<p>falsch</p>	<p>4</p>
	<p>Das Luftpolsterkissen ist mit Luft gefüllt und springt nicht. <i>[¬p q]</i></p>	<p>☺</p> <p>☹</p> <p>☹</p>	<p>irrelevant</p>	<p>3</p>
	<p>Der Tischtennisball ist mit Luft gefüllt und springt.</p> <p><i>[¬p ¬q]</i></p>	<p>☺</p> <p>☹</p> <p>☹</p>	<p>irrelevant</p>	<p>2</p>
<p>„ Dinge, die schwer sind, springen nicht.“</p> <p><i>[p q];</i></p> <p>Bestantwort: konditionales Antwortmuster]</p> <p>Aufgeblasener Luftballon</p> <p>Tonklumpen</p> <p>Federsäckchen</p> <p>Eisenkugel</p>	<p>Der Tonklumpen ist schwer und springt nicht.</p> <p><i>[p q]</i></p>	<p>☺</p> <p>☹</p> <p>☹</p>	<p>richtig</p>	<p>2</p>
	<p>Die Eisenkugel ist schwer und springt.</p> <p><i>[p ¬q]</i></p>	<p>☺</p> <p>☹</p> <p>☹</p>	<p>falsch</p>	<p>4</p>
	<p>Das Federsäckchen ist leicht und springt nicht.</p> <p><i>[¬p q]</i></p>	<p>☺</p> <p>☹</p> <p>☹</p>	<p>irrelevant</p>	<p>3</p>
	<p>Der aufgeblasene Luftballon ist leicht und springt.</p> <p><i>[¬p ¬q]</i></p>	<p>☺</p> <p>☹</p> <p>☹</p>	<p>irrelevant</p>	<p>1</p>

## Interviewleitfaden Schlussfolgerungstest (Querschnittstudie)

### Begrüßen und Vorstellen

„Guten Morgen Ich bin \_\_\_\_\_.“ (Namen nennen).

„Das ist \_\_\_\_\_.“ (Vorstellen der wissenschaftlichen Hilfskraft/wissenschaftlichen Mitarbeiters)

„Wie heißt du?“ (Antwort *abwarten*)..

„Heute möchte ich herausfinden, wie Kinder in deinem Alter über bestimmte Dinge denken.“

„Das ist kein Test und es gibt kein „richtig“ oder „falsch“. Wichtig ist, dass du die Antwort nennst, die du am besten findest.“

„Schau mal.“ *Papierkugel und Knetball (mit anderer Farbe als im Test) fallen lassen.*

„Was hast du beobachtet?“ (Antwort *abwarten*).

(Genau, das eine Ding ist gesprungen, das andere ist auf dem Boden lieengeblieben.)

„Wir sprechen nur von *springen*, wenn das Ding wieder ein bisschen hochkommt. Wenn das Ding liegenbleibt, sagen wir, dass es nicht springt. Hast du das verstanden? (ggf. noch einmal erklären).“

„Für die Denkaufgaben sind wir zu einem Kindergarten gefahren und haben Kinder gefragt, welche Dinge springen. Diese unterschiedlichen Vermutungen haben wir gesammelt.“

„Eine Vermutung ist etwas, was ich zwar glaube, was ich aber noch überprüfen muss.“



„Das ist Tim.“ (*Bild vom Kind zeigen*).

„Ich lese Tims Vermutung, welche Dinge springen, einmal vor“:

Übung „Dinge, die mit Luft sind, springen.“

„Dazu nehme ich diesen Tennisball hier aus der Kiste (*Tennisball zeigen*). Das Besondere an dem Ball ist, dass er mit Luft ist. Gut. Jetzt probieren wir aus, ob der Tennisball springt.“

„Schau genau hin (*Ball fallen lassen*). Hast du gesehen, was passiert ist? „(*Antwort des Kindes abwarten*). „Ja, richtig. Der Tennisball springt.“

„Wir sprechen nur von springen, wenn das Ding wieder ein wenig vom Boden hochkommt. Der Tennisball springt, weil er vom Boden wieder hochkommt. Ich fasse nochmal zusammen: Der Tennisball ist rund und springt.“

„Jetzt denke nochmal an Tim und seine Vermutung. Tim vermutet ja: „Dinge, die mit Luft sind, springen.“ Was zeigt dir der Tennisball über Tims Vermutung?“

„Du hast drei Möglichkeiten zu antworten.“

„Wenn du glaubst, Tims Vermutung stimmt, legst du den Tennisball auf den grünen Smiley. Genau und deshalb siehst du hier einen lachenden Smiley.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun, legst du den Tennisball auf den gelben Smiley.“

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du den Tennisball auf den roten, traurigen Smiley.“ (*Auf den Smiley zeigen*.)

*Kind legt Tennisball auf einen Smiley.*

1a „Dinge, die mit Luft sind, springen.“ Material: Spielball [p q]

„Jetzt schauen wir uns noch einmal Tims Vermutung an“

„Ich lese sie noch einmal vor: „Dinge, die mit Luft sind, springen.“

„Um Tims Vermutung zu überprüfen, nehmen wir nun einen Spielball“ (*Spielball zeigen*).

„Das Besondere an dem Spielball ist, dass er mit Luft gefüllt ist.“

„Gut. Jetzt probieren wir aus, und schauen, ob der Spielball springt. Schau genau hin.“ (*Spielball fallen lassen*). (*Antwort des Kindes abwarten*).

Der Spielball springt.

„Ich fasse noch einmal zusammen. Der Spielball ist mit Luft und springt.“

„Denk jetzt noch einmal an Tims Vermutung.“

„Er vermutet ja: Dinge, die rund sind, springen.“

„Was zeigt dir der Spielball über Tims Vermutung?

Zeigt der Spielball, dass Tims Vermutung stimmt, hat der Spielball gar nichts mit Tims Vermutung zu tun, oder zeigt der Spielball, dass die Vermutung nicht stimmt?“

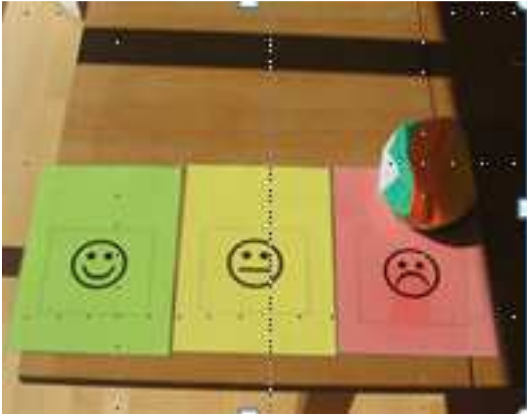
„Wenn du glaubst, Tims Vermutung stimmt, legst du den Spielball auf den grünen Smiley. Genau und deshalb siehst du hier einen grünen Smiley.“ (*Auf den Smiley zeigen*).

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun, legst du den Spielball auf den gelben Smiley“ (*Auf den Smiley zeigen*).

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du den Spielball auf den roten, traurigen Smiley.“ (*Auf den Smiley zeigen*).

*Kind legt den Spielball auf einen Smiley.*

Symbole (Smileys ) zur Evaluation von Ereignissen



TN Code:

\_\_\_\_\_ Kiga

\_\_\_\_\_ Datum:

\_\_\_\_\_

Frage	0	1	Bemerkungen
Woran erkennt die Mutter, dass Marius beim Klettern heruntergefallen ist?			
Woran erkennst du, dass Marius vom Baum gefallen ist?			
Was glaubst du, warum ist Marius vom Baum gefallen?			
Warum ist Marius vom Baum gefallen?			
Woran erkennt die Mutter, dass Elena sich bekleckert hat?			
Woran erkennst du, dass Elena sich bekleckert hat?			
Was glaubst du, warum dich Elena sich bekleckert hat?			
Warum hat Elena sich bekleckert?			
Woran erkennt die Oma, dass Struppi ganz nass ist?			
Woran erkennst du, dass der			

Struppi nass ist?			
Was glaubst du, warum ist Struppi ganz nass?			
Warum ist Struppi ganz nass?			

# Test zum Wissenschaftsverständnis

## Marius auf dem Baum

„Schau, ich lege dir vier Bilder hin. Die Bilder zeigen den Anfang und das Ende einer kleinen Geschichte.“

Bild 1 links außen, dann Bilder 2, 3 und 4 hinlegen, Bilder 2 und 3 verdeckt auflegen

„Auf dem ersten Bild siehst du Marius. Er klettert auf dem Baum herum.

Die Mutter kommt zum Baum (*zeigen auf Bild 4*). Sie guckt ganz erschrocken und sagt: „Oh je, du bist bestimmt beim Klettern runtergefallen!“

**GELB:** Was meinst du, woran erkennt die Mutter, dass Marius beim Klettern heruntergefallen ist?“ (*Produktion von Evidenz*)

Warten auf eine Antwort ( 30 s.)

**Bild 3 aufdecken**

„Hier siehst du, dass der Marius vom Baum gefallen ist. (*Zeigen auf Bild 3*)

**GRÜN:** „Woran siehst du, dass Marius vom Baum gefallen ist?“ (*Erkennen von Evidenz*) Schau genau auf das Bild.

Warten auf eine Antwort (30 s.)

„Jetzt fehlt uns noch ein wichtiges Bild.“

„Dieses Bild zeigt, warum der Marius vom Baum gefallen ist“ (*zeigen auf verdecktes Bild 2*)

**VERDECKT:** „Was glaubst du, warum ist Marius vom Baum gefallen?“ (*Produktion einer Theorie*)

Warten auf eine Antwort (30 s.)

**Bild 2 aufdecken**

„Hier siehst du, warum der Marius vom Baum gefallen ist.“ (*Zeigen auf Bild 2*)

**BLAU:** „Warum Marius ist vom Baum gefallen?“ (*Überprüfen der Theorie*)

## Elena und Sophie beim Essen

„Schau, ich lege dir vier Bilder hin. Die Bilder zeigen den Anfang und das Ende einer kleinen Geschichte.“

Bild 1 links außen, dann Bilder 2, 3 und 4 hinlegen, Bilder 2 und 3 verdeckt auflegen

„Auf dem ersten Bild siehst du Elena mit ihrer Schwester Sophie beim Mittagessen. Sie essen gemeinsam Pommes mit Ketchup.“

Auf dem letzten Bild (*zeigen auf Bild 4*) siehst du die Mutter der beiden Kinder. Sie sieht ärgerlich aus, weil Elena sich bekleckert hat.

**GELB:** „Woran erkennt die Mutter, dass Elena sich bekleckert hat?“ (*Produktion von Evidenz*)

Warten auf eine Antwort (30 s)

**Bild 3 aufdecken:**

„Hier siehst du Elena beim Essen. Sie hat sich bekleckert.“ (*Zeigen auf Bild 3*)

**GRÜN:** „Woran erkennst du, dass Elena sich bekleckert hat?“ (*Erkennen von Evidenz*). Schau genau auf das Bild.

Warten auf eine Antwort (30 s)

„Jetzt fehlt und noch ein wichtiges Bild.“

„Dieses Bild zeigt, warum sich Elena bekleckert hat.“ (*Zeigen auf verdecktes Bild 2*)

**VERDECKT:** „Was glaubst du, warum die Elena sich bekleckert hat?“ (*Produktion einer Theorie*)

Warten auf eine Antwort (30 s)

**Bild 2 aufdecken**

„Hier siehst du, was passiert ist.“ (*zeigen auf Bild 2*)

**BLAU:** „Warum hat Elena sich bekleckert?“ (*Überprüfen der Theorie*)

## Struppi und Marie auf ihrem Spaziergang

„Schau, ich lege dir vier Bilder hin. Die Bilder zeigen den Anfang und das Ende einer kleinen Geschichte.“

Bild 1 links außen, dann Bilder 2, 3 und 4 hinlegen, Bilder 2 und 3 verdeckt auflegen

„Auf dem ersten Bild siehst du Marie.

Hier sind Marie und Struppi wieder Zuhause.

Auf dem letzten Bild (*zeigen auf Bild 4*) siehst du die Oma von der Marie.“

Die Oma lacht und sagt: „Der Struppi ist ja ganz nass.“

**GELB:** Woran erkennt die Oma, dass der Struppi ganz nass ist?“ (*Produktion von Evidenz*)

Warten auf eine Antwort (30 s)

**Bild 3 aufdecken**

„Hier siehst du Struppi.“ (*zeigen auf Bild 3*)

**GRÜN:** „Woran erkennst du, dass der Struppi nass ist?“ (*Erkennen von Evidenz*).  
Schau genau auf das Bild.

Warten auf eine Antwort (30s)

„Jetzt fehlt noch ein wichtiges Bild.“

„Dieses Bild zeigt, warum Struppi ganz nass ist.“ (*Zeigen auf verdecktes Bild 2*)

**VERDECKT:** „Was glaubst du, warum ist Struppi ganz nass?“ (*Produktion einer Theorie*)

Warten auf eine Antwort (30s)

**Bild 2 aufdecken**

„Hier siehst du, warum Struppi nass ist.“ (*Zeigen auf Bild 2*)

**BLAU:** „Warum ist Struppi ganz nass?“ (*Prüfen der Evidenz*)



# Protokollbogen Wissenstest PRÄ (WIT\_prä)

Geschlecht: m w

Geburtsmonat: \_\_\_\_\_

Geburtsjahr: 20

ID: EL \_\_\_\_\_

1.

Das ist Mia [*Mia-Figur auf Tisch*].

Mia hat die Vermutung: **Dinge, die leicht sind, springen.**

[*Styroporkugel, Watte auf Tisch*]

Was denkst du über die Vermutung?

Bist du ganz sicher, dass Mias Vermutung stimmt, [*Bild mit zwei Daumen nach oben*]

glaubst du, dass die Vermutung stimmt, bist dir aber nicht ganz sicher [*Bild mit einem Daumen nach oben*]

oder glaubst du, dass die Vermutung nicht stimmt [*Bild mit einem Daumen nach unten*]

oder bist du ganz sicher, dass die Vermutung nicht stimmt [*auf Bild mit zwei Daumen nach unten*]

Lege das Kreuz auf das Antwortfeld, welches für dich am besten passt.

		stimmt	stimmt eher	stimmt eher nicht	stimmt nicht
Mia	Dinge, die leicht sind, springen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

**Warum glaubst du, dass die Vermutung [Antwortoption nennen]?  
bzw. Warum konntest du dich nicht entscheiden?**

---

**2.**

**Schauen wir mal, welche Vermutung Lena hat. [Lena-auf Tisch]**

**Lena vermutet: Dinge, die rund sind, springen.**

*[Spielball, Knetkugel]*

**Was denkst du über die Vermutung?**

**Bist du ganz sicher, dass Lenas Vermutung stimmt, [Bild mit zwei Daumen nach oben]**

**glaubst du, dass die Vermutung stimmt, bist dir aber nicht ganz sicher [Bild mit einem Daumen nach oben]**

**oder glaubst du, dass die Vermutung nicht stimmt [Bild mit einem Daumen nach unten]**

**oder bist du ganz sicher, dass die Vermutung nicht stimmt [auf Bild mit zwei Daumen nach unten]**

**Lege das Kreuz zuerst wieder auf das Antwortfeld, welches für dich am besten passt.**

		stimmt	stimmt eher	stimmt eher nicht	stimmt nicht
Lena	Dinge, die rund sind, springen.	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Warum glaubst du, dass die Vermutung *[Antwortoption nennen]*?  
bzw. Warum konntest du dich nicht entscheiden?

---

3.

Jetzt möchte ich dir noch Bens Vermutung zeigen. *[Ben auf Tisch]*

Ben vermutet: Dinge, die mit Luft sind, springen.

*[Wasserball, Luftballon]*

Was denkst du über die Vermutung?

Bist du ganz sicher, dass Bens Vermutung stimmt, *[Bild mit zwei Daumen nach oben]*

glaubst du, dass die Vermutung stimmt, bist dir aber nicht ganz sicher *[Bild mit einem Daumen nach oben]*

oder glaubst du, dass die Vermutung nicht stimmt *[Bild mit einem Daumen nach unten]*

oder bist du ganz sicher, dass die Vermutung nicht stimmt *[auf Bild mit zwei Daumen nach unten]*

Lege das Kreuz zuerst wieder auf das Antwortfeld, welches für dich am besten passt.



## Instruktion - Wissenstest PRE (WIT\_pre)

Hallo, [Name des Kindes],

Ich möchte dir gleich ein paar Fragen stellen.

Das ist aber kein Test. Es gibt also kein richtig oder falsch. Ich möchte nur wissen, was du denkst. Dabei ist jede Idee wichtig.

Damit ich keine deiner Antworten vergesse, haben ich ein Diktiergerät mitgebracht und würde unser Gespräch gerne aufnehmen, okay?

Gut, dann schalte ich es mal an. *[Diktiergerät einschalten]*

Okay, [Name des Kindes].

Ich habe hier außerdem noch einen Zettel *[Protokollbogen nehmen]*, auf dem ich noch einige deiner Antworten notiere. *[ID-Nummer Geschlecht im Protokollbogen eintragen]*.

Alles klar, dann lass uns mal beginnen. Ich habe dir etwas mitgebracht- schau einmal her:

*[Papierkugel und Knetkugel fallen lassen]*.

Es gibt Dinge, wie diese Papierkugel, die springen und Dinge, die bleiben liegen, wie diese Knetkugel.

Wenn etwas wieder ein Stückchen vom Boden hochkommt sagen wir, das Ding springt und wenn es auf dem Boden liegen bleibt sagen wir, es springt nicht.

Wir waren vor einiger Zeit schon in einem anderen Kindergarten und haben die Kinder gefragt, woran es denn liegen könnte, dass einige Dinge springen und andere Dinge nicht springen.

Die Vermutungen der Kinder haben wir dann gesammelt.

Eine Vermutung ist etwas, das man zwar glaubt, aber erst noch überprüfen muss.

*[Säckchen hervorholen und zeigen]* Ich vermute, dass in dem Säckchen ein Stoffhund versteckt ist.

Was glaubst du? *[Kind antworten lassen]*

Wie können wir herausbekommen, ob deine oder meine Vermutung stimmt? *[Kind antworten lassen]* Falls Kind keine Idee hat...

**(Wir können doch einfach nachschauen, wer Recht hat und so unsere Vermutungen überprüfen.)**

Dazu hast du gleich drei *[fünf]* Antwortmöglichkeiten *[Antworttafel mit Daumen vor das Kind legen]*:

Wenn du ganz sicher glaubst, dass die Vermutung stimmt, dann legst du das Kreuz *[Kind Antwortkreuz geben]* auf das Feld mit den „zwei Daumen“ *[auf entsprechendes Feld zeigen]*,

wenn du eher glaubst, dass die Vermutung stimmt, dir aber nicht ganz sicher bist, legst du das Kreuz auf das Feld „mit dem einen Daumen“ *[auf entsprechendes Feld zeigen]*,

**und wenn du eher glaubst, dass die Vermutung nicht, legst du das Kreuz *[Kind Antwortkreuz geben]***

**auf das Feld „mit dem „einen Daumen nach unten“** *[auf entsprechendes Feld zeigen],*

**und wenn du ganz sicher glaubst, dass die Vermutung nicht stimmt, dann legst du das Kreuz auf das Feld mit den „ zwei Daumen nach unten“**

*[auf entsprechendes Feld zeigen].*

**wenn du dich nicht entscheiden kannst, ob die Vermutung stimmt oder nicht stimmt, legst du das Kreuz auf das Feld „ mit dem Daumen zur Seite“,**

*[auf entsprechendes Feld zeigen],*

**Lass es uns mal ausprobieren:**

**Wo würdest du das Antwortkreuz hinlegen, wenn du ganz sicher glaubst, dass die Vermutung nicht stimmt?**

*[ggf. korrigieren]*

**Und wo würdest du das Kreuz hinlegen, wenn du eher glaubst, dass die Vermutung stimmt, dir aber nicht so sicher bist?**

*[ggf. korrigieren]*

Prima. Hast du noch eine Frage zu den Antwortmöglichkeiten, bevor wir loslegen?

*[ggf. noch einmal erklären, angefangen bei „Wenn du ganz sicher glaubst...“]*

Alles klar, dann lass uns beginnen.

*[weiter auf Protokollbogen]*

**Protokollbogen Laborstudie –Prä-Test– Tag 1 TN-Code: \_\_\_\_\_ Datum: \_\_\_\_\_**

Vermutung (+Materialreihenfolge)	Formen der Evidenz	Wahl des Kindes		Bestantw.
<b>„Dinge, die rund sind, springen.“</b> [p q]  Knetkugel Flummie Spielball Waschlappen	Der Spielball ist rund und springt. [p q]	☺ ☹ ☹	3	richtig
	Die farbige Knetkugel ist rund und springt nicht. [p ¬q]	☺ ☹ ☹	1	falsch
	Das Flummie ist nicht rund und springt. [¬p q]	☺ ☹ ☹	2	irrelevant
	Der W-lappen ist nicht rund und springt nicht. [¬p ¬q]	☺ ☹ ☹	4	irrelevant
<b>„Dinge, die weich sind, springen.“</b> [p q]  Filzball Buch Holzkugel Körnerkissen	Der Filzball ist weich und springt. [p q]	☺ ☹ ☹	1	richtig
	Das Körnerkissen ist weich und springt nicht. [p ¬q]	☺ ☹ ☹	4	falsch
	Die Holzkugel ist hart und springt. [¬p q]	☺ ☹ ☹	3	irrelevant
	Das Buch ist hart und springt nicht. [¬p ¬q]	☺ ☹ ☹	2	irrelevant
<b>„Dinge, die schwer sind, springen.“</b> [p q]  Wattebausch Bocciakugel Luftballon mit Sand Styroporkugel	Der Bocciakugel ist schwer und springt. [p q]	☺ ☹ ☹	2	richtig
	Der Luftballon mit Sand ist schwer und springt nicht. [p ¬q]	☺ ☹ ☹	3	falsch
	Die Styroporkugel ist leicht und springt. [¬p q]	☺ ☹ ☹	4	irrelevant
	Der Wattebausch ist leicht und springt nicht. [¬p ¬q]	☺ ☹ ☹	1	irrelevant

## Interviewleitfaden Prä-Test (Trainingstudie)

„Heute möchte ich herausfinden, wie Kinder in deinem Alter über bestimmte Dinge denken.“

„Das ist kein Test und es gibt kein „richtig“ oder „falsch“. Wichtig ist, dass du die Antwort nennst, die du am besten findest.“

„Schau mal.“ *Tischtennisball und Knetkugel (mit anderer Farbe als im Test) fallen lassen.*

„Was hast du beobachtet?“ *Antwort abwarten.*

(Genau, das eine Ding ist gesprungen, das andere ist auf dem Boden lieengeblieben.)

„Wir sprechen nur von *springen*, wenn das Ding wieder ein bisschen hochkommt. Wenn das Ding liegenbleibt, sagen wir, dass es nicht springt. Hast du das verstanden?“ (ggf. noch einmal erklären).“

„Für die Denkaufgaben sind wir zu einem Kindergarten gefahren und haben Kinder gefragt, welche Dinge springen. Diese unterschiedlichen Vermutungen haben wir gesammelt.“

„Eine Vermutung ist etwas, was ich zwar glaube, was ich aber noch überprüfen muss.“

„Wir wollen heute verschiedene Vermutungen überprüfen.“



## 1. Das ist Petra.

„Ich lese Petras Vermutung, welche Dinge springen, einmal vor“:

Übung „Dinge, die schwer sind, springen.“

„Dazu nehme ich diesen Golfball hier aus der Kiste.“ (*Golfball zeigen*). „Das Besondere an dem Ball ist, dass er schwer ist. Gut. Jetzt probieren wir aus, ob der Golfball springt.“

„Schaut genau hin.“ (*Ball fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig. Der Golfball springt.“

„Wir sprechen nur von springen, wenn das Ding wieder ein wenig vom Boden hochkommt. Der Tischtennisball springt, weil er vom Boden wieder hochkommt. Ich fasse nochmal zusammen: Der Golfball ist schwer und springt.“

„Jetzt denke nochmal an Petra und seine Vermutung. Paul vermutet ja: „Dinge, die schwer sind, springen.“ „Was zeigt dir Golfball über Pauls Vermutung?“

„Du hast drei Möglichkeiten zu antworten.“

„Wenn du glaubst, Petra Vermutung stimmt, legst du den Golfball hierhin.“ (*Auf den Smiley zeigen.*)

„Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du den Golfball auf den hierhin.“

„Wenn du glaubst, Petras Vermutung stimmt nicht, legst du den Golfball hierhin.“ (*Auf den Smiley zeigen.*)

*Kind legt Golfball auf einen Smiley.*

## **„Dinge, die schwer sind, springen.“**

Wattebausch [ $\neg p \neg q$ ], Bocciakugel [ $p q$ ], Luftballon mit Sand [ $p \neg q$ ], Styroporkugel [ $\neg p q$ ]

„Ich lese dir Petras Vermutung, welche Dinge springen, einmal vor“:

„Dinge, die schwer sind, springen.“

„Dazu nehme ich diesen Wattebausch hier aus der Kiste (*Wattebausch zeigen*). Das Besondere an dem Wattebausch ist, dass er leicht ist. Gut. Jetzt probieren wir aus, ob der Wattebausch springt.“

„Schau genau hin.“ (*Wattebausch fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig der Wattebausch ist leicht und springt nicht.“

„Jetzt denke nochmal an Petras Vermutung.“ „Petra vermutet ja: „Dinge, die schwer sind, springen.“ „Was zeigt dir der Wattebausch über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du den Wattebausch hierhin“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du den Wattebausch hierhin. Wenn du glaubst, Petras Vermutung stimmt nicht, legst du den Wattebausch hierhin.“ (*Auf den Smiley zeigen*.)

*Kind legt den Wattebausch auf einen Smiley.*

„Jetzt überprüfen wir Petras Vermutung mit einem anderen Ding.“

„Sie vermutet: „Dinge, die schwer sind, springen.“

„Dazu nehme ich diese Bocciakugel hier aus der Kiste (*Bocciakugel zeigen*). Das Besondere an der Bocciakugel ist, dass sie schwer ist. Gut. Jetzt probieren wir aus, ob die Bocciakugel springt.“

„Schau genau hin.“ (*Bocciakugel fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig die Bocciakugel ist schwer und springt.“

„Jetzt denke nochmal an Petras Vermutung.“ 2Petra vermutet ja: „Dinge, die schwer sind, springen.“ „Was zeigt dir die Bocciakugel über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du die hierhin.“ (*Auf den Smiley zeigen.*)

„Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du die Bocciakugel hierhin.“

„Wenn du glaubst, Petras Vermutung stimmt nicht, legst du die hierhin.“ (*Auf den Smiley zeigen.*)

*Kind legt die Bocciakugel auf einen Smiley.*

„Jetzt überprüfen wir Petras Vermutung mit einem weiteren Ding.“

Sie vermutet: „Dinge, die schwer sind, springen.“

„Dazu nehme ich diesen Luftballon mit Sand hier aus der Kiste.“ (Luftballon mit Sand zeigen). „Das Besondere an dem Luftballon mit Sand ist, dass er schwer ist. Gut. Jetzt probieren wir aus, ob der Luftballon mit Sand springt.“

„Schaut genau hin (Luftballon mit Sand *fallen lassen*). Hast du gesehen, was passiert ist?“ (Antwort des Kindes abwarten). „Ja, richtig. Der Luftballon mit Sand ist schwer und springt nicht.“

„Jetzt denke nochmal an Petra und ihre Vermutung.“ „Petra vermutet ja: „Dinge, die schwer sind, springen.“ „Was zeigt dir der Luftballon mit Sand über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du den Luftballon mit Sand hierhin“ (Auf den Smiley zeigen.)

Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du den Luftballon mit Sand hierhin “ (Auf den Smiley zeigen.)

„Wenn du glaubst, Petras Vermutung stimmt nicht, legst du den Luftballon mit Sand hierhin.“ (Auf den Smiley zeigen.)

*Kind legt Luftballon mit Sand auf einen Smiley.*

„Jetzt überprüfen wir Petras Vermutung mit einem letzten Ding.“

Sie vermutet: „Dinge, die schwer sind, springen.“

„Dazu nehme ich diese Styroporkugel hier aus der Kiste.“ (*Styroporkugel zeigen*). „Das Besondere an der Styroporkugel ist, dass sie leicht ist. Gut. Jetzt probieren wir aus, ob die Styroporkugel springt.“

„Schau genau hin.“ (*Styroporkugel fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig die Styroporkugel ist leicht und springt.“

„Jetzt denke nochmal an Petra und ihre Vermutung.“ Petra vermutet ja: „Dinge, die schwer sind, springen.“ „Was zeigt dir die Styroporkugel über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du die Styroporkugel hierhin“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du die Styroporkugel hierhin“ (*Auf den Smiley zeigen*.)

*Kind legt Styroporkugel auf einen Smiley.*

## „Dinge, die rund sind, springen“

Knetkugel [ $p \neg q$ ], Flummie [ $\neg p q$ ], Spielball [ $p q$ ], Waschlappen [ $\neg p \neg q$ ]

2. Das ist Tim.

„Tims Vermutung wollen wir jetzt überprüfen.“

Er vermutet: „Dinge, die rund sind, springen.“

„Dazu nehme ich diese Knetkugel hier aus der Kiste.“ (*Knetkugel zeigen*). „Das Besondere an der Kugel ist, dass sie rund ist.“ „Gut. Jetzt probieren wir aus, ob die Knetkugel springt.“

„Schau genau hin.“ (*Knetkugel fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig. Die Knetkugel ist rund und springt nicht.“

„Jetzt denke nochmal an Tim und seine Vermutung.“ Tim vermutet ja: „Dinge, die rund sind, springen.“ „Was zeigt dir die Knetkugel über Tims Vermutung?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du die Knetkugel hierhin“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun, legst du die Knetkugel hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du die Knetkugel hierhin.“ (*Auf den Smiley zeigen*.)

*Kind legt Knetkugel auf einen Smiley.*

„Jetzt überprüfen wir Tims Vermutung mit einem anderen Ding.“

„Er vermutet: Dinge, die rund sind, springen.“

„Dazu nehmen wir ein Flummie.“ (*Flummie zeigen*).

„Das Besondere am Flummie ist, dass es nicht rund ist.“

„Gut. Jetzt probieren wir aus, und schauen, ob das Flummie springt. Schau genau hin.“

(*Flummie fallen lassen*). (*Antwort des Kindes abwarten*)

„Das Flummie ist nicht rund und springt.“

„Denk jetzt noch einmal an Tims Vermutung.“

„Er vermutet ja: Dinge, die rund sind, springen.“

„Was zeigt dir das Flummie über Tims Vermutung?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du das Flummie\_hierhin“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun, legst du das Flummie hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du das Flummie hierhin.“ (*Auf den Smiley zeigen*.)

*Kind legt das Flummie auf einen Smiley.*

„Jetzt überprüfen wir Tims Vermutung mit einem weiteren Ding.“

„Er vermutet: Dinge, die rund sind, springen.“

„Dazu nehme ich diesen Spielball hier.“ (*Spielball zeigen*). „Das Besondere an dem Ball ist, dass er rund ist.“ „Gut. Jetzt probieren wir aus, ob der Spielball springt.“

„Schau genau hin“ (*Ball fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig.“

„Der Spielball ist rund und springt.“

„Jetzt denke nochmal an Tim und seine Vermutung.“ Tim vermutet ja: „Dinge, die rund sind, springen.“ „Was zeigt dir Spielball über Tims Vermutung?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du den Spielball hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun legst du den Spielball hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du den Spielball hierhin.“. (*Auf den Smiley zeigen*.)

*Kind legt Spielball auf einen Smiley.*



„Jetzt überprüfen wir Tims Vermutung mit einem anderen Ding.“

Er vermutet: „Dinge, die rund sind, springen.“

„Dazu nehmen wir nun einen Waschlappen“ (*Waschlappen zeigen*).

„Das Besondere an dem Waschlappen ist, dass es nicht rund ist.“ „Gut. Jetzt probieren wir aus, ob der Waschlappen springt.“ (*Waschlappen fallen lassen*). (Antwort des Kindes abwarten).

„Der Waschlappen ist nicht rund springt nicht.“

„Jetzt denk noch einmal an Tims Vermutung.“

„Er vermutet ja: Dinge, die rund sind, springen.“

„Was zeigt dir der Waschlappen über Tims Vermutung?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du den Waschlappen hierhin.“ (*Auf den Smiley zeigen*.)

Wenn du glaubst, Tims Vermutung hat damit nichts zu tun legst du den Waschlappen hierhin.“ (*Auf den Smiley zeigen*.)

Wenn du glaubst, Tims Vermutung stimmt nicht, legst du den Waschlappen hierhin.“ (*Auf den Smiley zeigen*.)

*Kind legt Waschlappen auf einen Smiley.*

## „Dinge, die weich sind, springen“

Filzball [ $p \ q$ ], Buch [ $\neg p \ \neg q$ ], Holzkugel [ $\neg p \ q$ ], Körnerkissen [ $p \ \neg q$ ]

„Jetzt überprüfen wir Pauls Vermutung.“

Paul vermutet: „Dinge, die weich sind, springen.“

„Dazu nehme ich diesen Filzball hier aus der Kiste.“ (*Filzball zeigen*). „Das Besondere am Filzball ist, dass er weich ist.“ „Gut. Jetzt probieren wir aus, ob der Filzball springt.“

„Schau genau hin.“ (*Filzball fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig der Filzball ist weich und springt.“

„Jetzt denke nochmal an Pauls Vermutung.“ „Er vermutet ja: „Dinge, die weich sind, springen.“ „Was zeigt dir der Filzball über Pauls Vermutung?“

„Wenn du glaubst, Pauls Vermutung stimmt, legst du den Filzball hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Pauls Vermutung hat damit nichts zu tun, legst du den Filzball hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Pauls Vermutung stimmt nicht, legst du den Filzball hierhin.“ (*Auf den Smiley zeigen*.)

*Kind legt den Filzball auf einen Smiley.*

„Jetzt überprüfen wir Pauls Vermutung mit einem anderen Ding.“

Er vermutet: „Dinge, die weich sind, springen.“

„Dazu nehme ich das Buch hier“ (*Buch zeigen*).

„Das Besondere an dem Buch ist, dass es hart ist. Gut. Jetzt probieren wir aus, ob das Buch springt.“

„Schau genau hin.“ (*Buch fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig das Buch ist hart und springt nicht.“

„Jetzt denke nochmal an Pauls Vermutung.“ „Er vermutet ja: „Dinge, die weich sind, springen.“ „Was zeigt dir das Buch über Pauls Vermutung?“

„Wenn du glaubst, Pauls Vermutung stimmt, legst du das Buch hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Pauls Vermutung hat damit nichts zu tun, legst du das Buch hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Pauls Vermutung stimmt nicht, legst du das Buch hierhin.“ (*Auf den Smiley zeigen*.)

*Kind legt das Buch auf einen Smiley.*

„Jetzt überprüfen wir Pauls Vermutung mit einem anderen Ding.“

Er vermutet: „Dinge, die weich sind, springen.“

„Dazu nehme ich diese Holzkugel.“ (*Holzkugel zeigen*). „Das Besondere an der Holzkugel ist, dass sie hart ist. Gut. Jetzt probieren wir aus, ob die Holzkugel springt.“

„Schau genau hin.“ (*Holzkugel fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig die Holzkugel ist hart und springt.“

„Jetzt denke nochmal an Pauls Vermutung. Er vermutet ja: „Dinge, die weich sind, springen.“

„Was zeigt dir die Holzkugel über Pauls Vermutung?“

„Wenn du glaubst, Pauls Vermutung stimmt, legst du die Holzkugel hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Pauls Vermutung hat damit nichts zu tun, legst du die Holzkugel hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Pauls Vermutung stimmt nicht, legst du die Holzkugel hierhin.“ (*Auf den Smiley zeigen*.)

*Kind legt die Holzkugel auf einen Smiley.*

„Jetzt überprüfen wir Pauls Vermutung mit einem anderen Ding.“

Er vermutet: „Dinge, die weich sind, springen.“

„Dazu nehme ich dieses Körnerkissen hier aus der Kiste.“ (*Körnerkissen zeigen*). „Das Besondere am Körnerkissen ist, dass es weich ist.“ Gut. Jetzt probieren wir aus, ob das Körnerkissen springt.“

„Schau genau hin.“ (*Körnerkissen fallen lassen*). „Hast du gesehen, was passiert ist? (*Antwort des Kindes abwarten*) Ja, richtig das Körnerkissen ist weich und springt nicht. „Jetzt denke nochmal an Pauls Vermutung.“ „Er vermutet ja: „Dinge, die weich sind, springen. „Was zeigt dir das Körnerkissen über Pauls Vermutung?“

„Wenn du glaubst, Pauls Vermutung stimmt, legst du das Körnerkissen hierhin“ (*Auf den Smiley zeigen.*)

„Wenn du glaubst, Pauls Vermutung hat damit nichts zu tun, legst du das Körnerkissen hierhin.“ (*Auf den Smiley zeigen.*)

„Wenn du glaubst, Pauls Vermutung stimmt nicht, legst du das Körnerkissen hierhin.“ (*Auf den Smiley zeigen.*)

*Kind legt das Körnerkissen auf einen Smiley.*

## Interviewleitfaden Post-Test

„Guten Morgen“ (*Name des Kindes notieren.*) „Heute wollen wir noch einmal Vermutungen von anderen Kindern überprüfen.“

„Hier ist noch einmal die weiche und die harte Unterlage.“

„Sortiere die Dinge danach, ob sie hart oder weich sind.“

Wolle, Jonglierball (weich)

Golfball, Knetklumpen (hart)

### 1. „Dinge, die weich sind, springen“

Wolle [p q], Jonglierball [p ¬q], Golfball [¬p q], Knetklumpen [¬p ¬q]

„Ich lese dir Pauls Vermutung, welche Dinge springen, einmal vor:

Er vermutet: „Dinge, die weich sind, springen.“

„Die Vermutung wollen wir jetzt überprüfen.“

„Dazu nehme ich diese Wolle hier aus der Kiste.“ (*Wolle zeigen*). „Das Besondere an der Wolle ist, dass sie weich ist. Gut. Jetzt probieren wir aus, ob die Wolle springt.“

„Schau genau hin.“ (*Wolle fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig die Wolle ist weich und springt.“

„Jetzt denke nochmal an Pauls Vermutung. Paul vermutet ja: „Dinge, die weich sind, springen. Was zeigt dir die Wolle über Pauls Vermutung?“

„Wenn du glaubst, Pauls Vermutung stimmt, legst du die Wolle hierhin“ (*auf den Smiley zeigen*.)

„Wenn du glaubst, Pauls Vermutung hat damit nichts zu tun, legst du die Wolle hierhin.“ (*auf den Smiley zeigen*.)

„Wenn du glaubst, Pauls Vermutung stimmt nicht, legst du die Wolle hierhin.“ (*auf den Smiley zeigen*.)

*Kind legt Wolle auf einen Smiley.*

„Jetzt überprüfen wir Pauls Vermutung mit einem anderen Ding.“

Er vermutet: „Dinge, die weich sind, springen.“

„Dazu nehme ich diesen Jonglierball hier aus der Kiste (*Jonglierball zeigen*). Das Besondere am Jonglierball ist, dass er weich ist. Gut. Jetzt probieren wir aus, ob der Jonglierball springt.“

„Schau genau hin.“ (*Jonglierball fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig der Jonglierball ist weich und springt nicht.“

„Jetzt denke nochmal an Pauls Vermutung. Paul vermutet: „Dinge, die weich sind, springen. Was zeigt dir der Jonglierball über Pauls Vermutung?“

„Wenn du glaubst, Pauls Vermutung stimmt, legst du den Jonglierball hierhin.“ (*auf den lachenden Smiley zeigen*.)

„Wenn du glaubst, Pauls Vermutung hat damit nichts zu tun, legst du den Jonglierball hierhin.“ (*auf den nachdenklichen Smiley zeigen*.)

„Wenn du glaubst, Pauls Vermutung stimmt nicht, legst du den Jonglierball hierhin.“ (*auf den traurigen Smiley zeigen*.)

*Kind legt den Jonglierball auf einen Smiley.*

Jetzt überprüfen wir Pauls Vermutung mit einem anderen Ding.

Er vermutet: „Dinge, die weich sind, springen.“

„Dazu nehme ich diesen Golfball.“ (*Golfball zeigen*). „Das Besondere an dem Golfball ist, dass er hart ist. Gut. Jetzt probieren wir aus, ob der Golfball springt.“

„Schau genau hin.“ (*Golfball fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig der Golfball ist hart und springt.“

„Jetzt denke nochmal an Pauls Vermutung. Paul vermutet ja: „Dinge, die weich sind, springen.“ Was zeigt dir der Golfball über Pauls Vermutung?“

„Wenn du glaubst, Pauls Vermutung stimmt, legst du den Golfball hierhin.“ (*Auf den lachenden Smiley zeigen.*)

„Wenn du glaubst, Pauls Vermutung hat damit nichts zu tun, legst du den Golfball hierhin.“ (*Auf den nachdenklichen Smiley zeigen.*)

„Wenn du glaubst, Pauls Vermutung stimmt nicht, legst du Golfball hierhin.“ (*Auf den traurigen Smiley zeigen.*)

*Kind legt den Golfball auf einen Smiley.*

„Jetzt wollen wir Pauls Vermutung mit einem anderen Ding überprüfen.“

Er vermutet. „Dinge, die weich sind, springen.“

„Dazu nehme ich den trockenen Knetklumpen hier aus der Kiste.“ (*Roten Knetklumpen zeigen.*)

„Das Besondere an dem Knetklumpen ist, dass es hart ist. Gut. Jetzt probieren wir aus, ob der Knetklumpen springt.“

„Schau genau hin.“ (*Knetklumpen fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig der Knetklumpen ist hart und springt nicht. Jetzt denke nochmal an Pauls Vermutung. Paul vermutet ja: „Dinge, die weich sind, springen.“ Was zeigt dir der Knetklumpen über Pauls Vermutung?“

„Wenn du glaubst, Pauls Vermutung stimmt, legst du den Knetklumpen hierhin.“ (*Auf den lachenden Smiley zeigen.*)

„Wenn du glaubst, Pauls Vermutung hat damit nichts zu tun, legst du den Knetklumpen hierhin.“ (*Auf den nachdenklichen Smiley zeigen.*)

„Wenn du glaubst, Pauls Vermutung stimmt nicht, legst du den Knetklumpen hierhin.“ (*Auf den traurigen Smiley zeigen.*)

*Kind legt den Knetklumpen auf einen Smiley.*



Das ist Tim.

## 2. „Dinge, die rund sind, springen.“

Flummi [ $p \rightarrow q$ ], Luftballon (nicht aufgeblasen) [ $\neg p \rightarrow \neg q$ ], Überraschungsei [ $\neg p \rightarrow q$ ], Knetkugel [ $p \rightarrow \neg q$ ],

„Ich lese Tims Vermutung, welche Dinge springen, einmal vor:

„Dinge, die rund sind, springen.“

„Jetzt wollen wir Tims Vermutung „Dinge, die rund sind, springen.“ überprüfen.“

„Dazu nehme ich diesen Flummi hier aus der Kiste.“ (*Flummi zeigen*). „Das Besondere an ihm ist, dass er rund ist. Gut. Jetzt probieren wir aus, ob der Flummi springt.“

„Schau genau hin.“ (*Flummi fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig. Der Flummi springt.“

„Der Flummi ist rund und springt.“

„Jetzt denke nochmal an Tim und seine Vermutung. Tim vermutet ja: „Dinge, die rund sind, springen.“ Was zeigt dir der Flummi über Tims Vermutung?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du den Flummi hierhin.“ (*Auf den lachenden Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun legst du den Flummi hierhin.“ (*Auf den nachdenklichen Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du den Flummi hierhin.“ (*Auf den traurigen Smiley zeigen*.)

*Kind legt Flummi auf einen Smiley.*

Jetzt wollen wir Tims Vermutung mit einem anderen Ding überprüfen.

Er vermutet. „Dinge, die rund sind, springen.“

„Um Tims Vermutung zu überprüfen, nehmen wir nun einen nicht aufgeblasenen Luftballon.“  
(*Luftballon zeigen*).

„Das Besondere an dem nicht aufgeblasenen Luftballon ist, dass er nicht rund ist. Gut. Jetzt probieren wir aus, und schauen, ob das der Luftballon springt. Schau genau hin.“ (*Luftballon fallen lassen*). (*Antwort des Kindes abwarten*).

„Der Luftballon springt nicht.“

„Denk jetzt noch einmal an Tims Vermutung.“

„Er vermutet ja: „Dinge, die rund sind, springen.“

„Was zeigt dir der Luftballon über Tims Vermutung?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du den nicht aufgeblasenen Luftballon hierhin.“ (*Auf den lachenden Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun legst du den nicht aufgeblasenen Luftballon hierhin.“ (*Auf den nachdenklichen Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du den nicht aufgeblasenen Luftballon hierhin.“ (*Auf den traurigen Smiley zeigen*.)

*Kind legt den Luftballon auf einen Smiley.*

„Jetzt wollen wir Tims Vermutung mit einem anderen Ding überprüfen.“

„Er vermutet. „Dinge, die rund sind, springen.“

„Dazu nehmen wir nun ein Überraschungsei“ (*Überraschungsei zeigen*).

„Das Besondere an dem Überraschungsei ist, dass es nicht rund ist.“

„Gut. Jetzt probieren wir aus und schauen, ob das Überraschungsei springt. Schau genau hin.“  
(*Überraschungsei fallen lassen*). „Was passiert?“ (*Antwort des Kindes abwarten*).

„Das Überraschungsei springt.“

„Das Überraschungsei ist nicht rund und springt.“

„Denk jetzt noch einmal an Tims Vermutung.“

„Er vermutet ja: „Dinge, die rund sind, springen.“

„Was zeigt dir das Überraschungsei über Tims Vermutung?“

„Zeigt das Überraschungsei, dass Tims Vermutung stimmt, hat das Überraschungsei gar nichts mit Tims Vermutung zu tun, oder zeigt das Überraschungsei, dass die Vermutung nicht stimmt?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du das Überraschungsei hierhin.“ (*Auf den lachenden Smiley zeigen.*)

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun, legst du das Überraschungsei hierhin.“ (*Auf den nachdenklichen Smiley zeigen.*)

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du das Überraschungsei hierhin.“ (*Auf den traurigen Smiley zeigen.*)

*Kind legt Überraschungsei auf einen Smiley.*

„Dazu nehme ich diese Knetkugel hier aus der Kiste.“ (*Knetkugel zeigen*). „Das Besondere an der Kugel ist, dass sie rund ist. Gut. Jetzt probieren wir aus, ob die Knetkugel springt. Schau genau hin.“ (*Knetkugel fallen lassen*).

„Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig. Die Knetkugel ist rund und springt nicht.“

„Jetzt denke nochmal an Tim und seine Vermutung. Tim vermutet ja: „Dinge, die rund sind, springen.“ Was zeigt dir Knetkugel über Tims Vermutung?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du den Knetkugel hierhin.“ (*Auf den lachenden Smiley zeigen.*)

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun, legst du die Knetkugel hierhin.“ (*Auf den nachdenklichen Smiley zeigen.*)

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du die Knetkugel hierhin.“ (*Auf den traurigen Smiley zeigen.*)

*Kind legt Knetkugel auf einen Smiley.*

### 3. „Dinge, die schwer sind, springen.“

Federsäckchen [ $\neg p \neg q$ ], Eisenkugel [ $p q$ ], Stein [ $p \neg q$ ], Luftballon [ $\neg p q$ ]  
(aufgeblasen)

„Das ist Petra.“

Ich lese dir Petras Vermutung, welche Dinge springen, einmal vor“:

„Dinge, die schwer sind, springen.“

„Jetzt wollen wir Petras Vermutung überprüfen.“

„Dazu nehme ich dieses Federsäckchen hier aus der Kiste.“ (*Federsäckchen zeigen*). „Das Besondere an dem Federsäckchen ist, dass es leicht ist. Gut. Jetzt probieren wir aus, ob das Federsäckchen springt.“

„Schau genau hin.“ (*Federsäckchen fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig das Federsäckchen ist leicht und springt nicht.“

„Jetzt denke nochmal an Petras Vermutung. Petra vermutet ja: „Dinge, die schwer sind, springen.“ Was zeigt dir das Federsäckchen über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du das Federsäckchen hierhin.“ (*Auf den lachenden Smiley zeigen*.)

„Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du das Federsäckchen hierhin.“ (*Auf den nachdenklichen Smiley zeigen*.)

„Wenn du glaubst, Petras Vermutung stimmt nicht, legst du das Federsäckchen hierhin.“ (*Auf den traurigen Smiley zeigen*.)

*Kind legt das Federsäckchen auf einen Smiley.*

„Jetzt wollen wir Petras Vermutung noch mit einem anderen Ding überprüfen.“

„Sie vermutet: „Dinge, die schwer sind, springen.“

„Dazu nehme ich diese Eisenkugel hier aus der Kiste.“ (*Eisenkugel zeigen*). „Das Besondere an der Eisenkugel ist, dass sie schwer ist. Gut. Jetzt probieren wir aus, ob die Eisenkugel springt.“

„Schau genau hin.“ (*Eisenkugel fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig die Eisenkugel ist schwer und springt.“

Jetzt denke nochmal an Petras Vermutung. Petra vermutet ja: „Dinge, die schwer sind, springen.“ Was zeigt dir die Eisenkugel über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du die Eisenkugel hierhin.“ (*Auf den lachenden Smiley zeigen*.)

„Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du die Eisenkugel hierhin.“(*Auf den nachdenklichen Smiley zeigen*)

„Wenn du glaubst, Petras Vermutung stimmt nicht, legst du die Eisenkugel hierhin.“ (*Auf den traurigen Smiley zeigen*.)

*Kind legt die Eisenkugel auf einen Smiley.*

„Jetzt wollen wir Petras Vermutung noch mit einem anderen Ding überprüfen.“

„Sie vermutet: „Dinge, die schwer sind, springen.“

„Dazu nehme ich diesen Stein hier aus der Kiste.“ (*Stein zeigen*). „Das Besondere an dem Stein ist, dass er schwer ist. Gut. Jetzt probieren wir aus, ob der Stein springt.“

„Schaut genau hin.“ (*Stein fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig. Der Stein ist schwer und springt nicht.“

„Jetzt denke nochmal an Petra und ihre Vermutung.“ „Petra vermutet ja: „Dinge, die schwer sind, springen.“ „Was zeigt dir der Stein über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du den Stein hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du den Stein hierhin.“ (*Auf den Smiley zeigen*).

„Wenn du glaubst, Petras Vermutung stimmt nicht, legst du den Stein hierhin.“ (*Auf Smiley zeigen*.)

*Kind legt den Stein auf einen Smiley.*

„Jetzt wollen wir Petras Vermutung noch mit einem anderen Ding überprüfen.“

„Sie vermutet: „Dinge, die schwer sind, springen.“

„Dazu nehme ich diesen aufgeblasenen Luftballon hier aus der Kiste.“ (*Luftballon zeigen*).

„Das Besondere an dem aufgeblasenen Luftballon ist, dass er leicht ist. Gut. Jetzt probieren wir aus, ob der aufgeblasene Luftballon springt.“

„Schau genau hin.“ (*Luftballon fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig, der Luftballon ist leicht und springt.“

„Jetzt denke nochmal an Petra und ihre Vermutung.“ „Petra vermutet ja: „Dinge, die schwer sind, springen.“ „Was zeigt dir der Luftballon über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du den Luftballon hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du den Luftballon hierhin.“ (*Auf den Smiley zeigen*).

„Wenn du glaubst, Petras Vermutung stimmt nicht, legst du den Luftballon hierhin.“ (*Auf den Smiley zeigen*.)

*Kind legt den Luftballon auf einen Smiley.*

## Interviewleitfaden Post-Test

„Guten Morgen“ (*Name des Kindes notieren.*) „Heute wollen wir noch einmal Vermutungen von anderen Kindern überprüfen.“

„Hier ist noch einmal die weiche und die harte Unterlage.“

„Sortiere die Dinge danach, ob sie hart oder weich sind.“

Wolle, Jonglierball (weich)

Golfball, Knetklumpen (hart)

### 1. „Dinge, die weich sind, springen“

Wolle [p q], Jonglierball [p ¬q], Golfball [¬p q], Knetklumpen [¬p ¬q]

„Ich lese dir Pauls Vermutung, welche Dinge springen, einmal vor:

Er vermutet: „Dinge, die weich sind, springen.“

„Die Vermutung wollen wir jetzt überprüfen.“

„Dazu nehme ich diese Wolle hier aus der Kiste.“ (*Wolle zeigen*). „Das Besondere an der Wolle ist, dass sie weich ist. Gut. Jetzt probieren wir aus, ob die Wolle springt.“

„Schau genau hin.“ (*Wolle fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten.*) „Ja, richtig die Wolle ist weich und springt.“

„Jetzt denke nochmal an Pauls Vermutung. Paul vermutet ja: „Dinge, die weich sind, springen. Was zeigt dir die Wolle über Pauls Vermutung?“

„Wenn du glaubst, Pauls Vermutung stimmt, legst du die Wolle hierhin“ (*auf den Smiley zeigen.*)

„Wenn du glaubst, Pauls Vermutung hat damit nichts zu tun, legst du die Wolle hierhin.“ (*auf den Smiley zeigen.*)

„Wenn du glaubst, Pauls Vermutung stimmt nicht, legst du die Wolle hierhin.“ (*auf den Smiley zeigen.*)



*Kind legt Wolle auf einen Smiley.*

„Jetzt überprüfen wir Pauls Vermutung mit einem anderen Ding.“

Er vermutet: „Dinge, die weich sind, springen.“

„Dazu nehme ich diesen Jonglierball hier aus der Kiste (*Jonglierball zeigen*). Das Besondere am Jonglierball ist, dass er weich ist. Gut. Jetzt probieren wir aus, ob der Jonglierball springt.“

„Schau genau hin.“ (*Jonglierball fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig der Jonglierball ist weich und springt nicht.“

„Jetzt denke nochmal an Pauls Vermutung. Paul vermutet: „Dinge, die weich sind, springen. Was zeigt dir der Jonglierball über Pauls Vermutung?“

„Wenn du glaubst, Pauls Vermutung stimmt, legst du den Jonglierball hierhin.“ (*auf den lachenden Smiley zeigen*.)

„Wenn du glaubst, Pauls Vermutung hat damit nichts zu tun, legst du den Jonglierball hierhin.“ (*auf den nachdenklichen Smiley zeigen*.)

„Wenn du glaubst, Pauls Vermutung stimmt nicht, legst du den Jonglierball hierhin.“ (*auf den traurigen Smiley zeigen*.)

*Kind legt den Jonglierball auf einen Smiley.*

Jetzt überprüfen wir Pauls Vermutung mit einem anderen Ding.

Er vermutet: „Dinge, die weich sind, springen.“

„Dazu nehme ich diesen Golfball.“ (*Golfball zeigen*). „Das Besondere an dem Golfball ist, dass er hart ist. Gut. Jetzt probieren wir aus, ob der Golfball springt.“

„Schau genau hin.“ (Golfball *fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig der Golfball ist hart und springt.“

„Jetzt denke nochmal an Pauls Vermutung. Paul vermutet ja: „Dinge, die weich sind, springen.“ Was zeigt dir der Golfball über Pauls Vermutung?“

„Wenn du glaubst, Pauls Vermutung stimmt, legst du den Golfball hierhin.“ (*Auf den lachenden Smiley zeigen.*)

„Wenn du glaubst, Pauls Vermutung hat damit nichts zu tun, legst du den Golfball hierhin.“ (*Auf den nachdenklichen Smiley zeigen.*)

„Wenn du glaubst, Pauls Vermutung stimmt nicht, legst du Golfball hierhin.“ (*Auf den traurigen Smiley zeigen.*)

*Kind legt den Golfball auf einen Smiley.*

„Jetzt wollen wir Pauls Vermutung mit einem anderen Ding überprüfen.“

Er vermutet. „Dinge, die weich sind, springen.“

„Dazu nehme ich den trockenen Knetklumpen hier aus der Kiste.“ (*Roten Knetklumpen zeigen*).

„Das Besondere an dem Knetklumpen ist, dass es hart ist. Gut. Jetzt probieren wir aus, ob der der Knetklumpen springt.“

„Schau genau hin.“ (Knetklumpen *fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig der Knetklumpen ist hart und springt nicht.

Jetzt denke nochmal an Pauls Vermutung. Paul vermutet ja: „Dinge, die weich sind, springen.“ Was zeigt dir der Knetklumpen über Pauls Vermutung?“

„Wenn du glaubst, Pauls Vermutung stimmt, legst du den Knetklumpen hierhin.“ (*Auf den lachenden Smiley zeigen.*)

„Wenn du glaubst, Pauls Vermutung hat damit nichts zu tun, legst du den Knetklumpen hierhin.“ (*Auf den nachdenklichen Smiley zeigen.*)

*„Wenn du glaubst, Pauls Vermutung stimmt nicht, legst du den Knetklumpen hierhin.“ (Auf den traurigen Smiley zeigen.)*

*Kind legt den Knetklumpen auf einen Smiley.*

Das ist Tim.

## 2. „Dinge, die rund sind, springen.“

Flummi [ $p \ q$ ], Luftballon (nicht aufgeblasen) [ $\neg p \ \neg q$ ], Überraschungsei [ $\neg p \ q$ ], Knetkugel [ $p \ \neg q$ ],

„Ich lese Tims Vermutung, welche Dinge springen, einmal vor:

„Dinge, die rund sind, springen.“

„Jetzt wollen wir Tims Vermutung „Dinge, die rund sind, springen.“ überprüfen.“

„Dazu nehme ich diesen Flummi hier aus der Kiste.“ (*Flummi zeigen*). „Das Besondere an ihm ist, dass er rund ist. Gut. Jetzt probieren wir aus, ob der Flummi springt.“

„Schau genau hin.“ (*Flummi fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig. Der Flummi springt.“

„Der Flummi ist rund und springt.“

„Jetzt denke nochmal an Tim und seine Vermutung. Tim vermutet ja: „Dinge, die rund sind, springen.“ Was zeigt dir der Flummi über Tims Vermutung?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du den Flummi hierhin.“ (*Auf den lachenden Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun legst du den Flummi hierhin.“ (*Auf den nachdenklichen Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du den Flummi hierhin.“ (*Auf den traurigen Smiley zeigen*.)

*Kind legt Flummi auf einen Smiley.*

Jetzt wollen wir Tims Vermutung mit einem anderen Ding überprüfen.

Er vermutet. „Dinge, die rund sind, springen.“

„Um Tims Vermutung zu überprüfen, nehmen wir nun einen nicht aufgeblasenen Luftballon.“ (*Luftballon zeigen*).

„Das Besondere an dem nicht aufgeblasenen Luftballon ist, dass er nicht rund ist. Gut. Jetzt probieren wir aus, und schauen, ob das der Luftballon springt. Schau genau hin.“ (*Luftballon fallen lassen*). (*Antwort des Kindes abwarten*).

„Der Luftballon springt nicht.“

„Denk jetzt noch einmal an Tims Vermutung.“

„Er vermutet ja: „Dinge, die rund sind, springen.“

„Was zeigt dir der Luftballon über Tims Vermutung?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du den nicht aufgeblasenen Luftballon hierhin.“ (*Auf den lachenden Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun legst du den nicht aufgeblasenen Luftballon hierhin.“ (*Auf den nachdenklichen Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du den nicht aufgeblasenen Luftballon hierhin.“ (*Auf den traurigen Smiley zeigen*.)

*Kind legt den Luftballon auf einen Smiley.*

„Jetzt wollen wir Tims Vermutung mit einem anderen Ding überprüfen.“

„Er vermutet. „Dinge, die rund sind, springen.“

„Dazu nehmen wir nun ein Überraschungsei“ (*Überraschungsei zeigen*).

„Das Besondere an dem Überraschungsei ist, dass es nicht rund ist.“

„Gut. Jetzt probieren wir aus und schauen, ob das Überraschungsei springt. Schau genau hin.“ (*Überraschungsei fallen lassen*). „Was passiert?“ (*Antwort des Kindes abwarten*).

„Das Überraschungsei springt.“

„Das Überraschungsei ist nicht rund und springt.“

„Denk jetzt noch einmal an Tims Vermutung.“

„Er vermutet ja: „Dinge, die rund sind, springen.“

„Was zeigt dir das Überraschungsei über Tims Vermutung?“

„Zeigt das Überraschungsei, dass Tims Vermutung stimmt, hat das Überraschungsei gar nichts mit Tims Vermutung zu tun, oder zeigt das Überraschungsei, dass die Vermutung nicht stimmt?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du das Überraschungsei hierhin.“ (*Auf den lachenden Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun, legst du das Überraschungsei hierhin.“ (*Auf den nachdenklichen Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du das Überraschungsei hierhin.“ (*Auf den traurigen Smiley zeigen*.)

*Kind legt Überraschungsei auf einen Smiley.*

„Dazu nehme ich diese Knetkugel hier aus der Kiste.“ (*Knetkugel zeigen*). „Das Besondere an der Kugel ist, dass sie rund ist. Gut. Jetzt probieren wir aus, ob die Knetkugel springt.

Schau genau hin.“ (*Knetkugel fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig. Die Knetkugel ist rund und springt nicht.“

„Jetzt denke nochmal an Tim und seine Vermutung. Tim vermutet ja: „Dinge, die rund sind, springen.“ Was zeigt dir Knetkugel über Tims Vermutung?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du den Knetkugel hierhin.“ (*Auf den lachenden Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun, legst du die Knetkugel hierhin.“ (*Auf den nachdenklichen Smiley zeigen*.)

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du die Knetkugel hierhin.“ (*Auf den traurigen Smiley zeigen*).

*Kind legt Knetkugel auf einen Smiley.*

### 3. „Dinge, die schwer sind, springen.“

Federsäckchen [ $\neg p \neg q$ ], Eisenkugel [ $p q$ ], Stein [ $p \neg q$ ], Luftballon [ $\neg p q$ ] (aufgeblasen)

„Das ist Petra.“

Ich lese dir Petras Vermutung, welche Dinge springen, einmal vor“:

„Dinge, die schwer sind, springen.“

„Jetzt wollen wir Petras Vermutung überprüfen.“

„Dazu nehme ich dieses Federsäckchen hier aus der Kiste.“ (*Federsäckchen zeigen*). „Das Besondere an dem Federsäckchen ist, dass es leicht ist. Gut. Jetzt probieren wir aus, ob das Federsäckchen springt.“

„Schau genau hin.“ (*Federsäckchen fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig das Federsäckchen ist leicht und springt nicht.“

„Jetzt denke nochmal an Petras Vermutung. Petra vermutet ja: „Dinge, die schwer sind, springen.“ Was zeigt dir das Federsäckchen über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du das Federsäckchen hierhin.“ (*Auf den lachenden Smiley zeigen*.)

„Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du das Federsäckchen hierhin.“ (*Auf den nachdenklichen Smiley zeigen*).

„Wenn du glaubst, Petras Vermutung stimmt nicht, legst du das Federsäckchen *hierhin*.“ (*Auf den traurigen Smiley zeigen*.)

*Kind legt das Federsäckchen auf einen Smiley.*



„Jetzt wollen wir Petras Vermutung noch mit einem anderen Ding überprüfen.“

„Sie vermutet: „Dinge, die schwer sind, springen.“

„Dazu nehme ich diese Eisenkugel hier aus der Kiste.“ (*Eisenkugel zeigen*). „Das Besondere an der Eisenkugel ist, dass sie schwer ist. Gut. Jetzt probieren wir aus, ob die Eisenkugel springt.“

„Schau genau hin.“ (*Eisenkugel fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig die Eisenkugel ist schwer und springt.“

Jetzt denke nochmal an Petras Vermutung. Petra vermutet ja: „Dinge, die schwer sind, springen.“ Was zeigt dir die Eisenkugel über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du die Eisenkugel hierhin.“ (*Auf den lachenden Smiley zeigen.*)

„Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du die Eisenkugel hierhin.“ (*Auf den nachdenklichen Smiley zeigen*)

„Wenn du glaubst, Petras Vermutung stimmt nicht, legst du die Eisenkugel hierhin.“ (*Auf den traurigen Smiley zeigen.*)

*Kind legt die Eisenkugel auf einen Smiley.*

„Jetzt wollen wir Petras Vermutung noch mit einem anderen Ding überprüfen.“

„Sie vermutet: „Dinge, die schwer sind, springen.“

„Dazu nehme ich diesen Stein hier aus der Kiste.“ (*Stein zeigen*). „Das Besondere an dem Stein ist, dass er schwer ist. Gut. Jetzt probieren wir aus, ob der Stein springt.“

„Schaut genau hin.“ (*Stein fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig. Der Stein ist schwer und springt nicht.“

„Jetzt denke nochmal an Petra und ihre Vermutung.“ „Petra vermutet ja: „Dinge, die schwer sind, springen.“ „Was zeigt dir der Stein über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du den Stein hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du den Stein hierhin.“ (*Auf den Smiley zeigen*).

„Wenn du glaubst, Petras Vermutung stimmt nicht, legst du den Stein hierhin.“ (*Auf Smiley zeigen*.)

*Kind legt den Stein auf einen Smiley.*

„Jetzt wollen wir Petras Vermutung noch mit einem anderen Ding überprüfen.“

„Sie vermutet: „Dinge, die schwer sind, springen.“

„Dazu nehme ich diesen aufgeblasenen Luftballon hier aus der Kiste.“ (*Luftballon zeigen*).

„Das Besondere an dem aufgeblasenen Luftballon ist, dass er leicht ist. Gut. Jetzt probieren wir aus, ob der aufgeblasene Luftballon springt.“

„Schau genau hin.“ (*Luftballon fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig, der Luftballon ist leicht und springt.“

„Jetzt denke nochmal an Petra und ihre Vermutung.“ „Petra vermutet ja: „Dinge, die schwer sind, springen.“ „Was zeigt dir der Luftballon über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du den Luftballon hierhin.“ (*Auf den Smiley zeigen*.)

„Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du den Luftballon hierhin.“ (*Auf den Smiley zeigen*).

„Wenn du glaubst, Petras Vermutung stimmt nicht, legst du den Luftballon hierhin.“ (*Auf den Smiley zeigen*.)

*Kind legt den Luftballon auf einen Smiley.*

## Protokollbogen Laborstudie 1. Trainingssitzung (Scaffolding nur mit adaptivem Nachfragen)

Geschlecht:    m w

ID:   

In den einzelnen Spalten wird jeweils angekreuzt wie das Kind antwortet: **Richtig (r)**, **falsch (f)**; 3. Und 7. Wird in Stichpunkten protokolliert

Vermutung: „Dinge, die hart sind, springen“		1. Eigen- schaft			2. Spru- ngver- halten			3. Zusam- menfas- sung			4. Jens			5. Zuord- nung Ball			6. Zusammen- fassung			7. Wahl des Kindes	
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b			
a) Knetkugel [p-q]	Der Murmel ist hart und springt. [p q]																			☺	☺
b) Softball [¬p q]	Die Knetk. ist hart und springt nicht. [p ¬q]																			☺	☺
c) Murmel [p q]																				☺	☺
d) Jonglierball [¬p ¬q]	Der Softball ist weich und springt [¬p q]																			☺	☺
	Der Jonglierball ist weich und springt nicht. [¬p ¬q]																			☺	☺

## Protokollbogen Laborstudie Intervention II (ohne Modeling)

Geschlecht:     m w            ID:

In den einzelnen Spalten wird jeweils angekreuzt wie das Kind antwortet: **Richtig (r)**, **falsch (f)**; 3. Und 7. Wird in Stichpunkten protokolliert

Vermutung: „Dinge, die mit Luft sind, springen“		1. Eigen- schaft			2. Spru- ngver- halten			3. Zusam- menfas- sung			4. Jens			5. Zuord- nung Ball			6. Zusammen- fassung			7. Wahl des Kindes	
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b
a) Wasserball [p -q]	Der Spielball ist mit Luft und springt. [p q]																			☺	☺
b) Spielball [p q]	Der Wasserball ist mit Luft und springt nicht. [p -q]																			☺	☺
c) Würfel [-p q]	Der Würfel ist ohne Luft und springt [¬p q]																			☺	☺
d) Aufg. Luftballon [-p -q]	Der aufgeschnittene Luftballon ist ohne Luft und springt nicht. [¬p -q]																			☺	☺



**Interviewleitfaden**  
**1. Trainingsitzung**  
**(Scaffolding nur mit**  
**adaptivem Nachfragen)**

**Scaffolding nur mit**  
**adaptivem**  
**Nachfragen**

Vermutung „Dinge, die hart sind, springen.“

(Knetkugel [ $p \neg q$ ]; Softball [ $\neg p q$ ]; Murmel [ $p q$ ]; Jonglierball)

Explizieren des Gedankenganges komplett

1. Sortieren der Gegenstände
  - nach Eigenschaften
  - nach Sprungverhalten
  
2. Präsentieren der Vermutung
  - Spielfigur einführen
  
3. Überprüfen der Vermutung
  - Kind trifft Entscheidung selbst

# Sortieren der Gegenstände nach Eigenschaften

Jens Vermutung: „Dinge, die hart sind, springen.“

(trockener Knetkugel [p -q]; Softball [-p q]; Murmel [p q]; Jonglierball [-p -q])

„Schau mal, hier sind vier Dinge: ein Softball, eine Murmel, eine trockene Knetkugel und ein Jonglierball.“

*(Bälle zeigen)*

„Hier habe ich zwei Tafeln mitgebracht. Eine ist ganz hart *(auf Metallplatte zeigen)* und die andere ganz weich *(auf Platte mit Watte zeigen)*.

„Nun sortiere doch mal die vier Dinge nach hart und weich. Die harten Dinge legst du hier hin *(auf Metallplatte zeigen)* und die weichen Dinge legst du hier hin *(auf Platte mit Watte zeigen)*.

Kind sortieren lassen. [Richtige Lösung: Murmel, trockener Knetkugel hart; Softball, Jonglierball: weich]

Protokoll 1a)

Bei Fehler:

„Fühl doch noch einmal ganz genau. Ist das Ding hart oder weich?“

*(Kind anfassen lassen)*.

„Und wo würdest du es dann hinlegen?“

*So lange fragen, bis alle Dinge auf der richtigen Platte liegen.*

Protokoll 1b)

Protokoll 1c)



# Sortieren der Gegenstände nach Sprungverhalten

„Nun wollen wir mal sehen, welche dieser Dinge springen und welche nicht springen können.“

„Dafür habe ich hier noch etwas mitgebracht. Diese Karte bedeutet, dass das Ding springen kann. (*Karte für „springt“ zeigen*) und diese bedeutet, dass das Ding nicht springen kann, also nur fällt und auf liegen bleibt. (*Karte für springt nicht zeigen*).“  
(*Symbolkarten für springt und springt nicht auf jede Platte legen*).

„Nun lass die Dinge mal fallen, um zu sehen, ob sie springen können oder nicht und sortiere die vier Dinge danach, ob sie springen oder nicht.“

„Die Dinge, die hart sind und springen legst du hier hin.“ (Symbolkarte „springt“ im Feld „hart“), die Dinge, die hart sind und nicht springen legst du hier hin (Symbolkarte „springt nicht“ im Feld „hart“).

„Die Dinge, die weich sind und springen kommen hier hin (Symbolkarte „springt“ im Feld „weich“) und die Dinge, die weich sind und nicht springen legst du hier hin (Symbolkarte „springt nicht“ im Feld „weich“).

(*Kind sortieren lassen*). [richtige Lösung: Murmel, Knetkugel hart, springt nicht; Softball, Jonglierball: weich, springt nicht]

## Protokoll 2a)

Bei Fehler:

„Wir haben ja eben schon gesehen, dass dieses Ding hart / weich ist. Und nun sag mir doch noch mal wo denn die harten / weichen Dinge hinkommen? (falls schon gesagt: Deswegen haben wir es ja auf das harte / weiche Feld gelegt.) und nun probiere doch noch einmal aus, ob das Ding springt oder nicht. (*Kind ausprobieren lassen*) Wo würdest du es jetzt hinlegen?“

*So lange fragen, bis alle Dinge auf der richtigen Platte liegen*

# Zusammenfassung

„Jetzt wissen wir, welche Dinge hart und welche weich sind.“  
(Bälle wegnehmen).

„Und jetzt sag mir doch nochmal warum du die Bälle auf die jeweiligen Felder gelegt hast. Also warum liegt zum Beispiel die Knetkugel dort?“

[auf Murmel bzw. Feld zeigen]

(Kind antworten lassen).

Protokoll 3a)

„Und warum liegt der Softball dort?“

(Kind antworten lassen).

Protokoll 3b)

„Was ist mit der Murmel?“

(Kind antworten lassen).

Protokoll 3c)

„Und warum liegt der Jonglierball auf diesem Feld?“

(Kind antworten lassen).

Protokoll 3d)

„Jetzt wissen wir, welche harten Dinge springen und welche nicht. Wir wissen welche weichen Dinge springen und welche nicht.“

(Bälle wegnehmen).

## Erklärung: Was ist eine Vermutung?

„Wir wollen heute verschiedene Vermutungen überprüfen.“

Eine Vermutung ist etwas, das ich zwar glaube, aber noch überprüfen muss.“

*(Säckchen hervorholen und zeigen).*

„Ich vermute, dass in dem Säckchen Orangen versteckt sind. Was glaubst du?“

*(Kind antworten lassen).*

„Wie können wir herausbekommen, ob deine oder meine Vermutung stimmt?“

*(Kind antworten lassen).*

Stummer Impuls: Säckchen mit Öffnung zum Kind halten *(Kind nachschauen und antworten lassen).*

„Ja richtig, im Säckchen ist ein Bär.“

„Welche Vermutung stimmt, deine oder meine Vermutung? Oder hat keiner von uns Recht?“ *(Kind antworten lassen).*

# Präsentation der Vermutung:

„Nun haben wir ja schon einige Kinder gefragt, was sie denken, warum ein Ball springt. Jens meint, dass Dinge, die hart sind, springen.“

„Das hier ist Jens.“ (*auf männliche Playmobilfigur zeigen*).

Wo würdest du Jens nun hinstellen? Jens meint ja, dass Dinge die hart sind, springen. Die Dinge, von denen er spricht, müssen also hart sein und springen können. [Richtige Antwort: „Auf die Symbolkarte „springt“ im Feld „hart“].

Protokoll 4a)

Bei Fehler:

„Spricht Jens von harten oder von weichen Dingen? Auf welches Feld würdest du Jens also stellen?“

Kind antworten lassen [richtige Antwort: hart].

„Und, meint Jens, dass diese Dinge springen können oder, dass sie nicht springen können?“

Kind antworten lassen [richtige Antwort: springen].

„Auf welches Feld musst du Jens also stellen?“

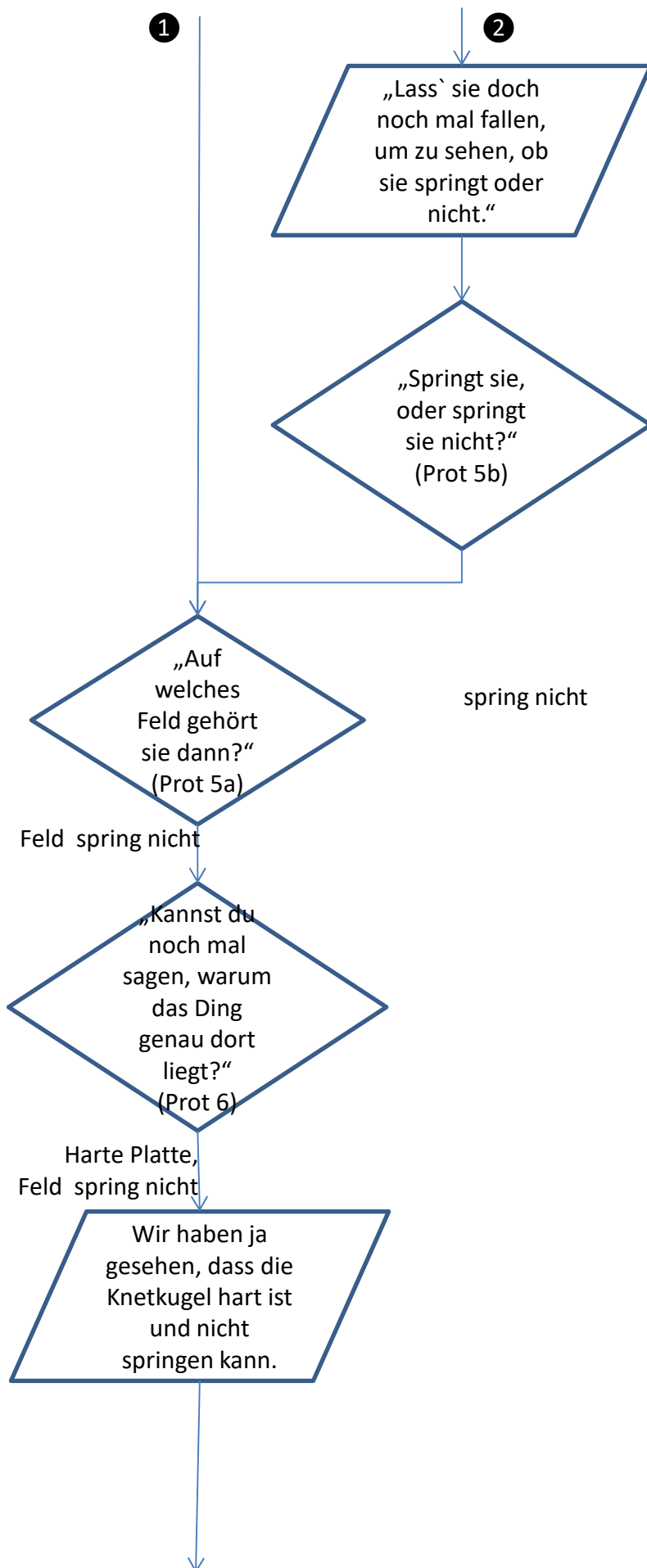
Kind antworten lassen [richtige Antwort: Feld „hart“, Symbolkarte „springt“].

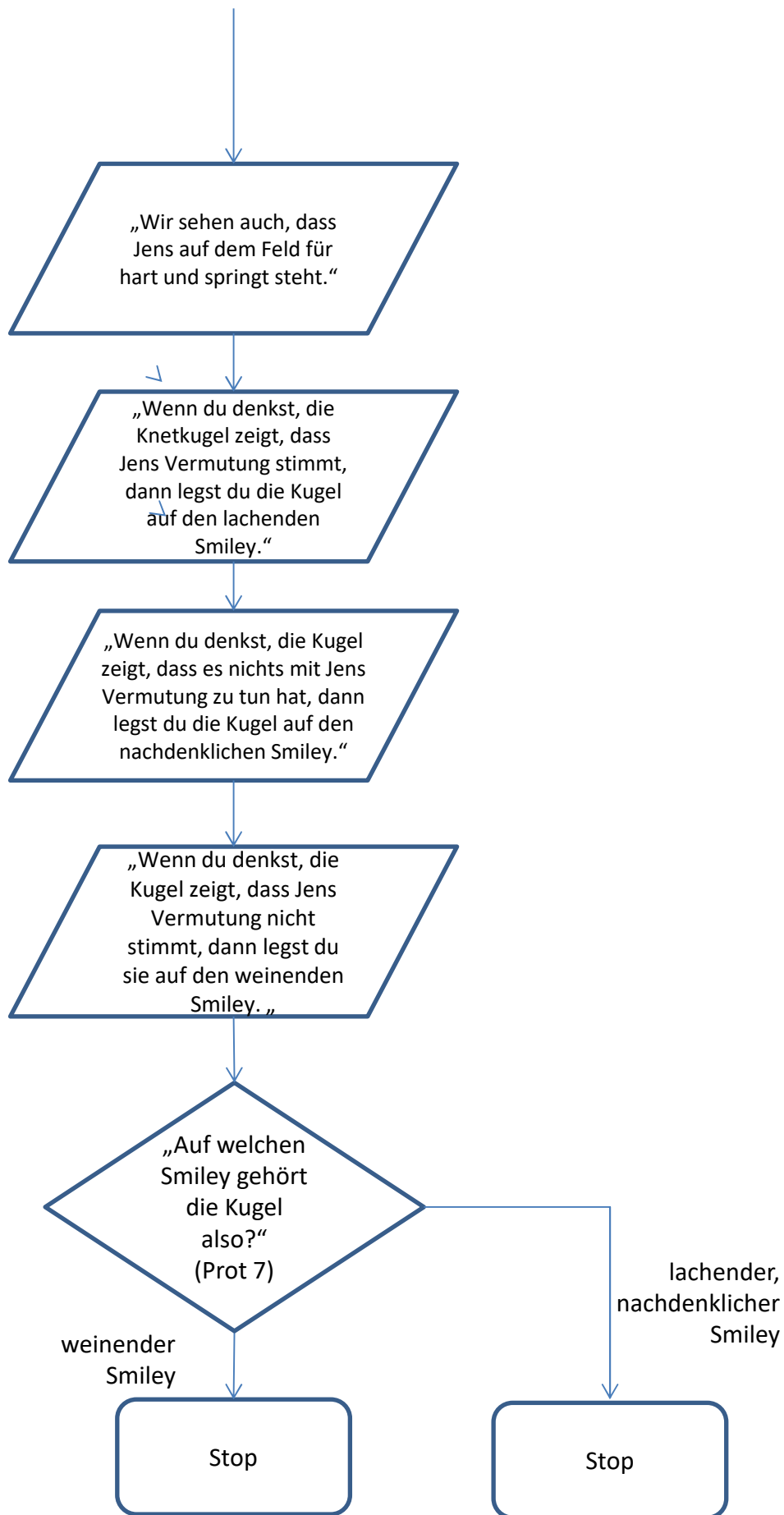
*Solange fragen, bis die Spielfigur auf dem richtigen Feld steht.*

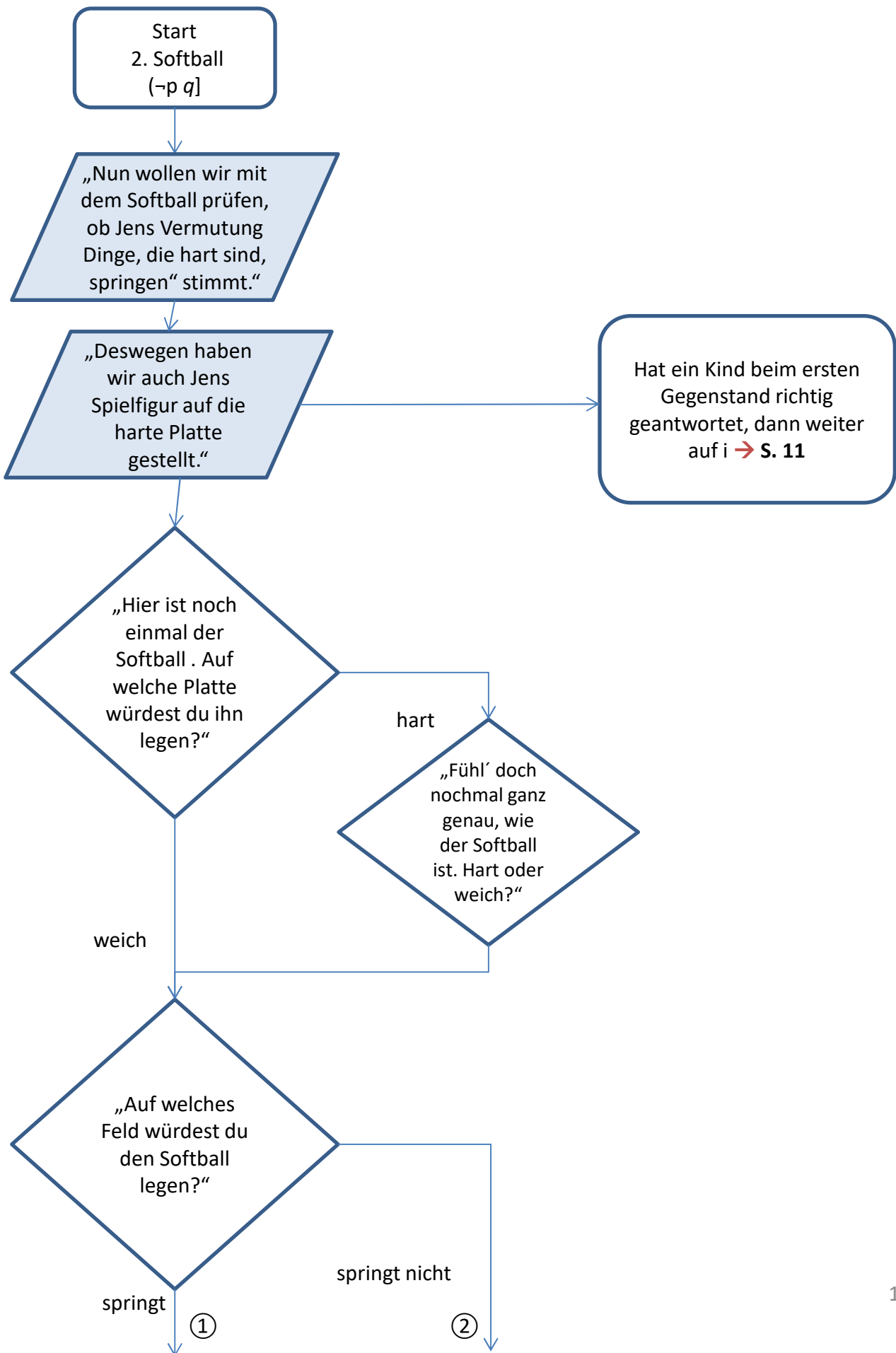
Protokoll 4b)

Protokoll 4c)

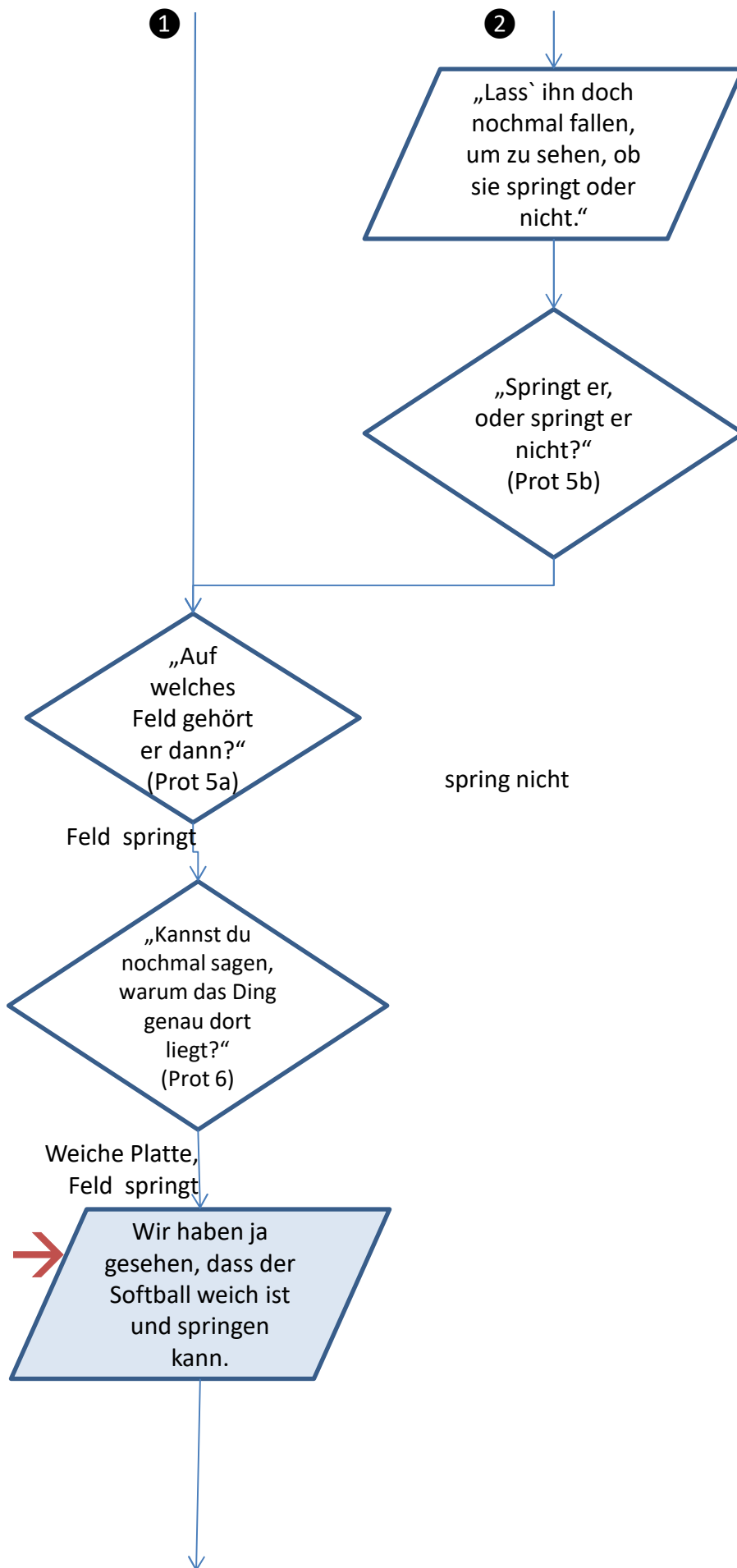


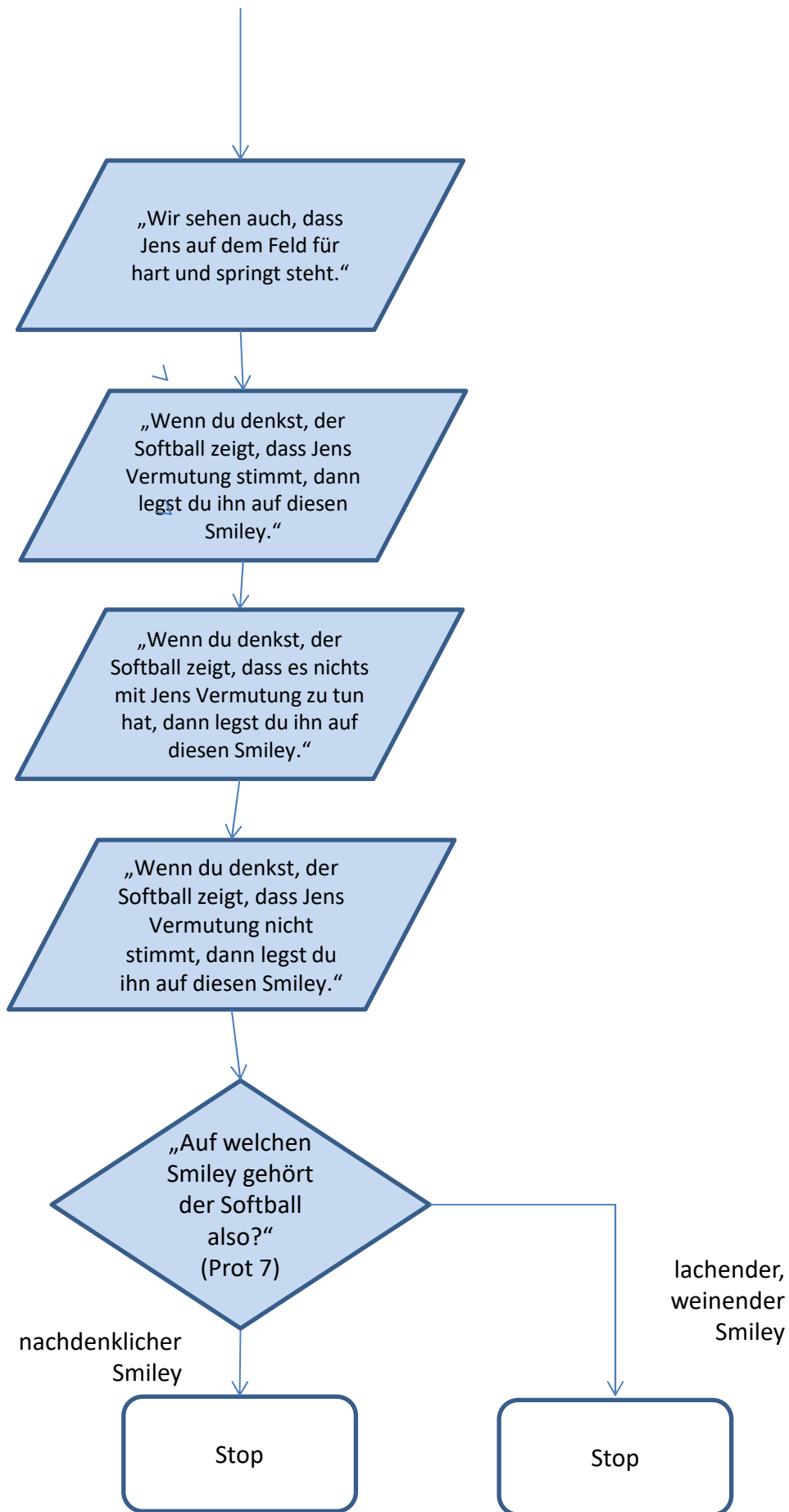


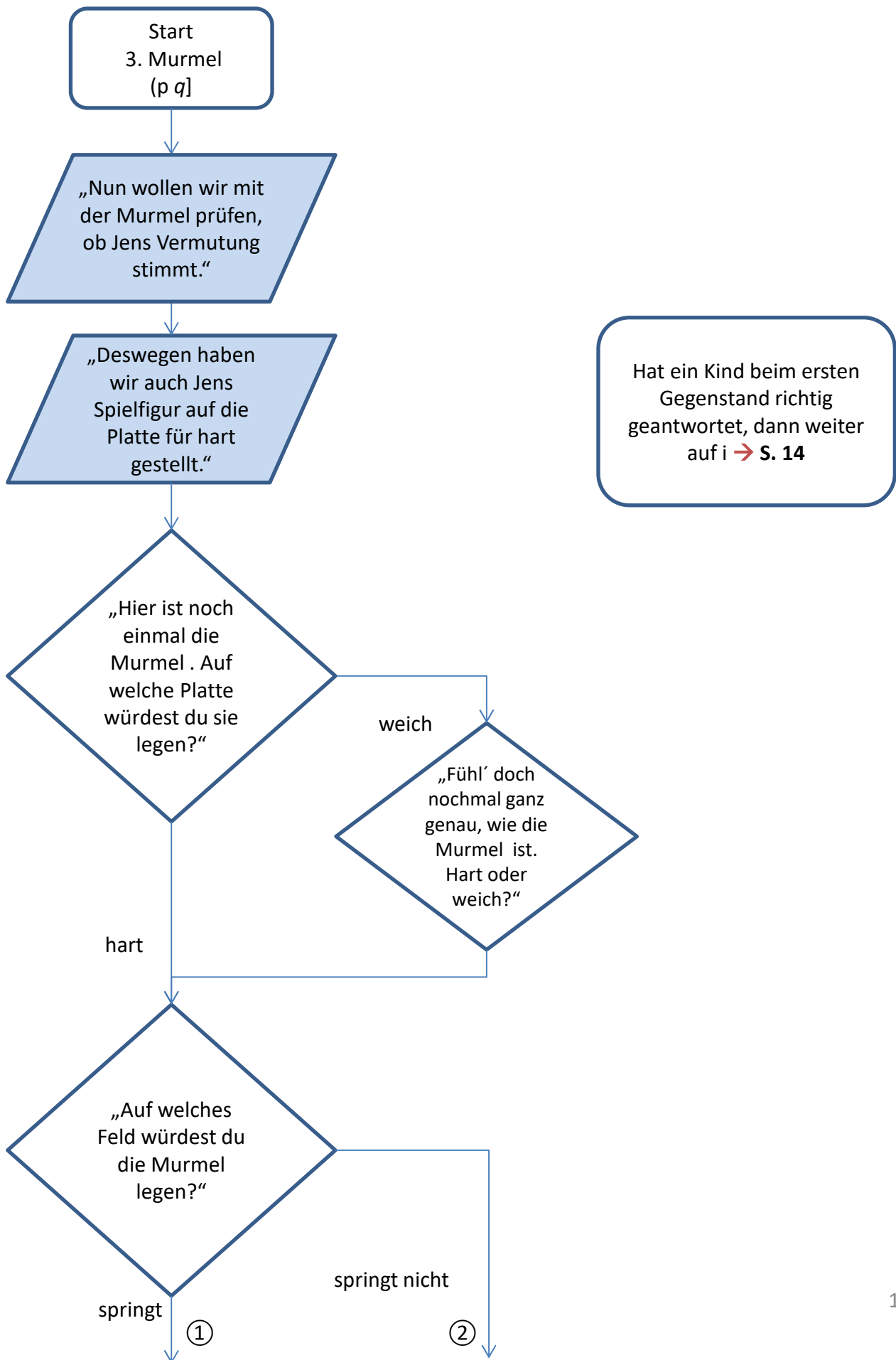


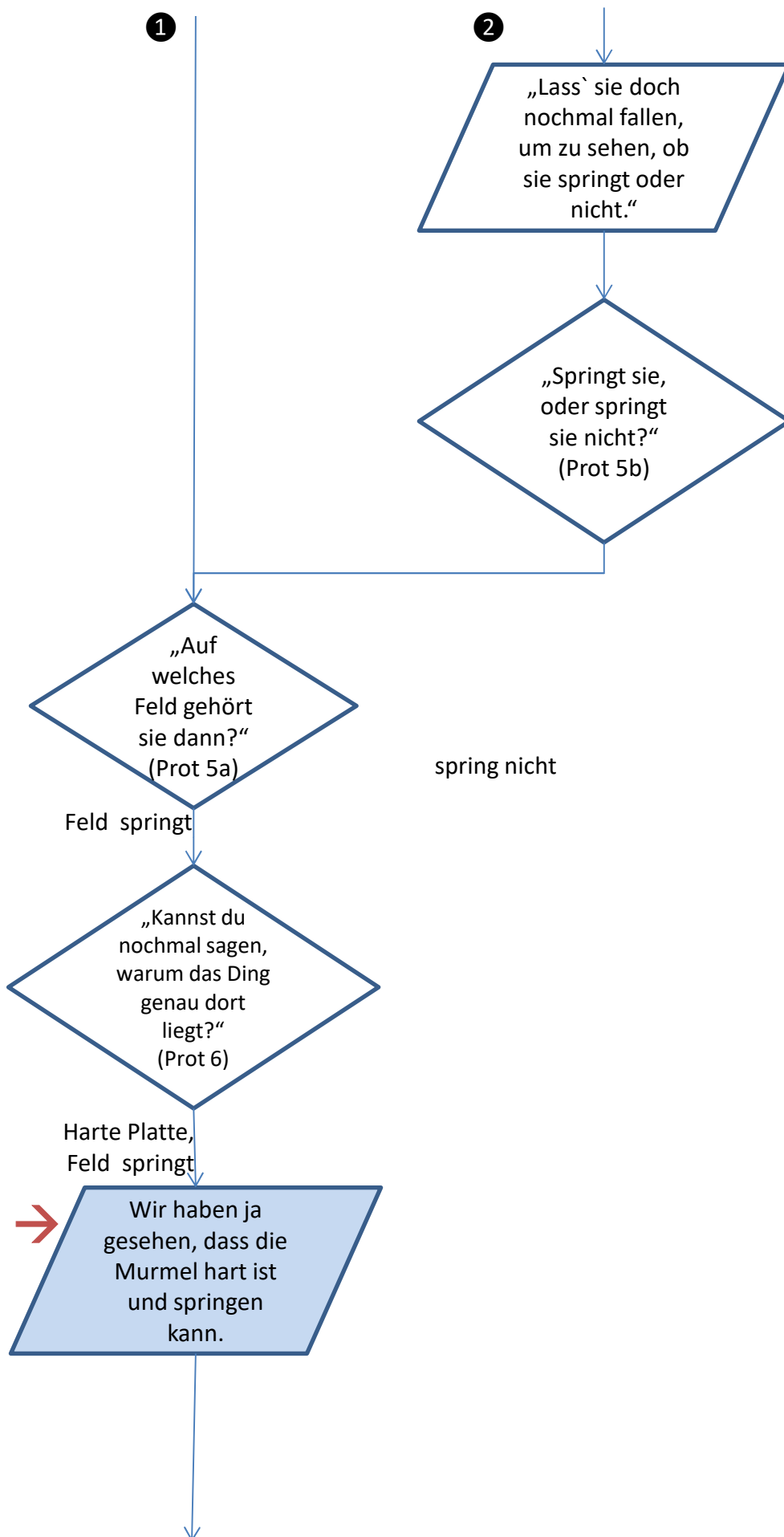


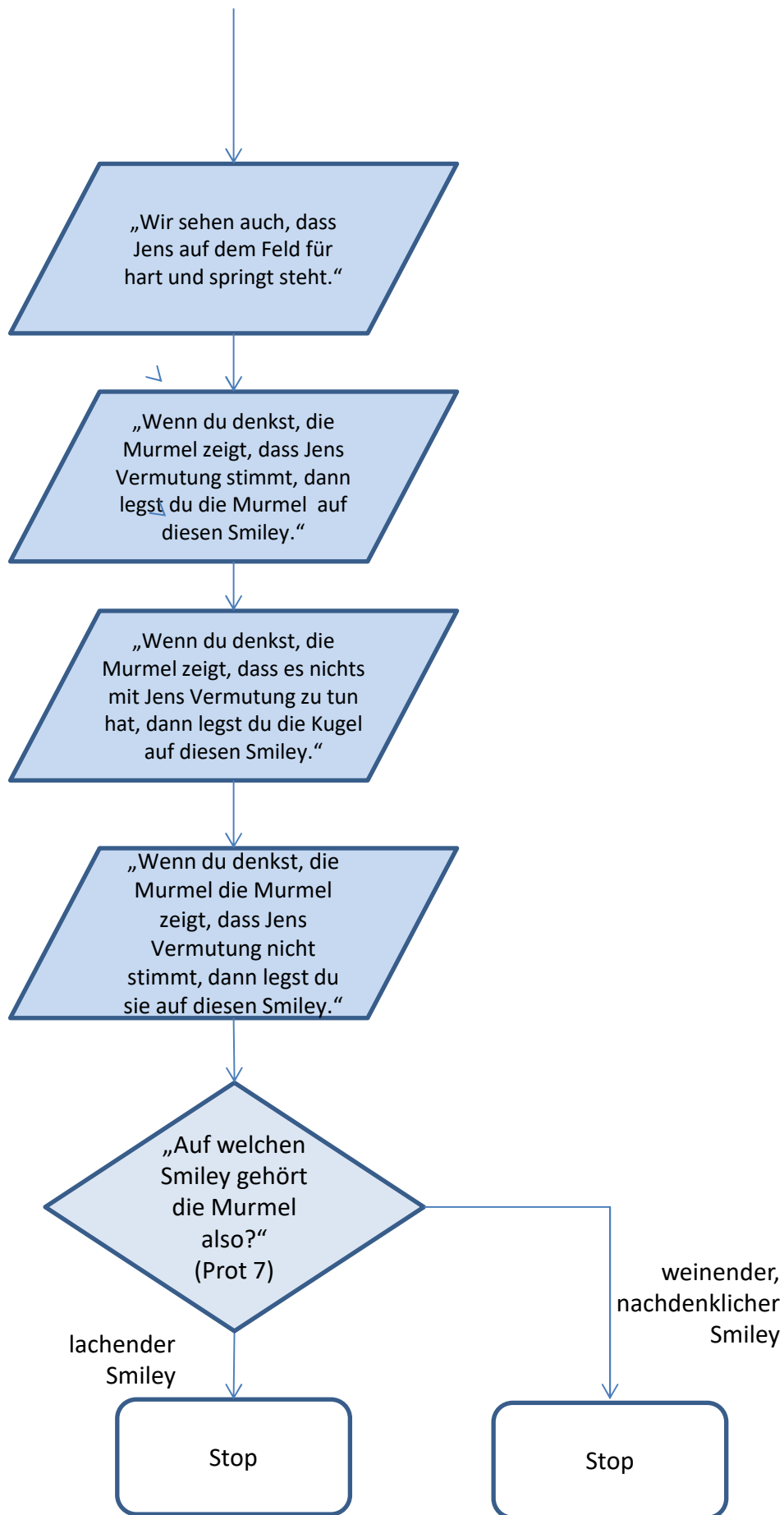


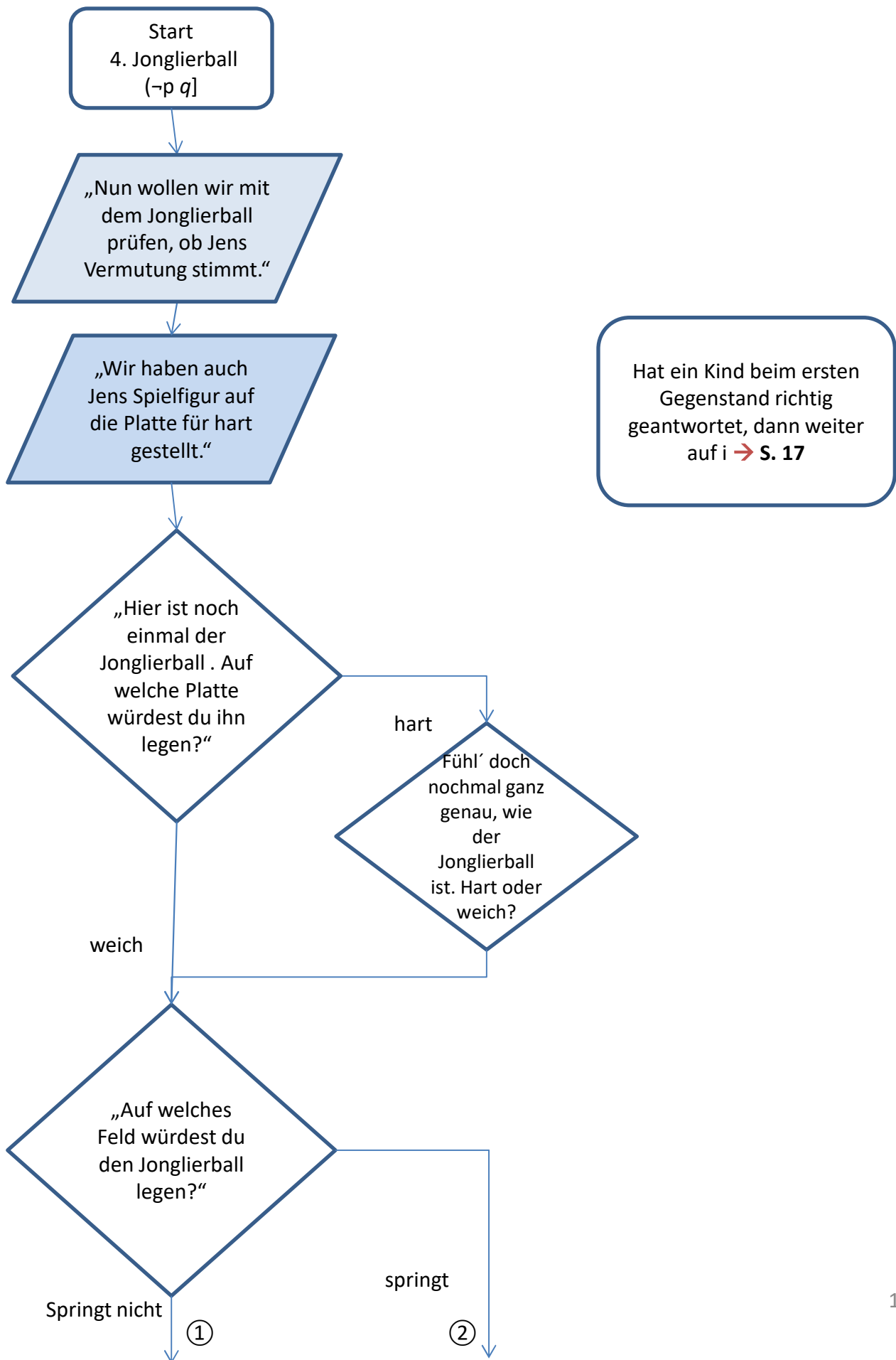




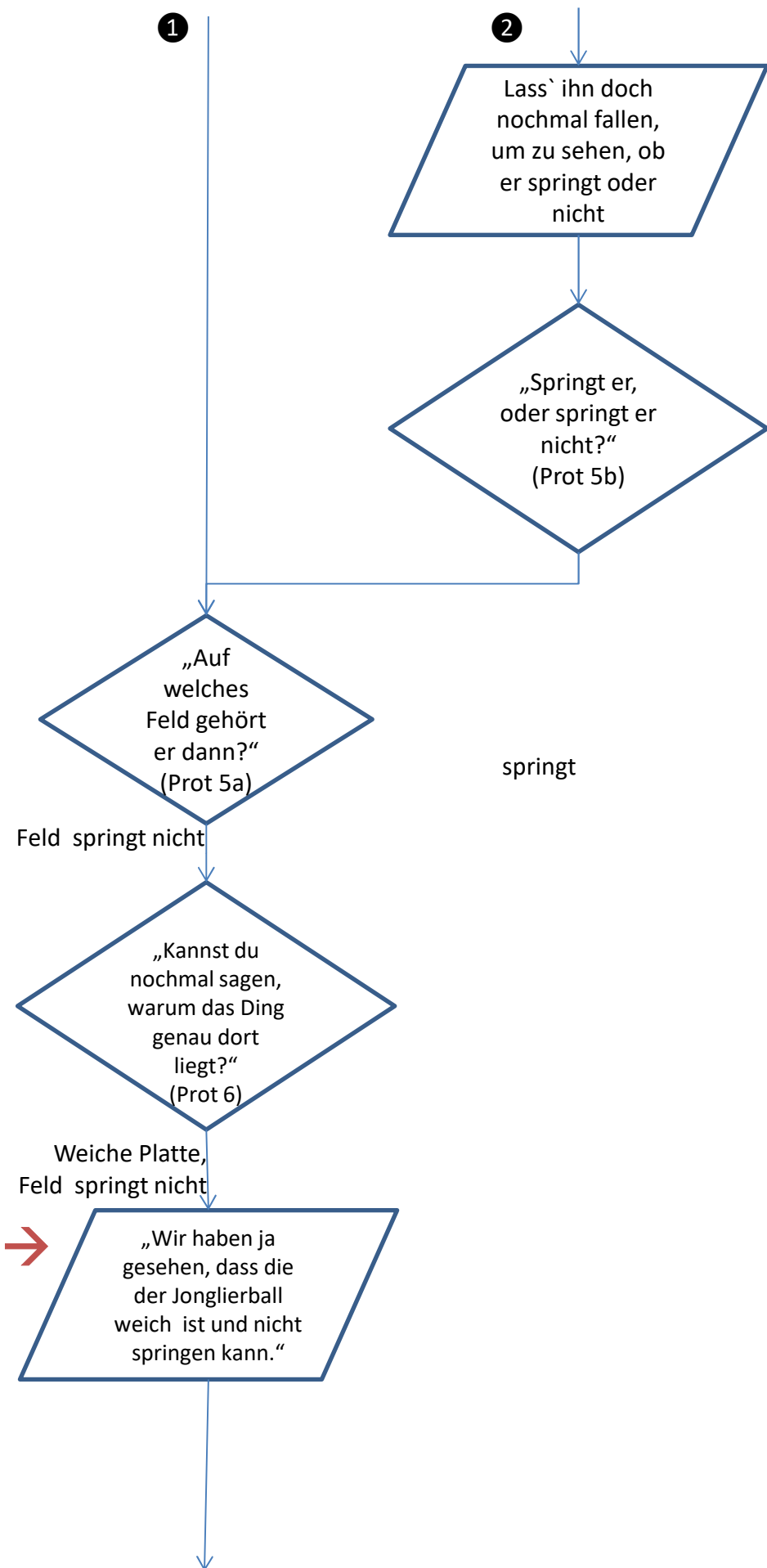


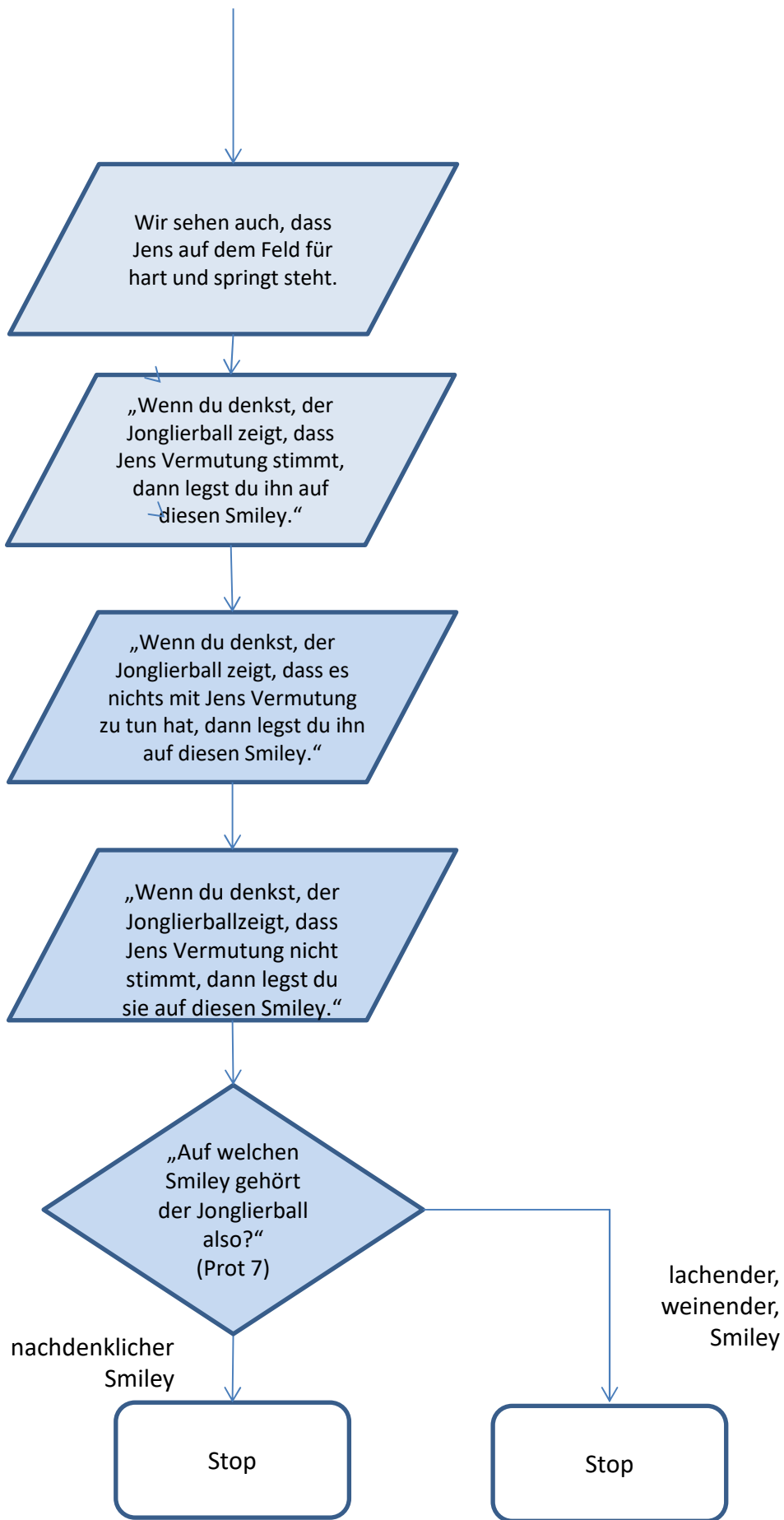






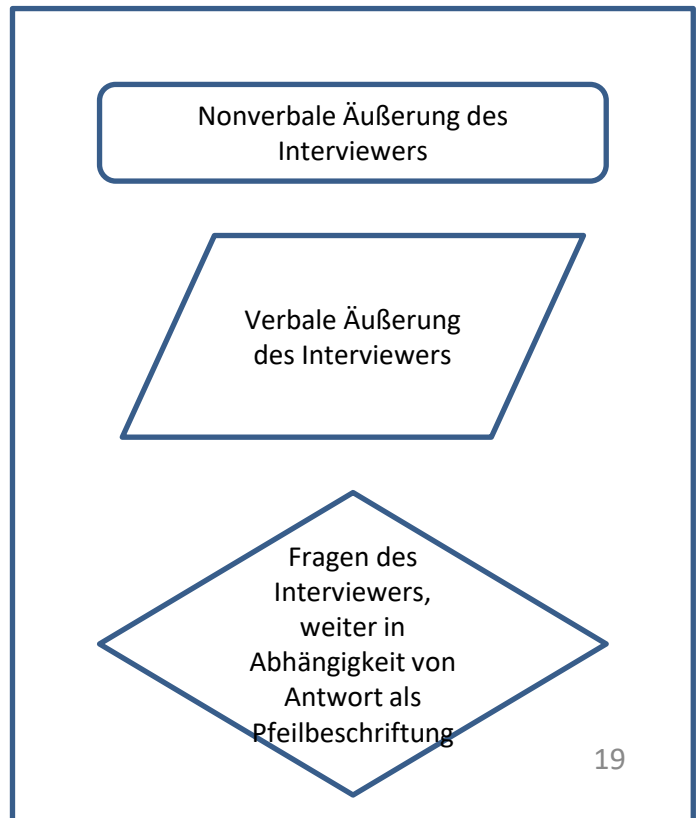
Hat ein Kind beim ersten Gegenstand richtig geantwortet, dann weiter auf i → **S. 17**







## Legende:



## Protokollbogen Laborstudie 2. Trainings Sitzung (Scaffolding mit adaptivem Nachfragen)

Geschlecht: m w

TN:

In den einzelnen Spalten wird jeweils angekreuzt wie das Kind antwortet: **Richtig (r)**, **falsch (f)**; 3. Und 7. Wird in Stichpunkten protokolliert

Vermutung: „Dinge, die mit Luft sind, springen“		1. Eigen- schaft			2. Sprungver- halten			3. Zusam- menfas- sung			4. Jens			5. Zuord- nung Ball			6. Zusammen- fassung			7. Wahl des Kindes	
		a	b	c	a	b	c	a	b	c	a	b	c	a	b	c	a	b	a	b	
a) Wasserball [p -q]	Der Tennisball ist mit Luft und springt. [p q]																			☺	☺
b) Spielball [p q]	Der Wasserball ist mit Luft und springt nicht. [p -q]																			☺	☺
c) Würfel [¬p q]	Der Würfel ist ohne Luft und springt [¬p q]																			☺	☺
d) Aufg. Luftballon [¬p ¬q]	Der aufgeschnittene Luftballon ist ohne Luft und springt nicht. [¬p ¬q]																			☺	☺

# Interviewleitfaden

## 2. Trainingsitzung (Scaffolding nur mit adaptivem Nachfragen)

### Scaffolding nur mit adaptivem Nachfragen

Vermutung „Dinge, die mit Luft gefüllt sind, springen.“  
(Wasserball mit wenig Luft [p -q]; Spielball[p q]; Würfel [-p q];  
aufgeschnittener Luftballon [-p -q];)

adaptives Nachfragen

1. Sortieren der Gegenstände
  - nach Eigenschaften
  - nach Sprungverhalten
  - adaptives Nachfragen bei Fehler
  
2. Präsentieren der Vermutung
  - Spielfigur einführen
  
3. Überprüfen der Vermutung
  - Kind entscheidet selbst

# Sortieren der Gegenstände nach Eigenschaften

„Schau´ mal, hier sind vier Dinge: ein aufgeschnittenen Luftballon, ein Spielball, ein aufblasbarer Wasserball und einen Würfel.“

*(Gegenstände zeigen).*

„Hier habe ich zwei kleine Luftkissen mitgebracht. Eins ist nicht aufgeblasen und flach, also ohne Luft *(auf nicht aufgeblasenes Luftkissen zeigen)*.

und eins aufgeblasen, also mit Luft *(auf aufgeblasenes Luftkissen zeigen)*.

„Nun sortiere doch mal die vier Dinge danach, ob sie mit Luft oder ohne Luft sind.

Die Dinge ohne Luft legst du hier hin *(auf nicht aufgeblasenes Luftkissen zeigen)* .

und die Dinge mit Luft legst du hier hin *(auf aufgeblasenes Luftkissen zeigen)*.

*(Kind sortieren lassen)*. [richtige Lösung: Spielball, aufblasbarer Wasserball: mit Luft; aufgeschnittener Luftballon, Würfel: ohne Luft].

## Protokoll 1a)

Bei Fehler:

„Schau doch noch einmal genau hin. Ist das Ding mit Luft oder ohne Luft?“

Kind anfassen lassen.

„Und wo würdest du es dann hinlegen?“

*So lange fragen, bis alle Dinge auf dem richtigen Kissen liegen*

## Protokoll 1b)

## Protokoll 1c)

„Jetzt wissen wir, welche Dinge mit Luft und welche ohne Luft sind.“

*(Bälle wegnehmen).*

# Sortieren der Gegenstände nach Sprungverhalten:

„Nun wollen wir mal sehen, welche dieser Dinge springen und welche nicht springen können.“

„Dafür habe ich hier noch etwas mitgebracht.“

*(Symbolkarten für springt und springt nicht auf jedes Luftkissen legen).*

„Nun lass die Dinge mal fallen um zu sehen, ob sie springen können oder nicht.“

*(Kind ausprobieren lassen).*

„Und nun sortiere‘ die vier Dinge danach, ob sie springen oder nicht.“

„Die Dinge, die mit Luft sind und springen legst du hier hin.“ (Symbolkarte „springt“ auf aufgeblasenem Luftkissen).

„Die Dinge, die mit Luft sind und nicht springen legst du hier hin.“  
(Symbolkarte „springt nicht“ auf aufgeblasenem Luftkissen).

„Die Dinge, die ohne Luft sind und springen kommen hier hin. (Symbolkarte „springt“ auf nicht- aufgeblasenem Luftkissen).

und die Dinge, die ohne Luft sind und nicht springen legst du hier hin.  
(Symbolkarte „springt nicht“ auf nicht- aufgeblasenem Luftkissen).

Kind sortieren lassen. [Richtige Lösung: Tennisball: mit Luft, springt; aufblasbarer Gegenstand: mit Luft, springt nicht; aufgeschnittener Luftballon: ohne Luft, springt nicht; Würfel: ohne Luft, springt].

Protokoll 2a)

Bei Fehler:

„Schau doch noch einmal genau hin. Ist das Ding mit Luft oder ohne Luft?“  
(Kind antworten lassen).

„Und nun probiere doch noch einmal aus, ob das Ding springt oder nicht.“  
(Kind ausprobieren lassen).

„Wo würdest du es jetzt hinlegen?“

So lange fragen, bis alle Dinge auf der richtigen Platte liegen.

# Zusammenfassung

„Und jetzt sag mir doch nochmal warum du die Bälle auf die jeweiligen Felder gelegt hast. Also warum liegt zum Beispiel der Wasserball dort?“

[auf Wasserball bzw. Feld zeigen]

*(Kind antworten lassen).*

Protokoll 3a)

„Und warum liegt der Spielball dort?“

*(Kind antworten lassen).*

Protokoll 3b)

„Was ist mit dem aufgeschnittene Luftballon?“

*(Kind antworten lassen).*

Protokoll 3c)

„Und warum liegt der Würfel auf diesem Feld?“

*(Kind antworten lassen).*

Protokoll 3d)

„Jetzt wissen wir, welche Dinge ohne Luft springen und welche nicht. Wir wissen welche Dinge mit Luft springen und welche nicht.“

# Erklärung Was ist eine Vermutung

„Wir wollen heute verschiedene Vermutungen überprüfen.“

„Eine Vermutung ist etwas, das ich zwar glaube, aber noch überprüfen muss.“  
(Kästchen hervorholen und zeigen).

„Ich vermute, dass in dem Kästchen ein Boxer versteckt ist. Was glaubst du?“  
(Kind antworten lassen).

„Wie können wir herausbekommen, ob deine oder meine Vermutung stimmt?“  
(Kind antworten lassen).

„Wir können doch einfach nachschauen, um festzustellen, ob deine oder meine Vermutung stimmt.“ (Kind nachschauen lassen).

„Welche Vermutung stimmt, deine oder meine Vermutung?“ (Kind antworten lassen). „Da ist eine Tänzerin drin, also stimmt meine Vermutung nicht.“  
(mögliche Antwort des Kindes „deine auch nicht“ oder „deine schon“).

# Präsentation der Vermutung

„Nun haben wir ja schon einige Kinder gefragt, was sie denken, warum ein Ball springt. Mia meint, dass Dinge die mit Luft sind, springen.“

„Das hier ist Mia.“ (auf weibliche Playmobilfigur zeigen).

„Wo würdest du Mia nun hinstellen? Mia meint ja, dass Dinge die mit Luft sind springen. Die Dinge von denen sie spricht müssen also mit Luft sein und springen können.“ [Richtige Antwort: „auf Kissen mit Luft“, Symbolkarte „springt“ ].

Protokoll 4a)

Bei Fehler:

„Spricht Mia von Dingen mit Luft oder ohne Luft? Auf welches Kissen würdest du Mia also stellen?“

*Kind antworten lassen* [Richtige Antwort: auf das aufgeblasene Luftkissen].

„Und, meint Mia, dass diese Dinge springen können oder, dass sie nicht springen können?“

*(Kind antworten lassen)*. [Richtige Antwort: springen].

„Auf welches Feld musst du Mia also stellen?“

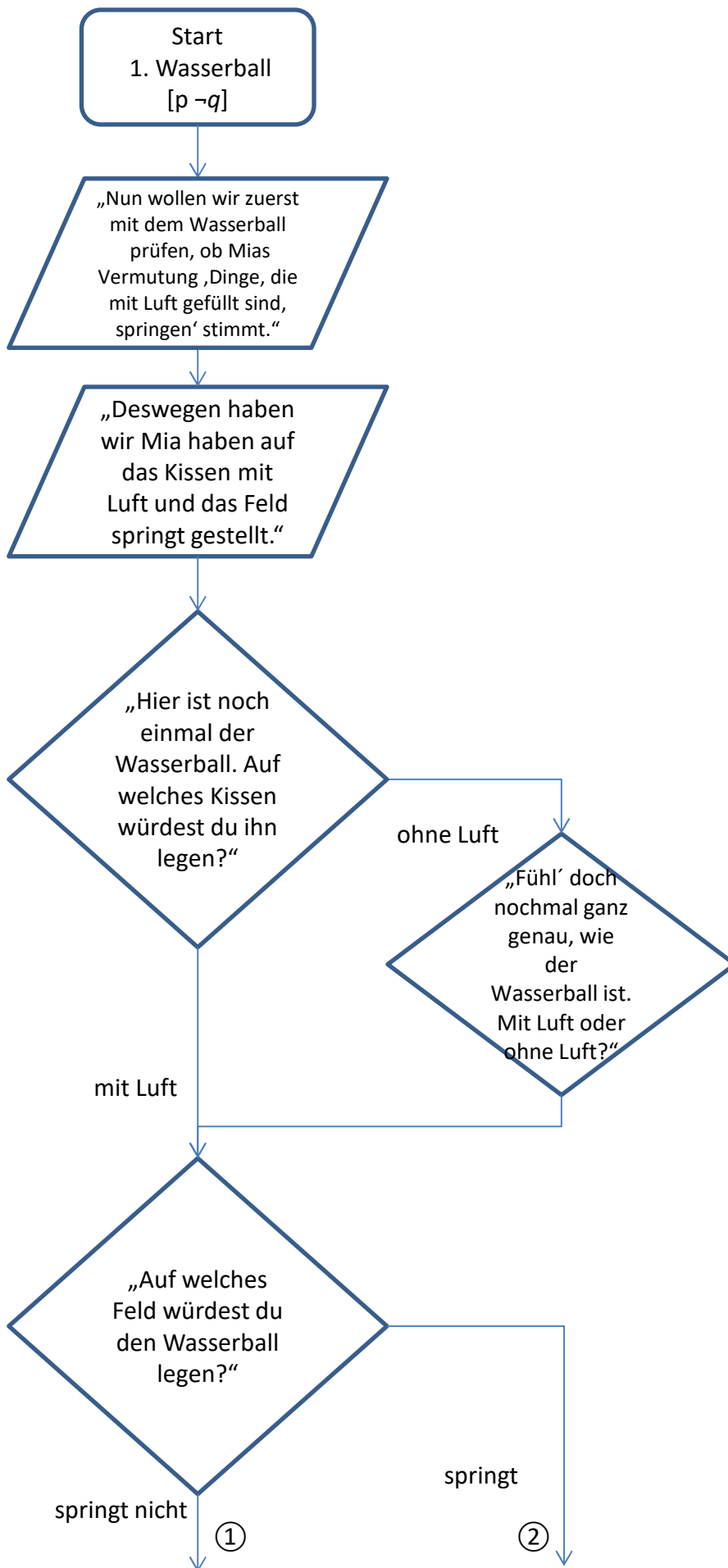
Kind antworten lassen (richtige Antwort: „Kissen mit Luft“, Symbolkarte „springt“)

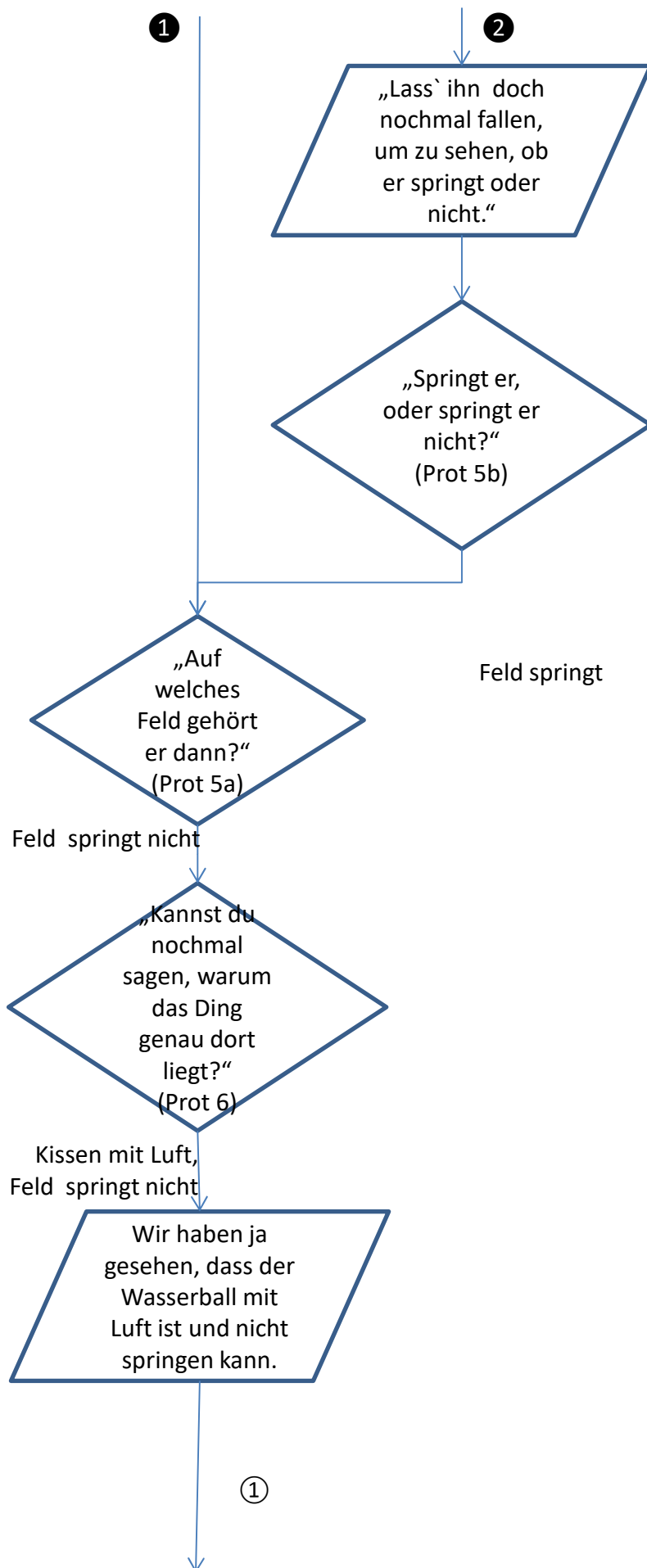
*Solange fragen, bis die Spielfigur auf dem richtigen Feld steht.*

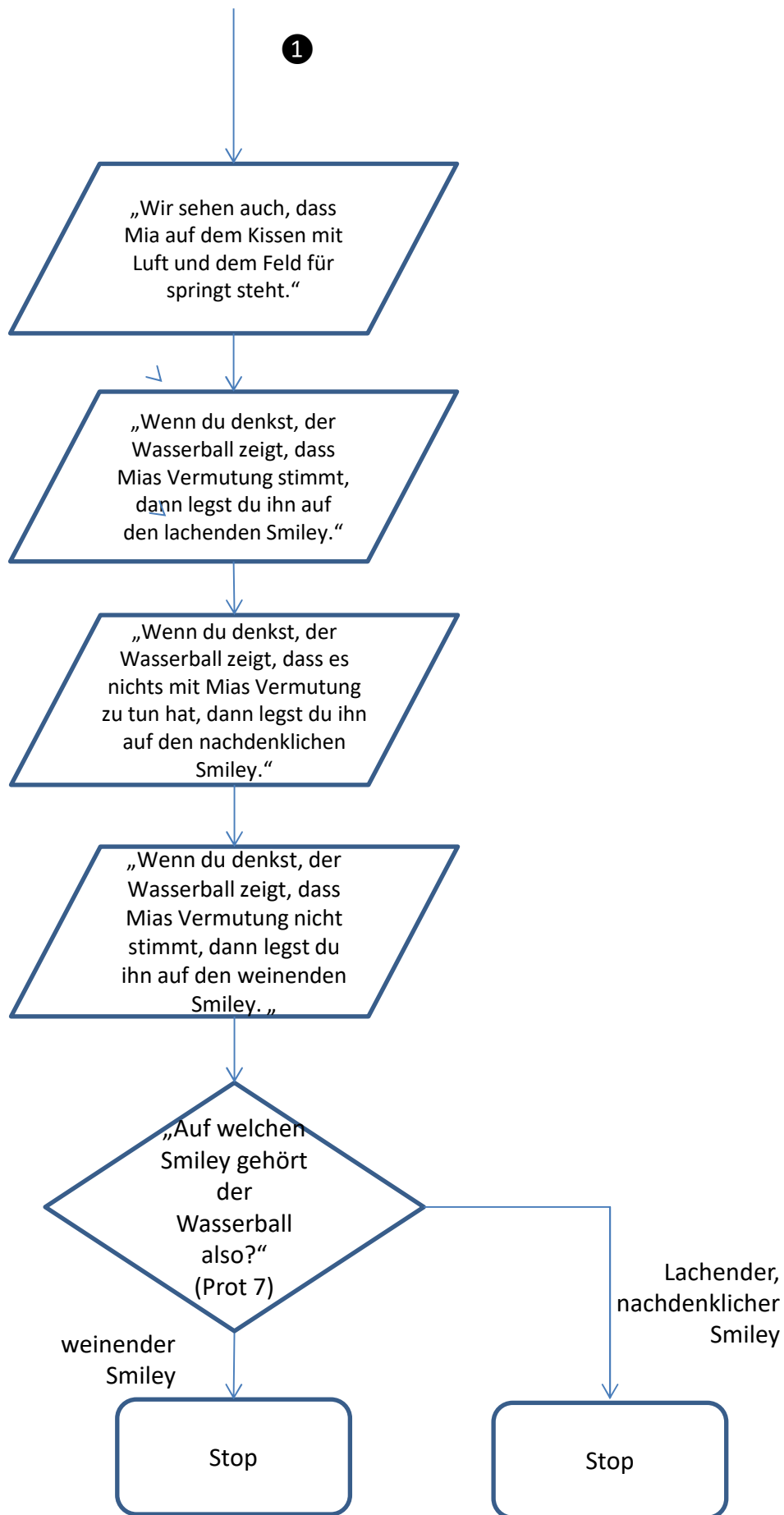
Protokoll 4b)

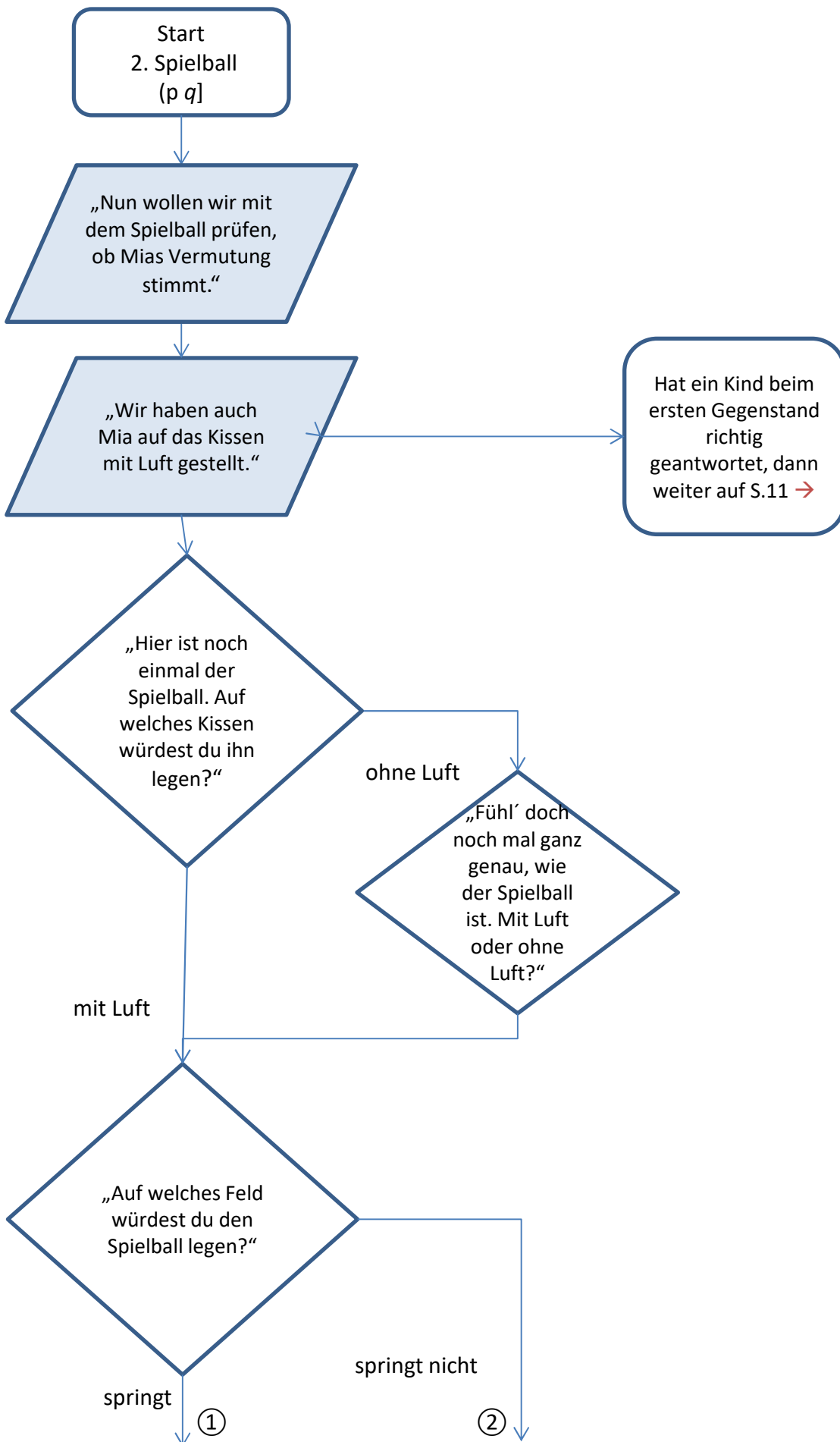
Protokoll 4c)

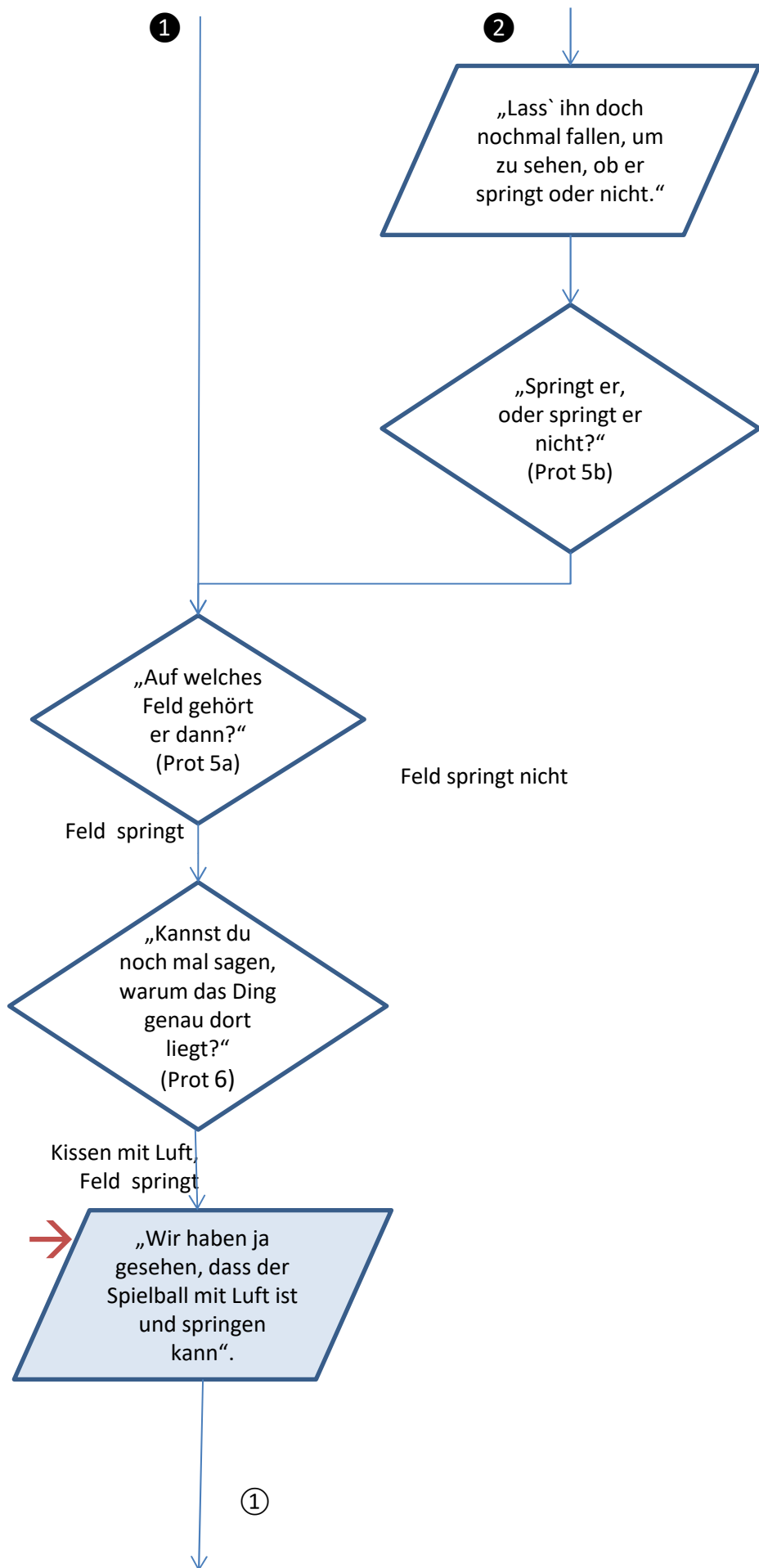


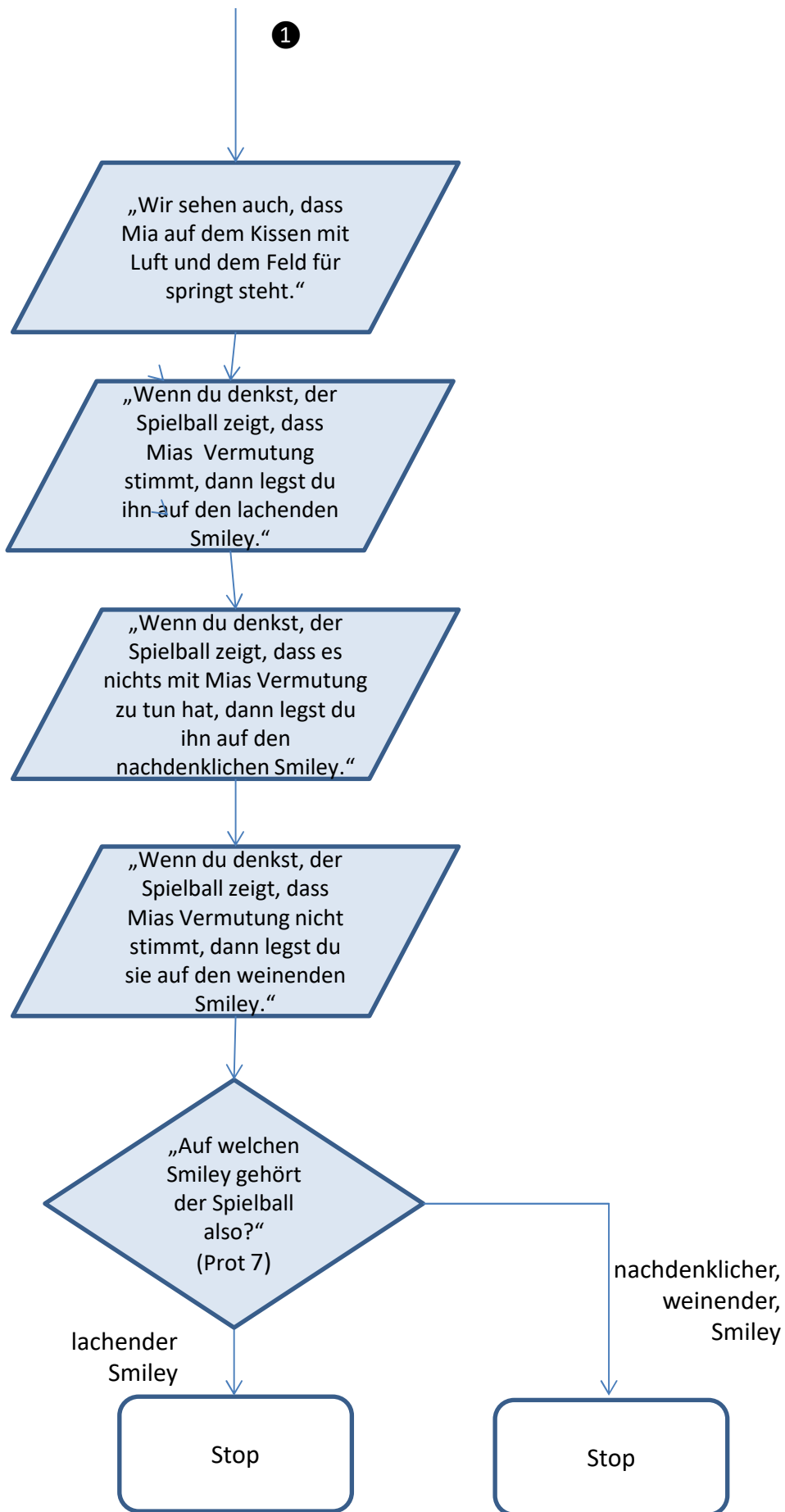


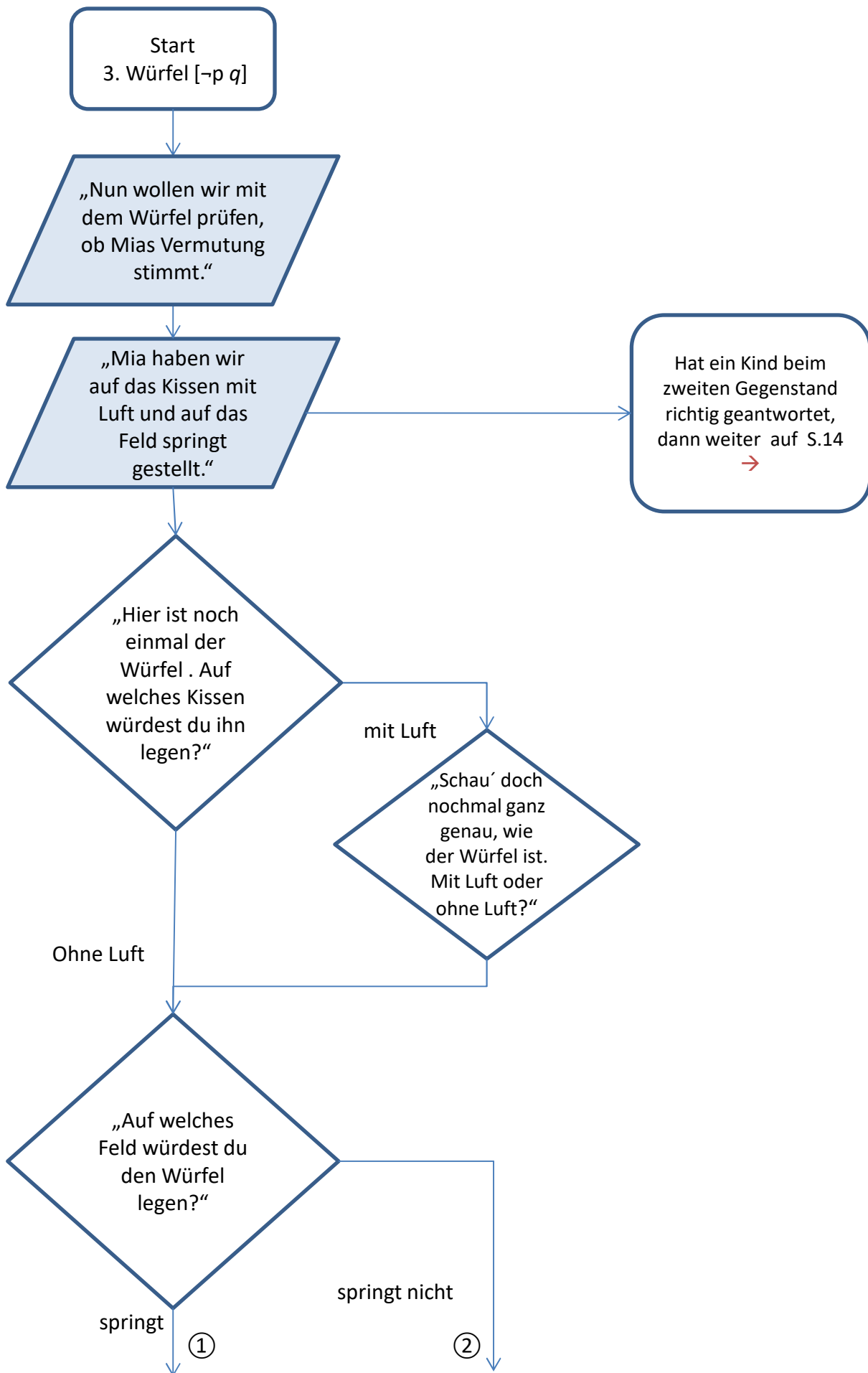


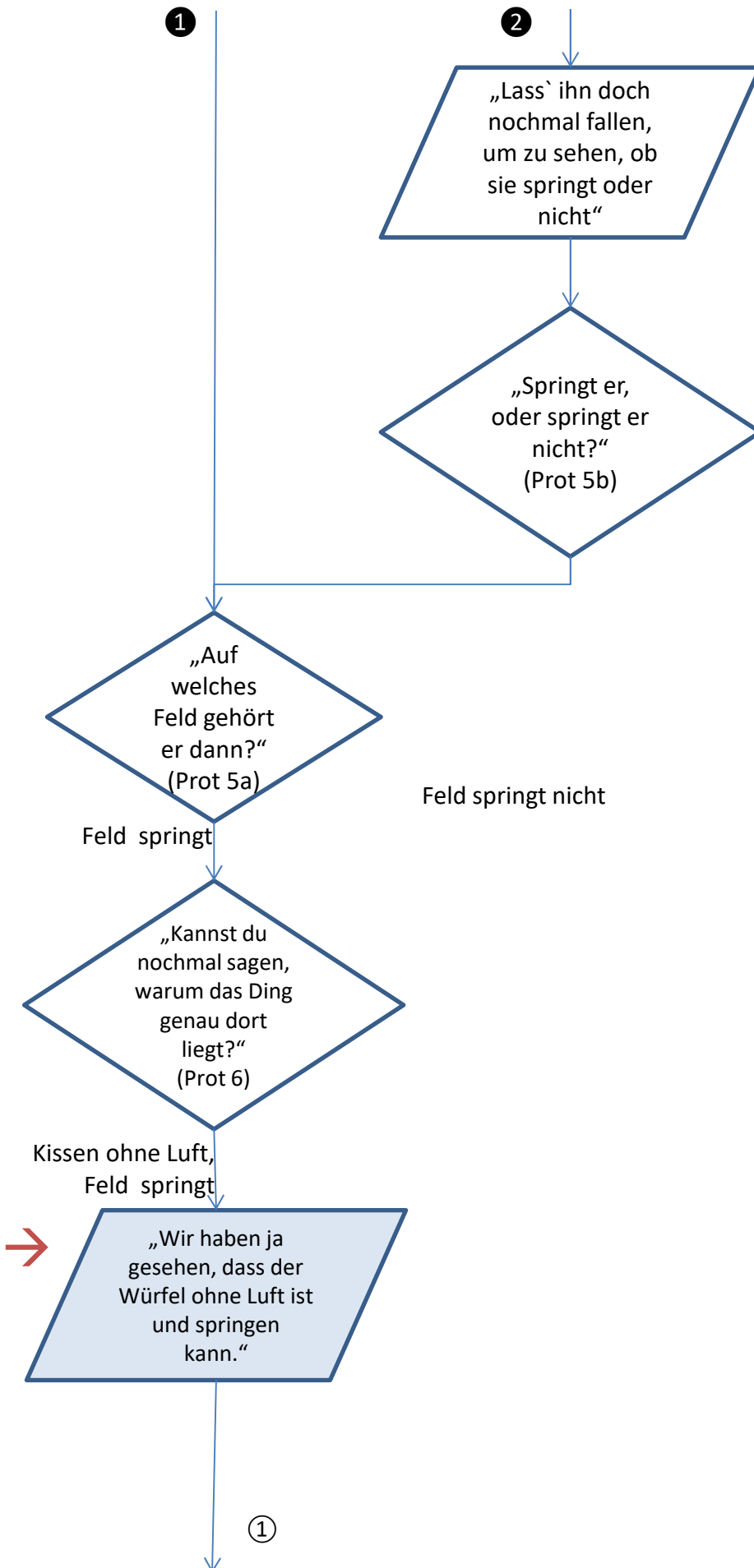




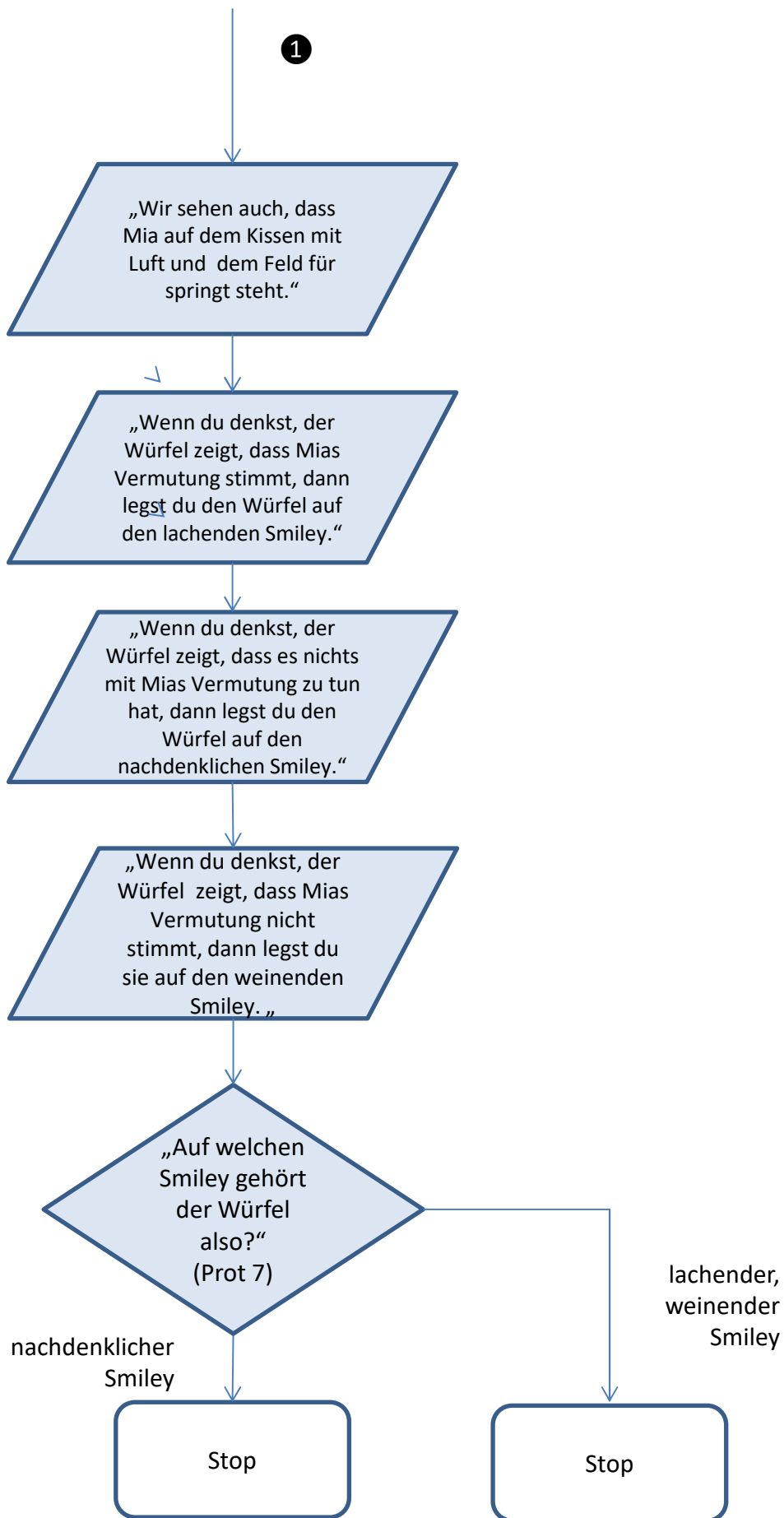


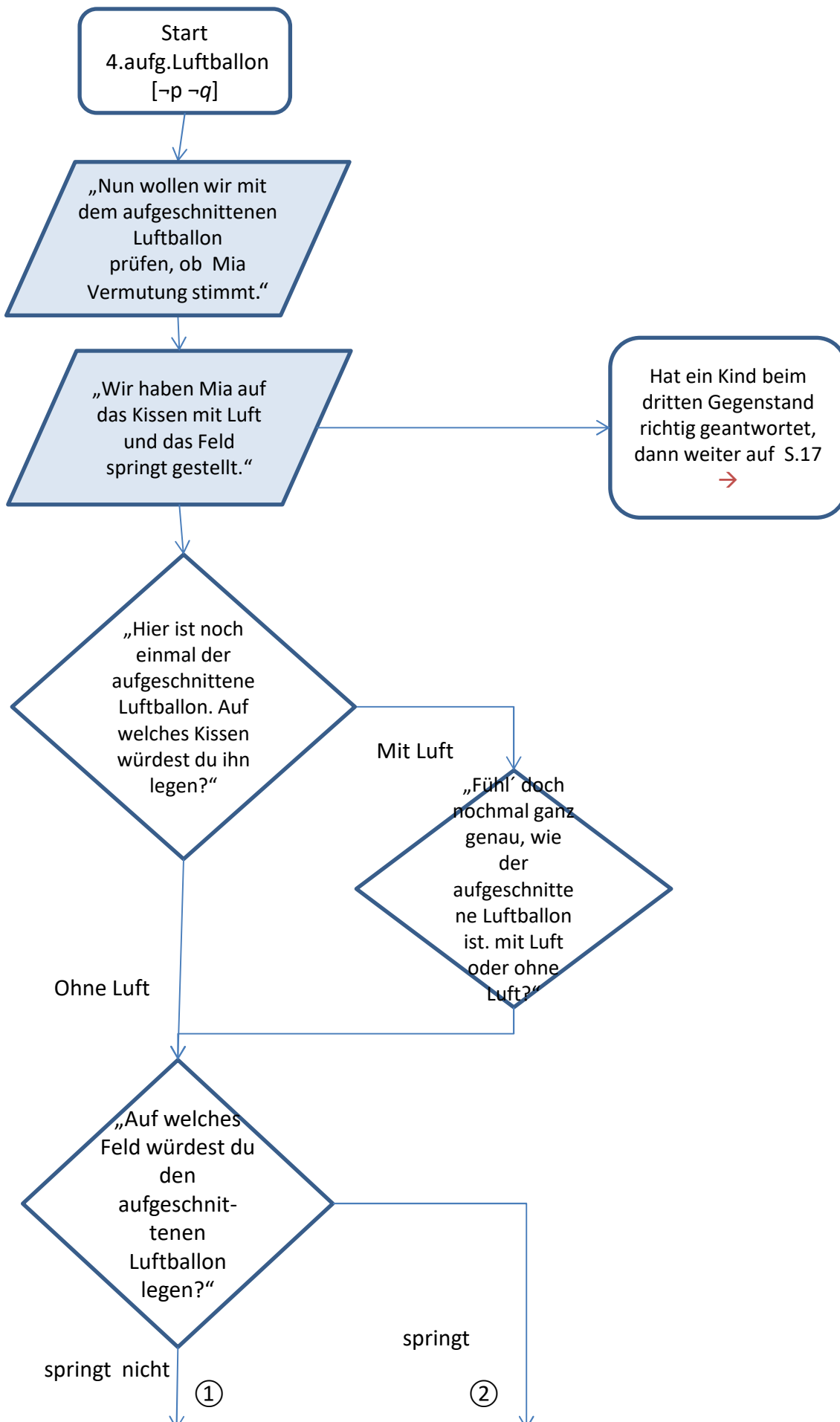


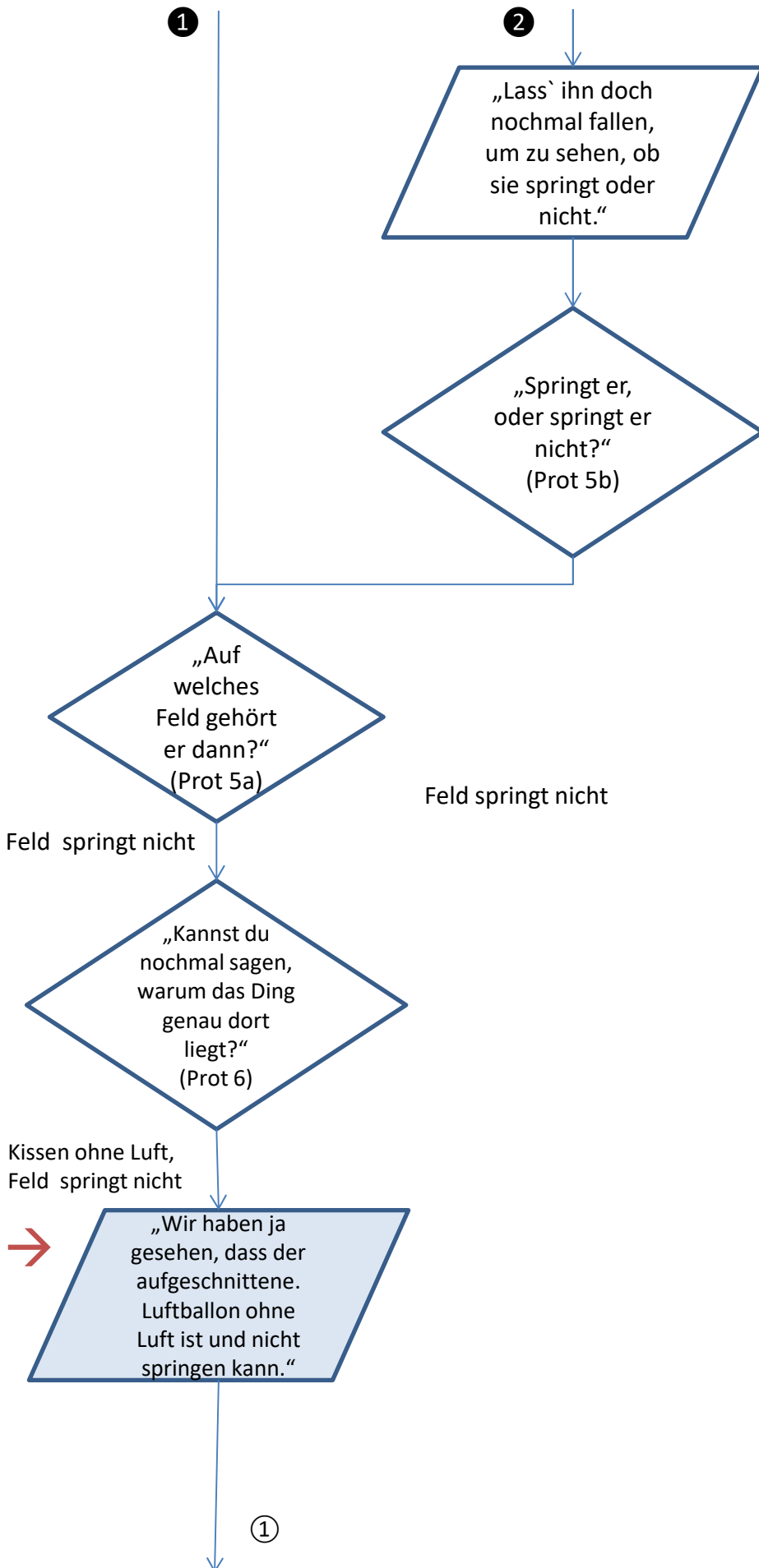


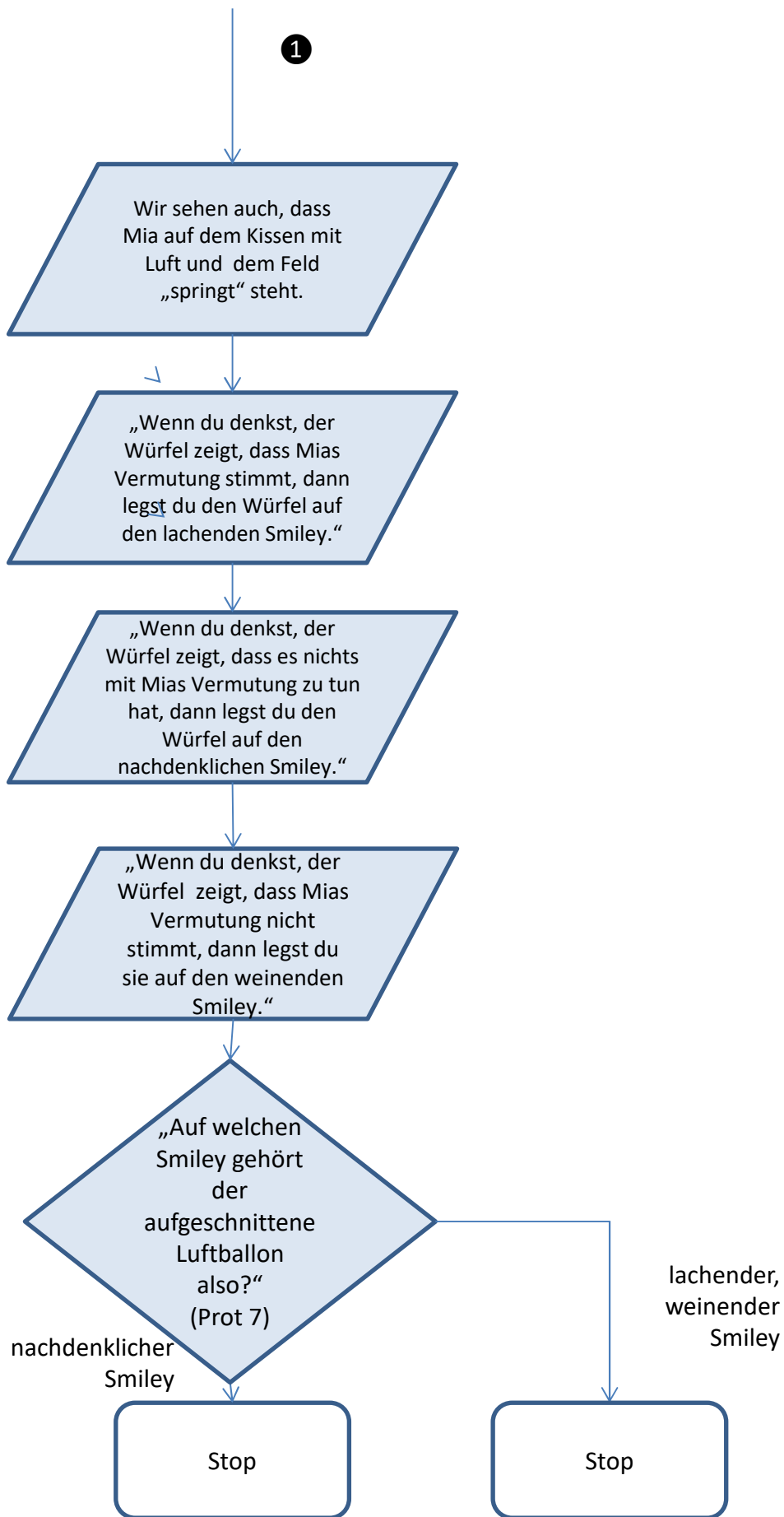




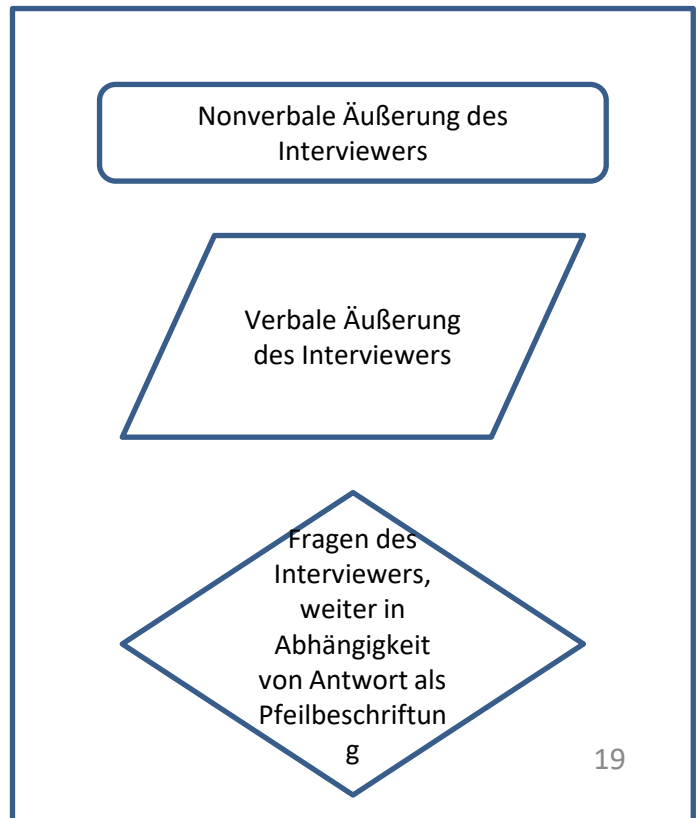








## Legende:



# Interviewleitfaden

## Intervention I mit Modeling

Vermutung „Dinge, die hart sind, springen.“

(lila Knetkugel [ $p \neg q$ ]; Softball [ $\neg p q$ ]; Murmel [ $p q$ ]; Jonglierball)

Explizieren des Gedankenganges komplett

1. Sortieren der Gegenstände

→ nach Eigenschaften

→ nach Sprungverhalten

2. Präsentieren der Vermutung

→ Spielfigur einführen

3. Überprüfen der Vermutung

→ Modeling bis zum Schluss (Explizieren des kompletten Gedankenganges)

# Sortieren der Gegenstände nach Eigenschaften:

**Jens Vermutung: „Dinge, die hart sind, springen.“**

(trockener Knetkugel [p -q]; Softball [-p q]; Murmel [pq]; Jonglierball [-p -q])

**„Schau mal, hier sind vier Dinge: ein Softball, eine Murmel, ein trockener Knetball und ein Jonglierball.“**

(Bälle zeigen)

**„Hier habe ich zwei Tafeln mitgebracht. Eine ist ganz hart (auf Metallplatte zeigen)**

**und die andere ganz weich (auf Platte mit Watte zeigen).**

**„Nun sortier doch mal die vier Dinge nach hart und weich. Die harten Dinge legst du**

**hier hin (auf Metallplatte zeigen) und die weichen Dinge legst du hier hin (auf Platte**

mit Watte zeigen).

Kind sortieren lassen. [richtige Lösung: Murmel: hart; trockener Knetball: hart; Softball: weich; Jonglierball: weich]

Protokoll 1a)

Bei Fehler:

**I: „Fühl doch noch einmal ganz genau. Ist das Ding hart oder weich?“**

Kind anfassen lassen

**I: „Und wo würdest du es dann hinlegen?“**

So lange fragen, bis alle Dinge auf der richtigen Platte liegen

Protokoll 1b)

Protokoll 1c) Bälle wegnehmen

# Sortieren der Gegenstände nach Sprungverhalten

**„Nun wollen wir mal sehen, welche dieser Dinge springen und welche nicht springen können.“**

**„Dafür habe ich hier noch etwas mitgebracht. Diese Karte bedeutet, dass das Ding springen kann. (Karte für springt zeigen) und diese bedeutet, dass das Ding nicht springen kann, also nur fällt und auf liegen bleibt. (Karte für springt nicht zeigen).“**  
(Symbolkarten für springt und springt nicht auf jede Platte legen.)

**„Nun lass die Dinge mal fallen, um zu sehen, ob sie springen können oder nicht.“**  
**Und sortiere die vier Dinge danach, ob sie springen oder nicht.**

**Die Dinge, die hart sind und springen legst du hier hin** (Symbolkarte „springt“ im Feld „hart“), **die Dinge, die hart sind und nicht springen legst du hier hin** (Symbolkarte „springt nicht“ im Feld „hart“),

**die Dinge, die weich sind und springen kommen hier hin** (Symbolkarte „springt“ im Feld „weich“)

**und die Dinge, die weich sind und nicht springen legst du hier hin** (Symbolkarte „springt nicht“ im Feld „weich“).

Kind sortieren lassen. [richtige Lösung: Murmel hart, springt; Knetkugel hart, springt nicht; Softball: weich, springt; Jonglierball: weich, springt nicht]

Protokoll 2a)

Bei Fehler:

**I: „Wir haben ja eben schon gesehen, dass dieses Ding hart / weich ist. Und nun sag mir doch noch mal wo denn die harten / weichen Dinge hinkommen? (falls schon gesagt: Deswegen haben wir es ja auf das harte / weiche Feld gelegt.) und nun probiere doch noch einmal aus, ob das Ding springt oder nicht.“**

Kind ausprobieren lassen (springt)

**Wo würdest du es jetzt hinlegen?“**

So lange fragen, bis alle Dinge auf der richtigen Platte liegen

Bälle liegen lassen



# Zusammenfassung

**„Jetzt wissen wir, welche Dinge hart und welche weich sind.“**

Bälle wegnehmen.

**„Und jetzt sag mir doch nochmal warum du die Bälle auf die jeweiligen Felder gelegt hast. Also warum liegt zum Beispiel die Knetkugel dort?“**

[auf Murmel bzw. Feld zeigen]

Kind antworten lassen

Protokoll 3a)

**„Und warum liegt der Softball dort?“**

Kind antworten lassen

Protokoll 3b)

**„Was ist mit der Murmel?“**

Kind antworten lassen

Protokoll 3c)

**„Und warum liegt der Jonglierball auf diesem Feld?“**

Kind antworten lassen

Protokoll 3d)

**„Jetzt wissen wir, welche harten Dinge springen und welche nicht. Wir wissen welche weichen Dinge springen und welche nicht.“**

Bälle wegnehmen.

# Erklärung: Was ist eine Vermutung?

**Wir wollen heute verschiedene Vermutungen überprüfen.**

**Eine Vermutung ist etwas, das ich zwar glaube, aber noch überprüfen muss.**

(Kästchen hervorholen und zeigen)

**Ich glaube, dass in dem Säckchen Orangen versteckt sind. Was glaubst du?** (Kind antworten lassen)

**Wie können wir herausbekommen, ob deine oder meine Vermutung stimmt?** (Kind antworten lassen)

Stummer Impuls: Säckchen mit Öffnung zum Kind halten (Kind nachschauen lassen).

Es sind Äpfel drin.

Welche Vermutung stimmt, deine oder meine Vermutung? Oder hat keiner von uns beiden Recht? (Kind antworten lassen)

Eine Vermutung kann stimmen, kann aber auch falsch sein. Das müssen wir immer überprüfen.

# Präsentation der Vermutung:

**„Nun haben wir ja schon einige Kinder gefragt, was sie denken, warum ein Ball springt. Jens meint, dass Dinge, die hart sind, springen.“**

**„Das hier ist Jens. (auf männliche Playmobilfigur zeigen)**

**Wo würdest du Jens nun hinstellen? Jens meint ja, dass Dinge die hart sind springen. Die Dinge von denen er spricht müssen also hart sein und springen können. (Richtige Antwort: „Auf die Symbolkarte „springt“ im Feld „hart“.)**

Protokoll 4a)

Bei Fehler:

**„Spricht Jens von harten oder von weichen Dingen? Auf welches Feld würdest du Jens also stellen?“**

Kind antworten lassen (richtige Antwort: hart)

**„Und, meint Jens, dass diese Dinge springen können oder, dass sie nicht springen können?“**

Kind antworten lassen (richtige Antwort: springen)

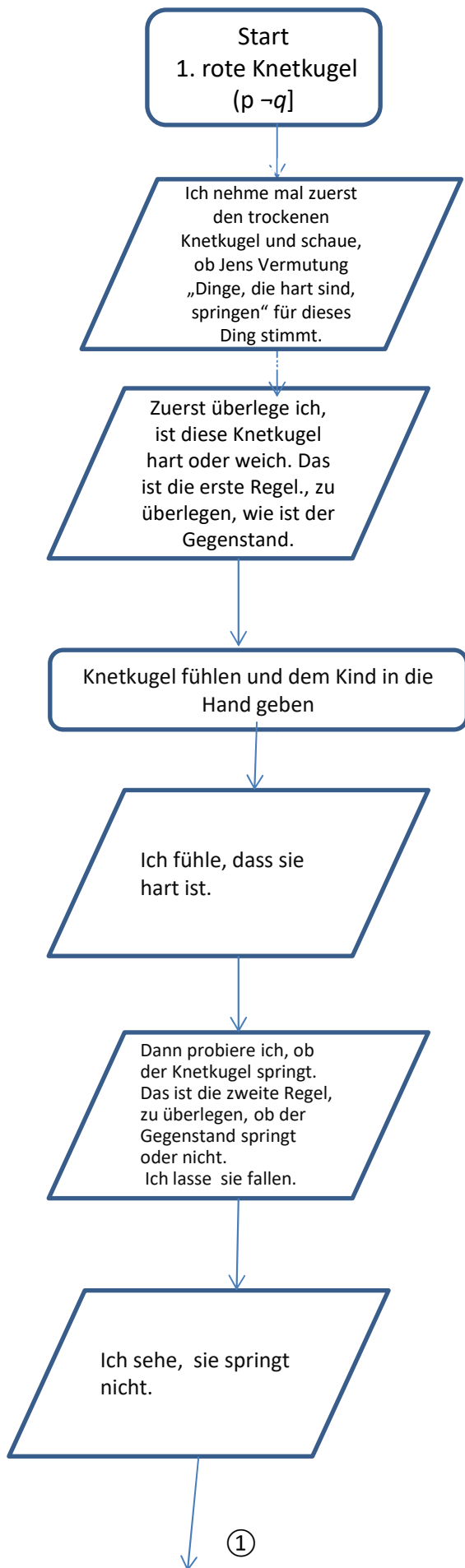
**„Auf welches Feld musst du Jens also stellen?“**

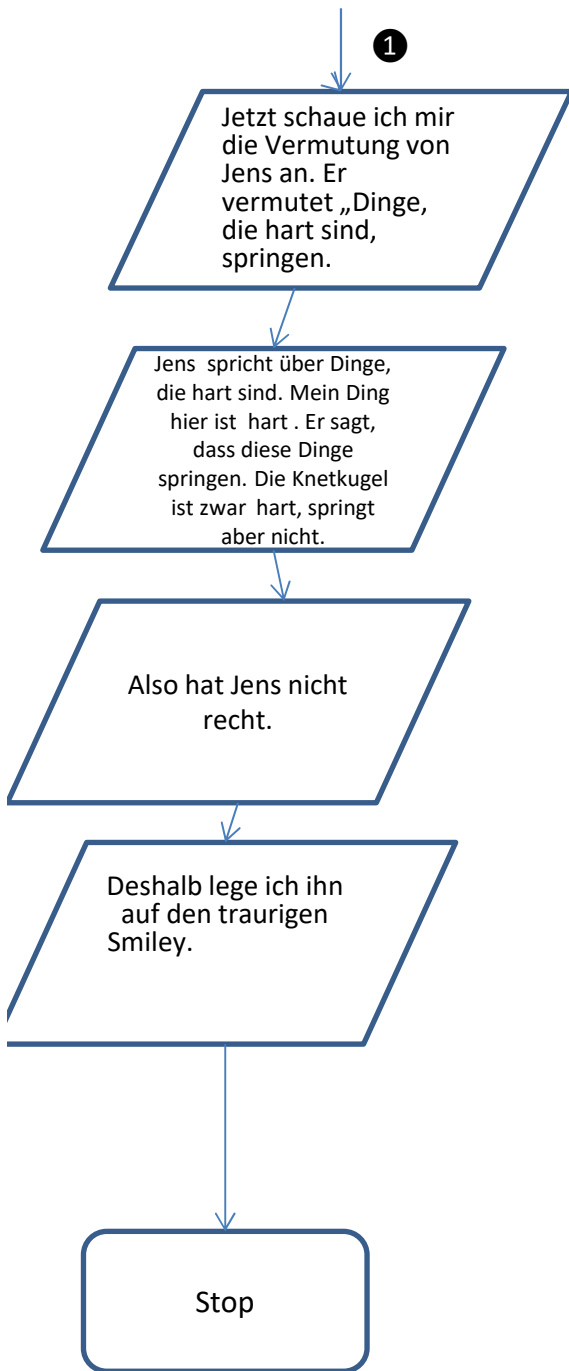
Kind antworten lassen (richtige Antwort: Feld „hart“, Symbolkarte „springt“)

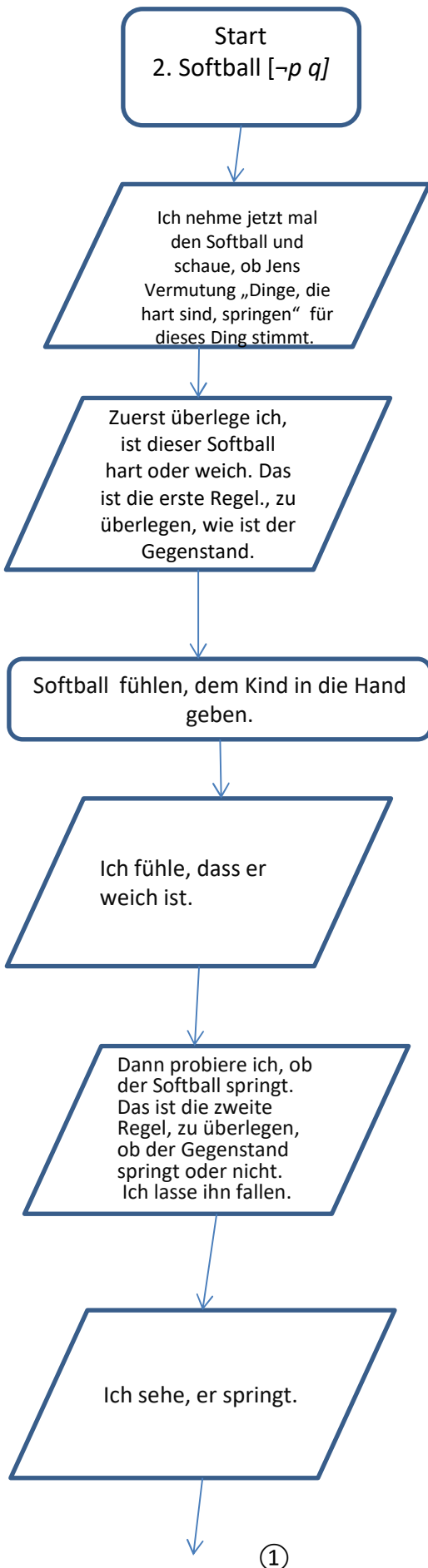
Solange fragen, bis die Spielfigur auf dem richtigen Feld steht.

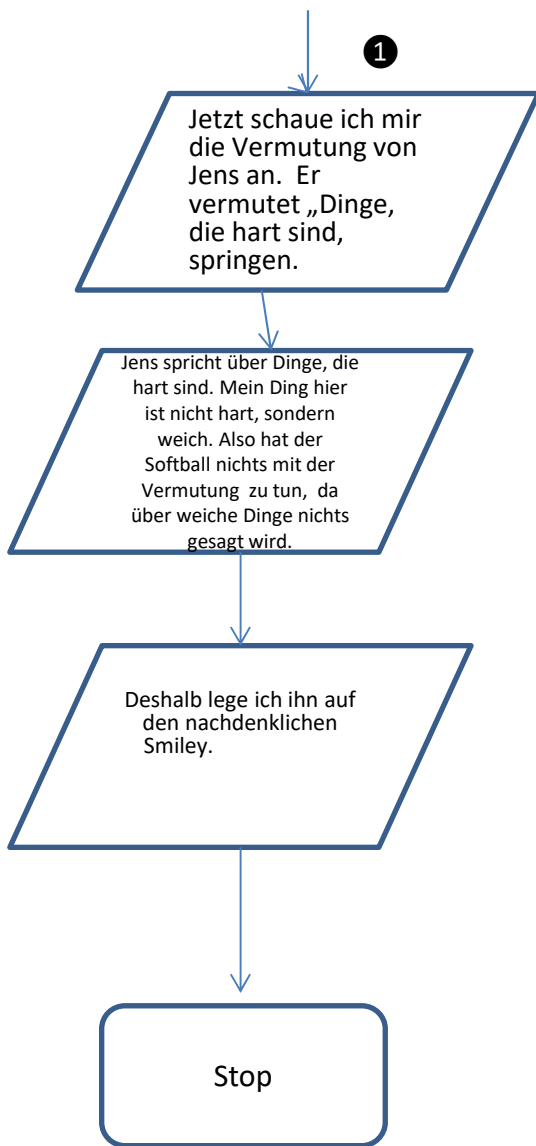
Protokoll 4b)

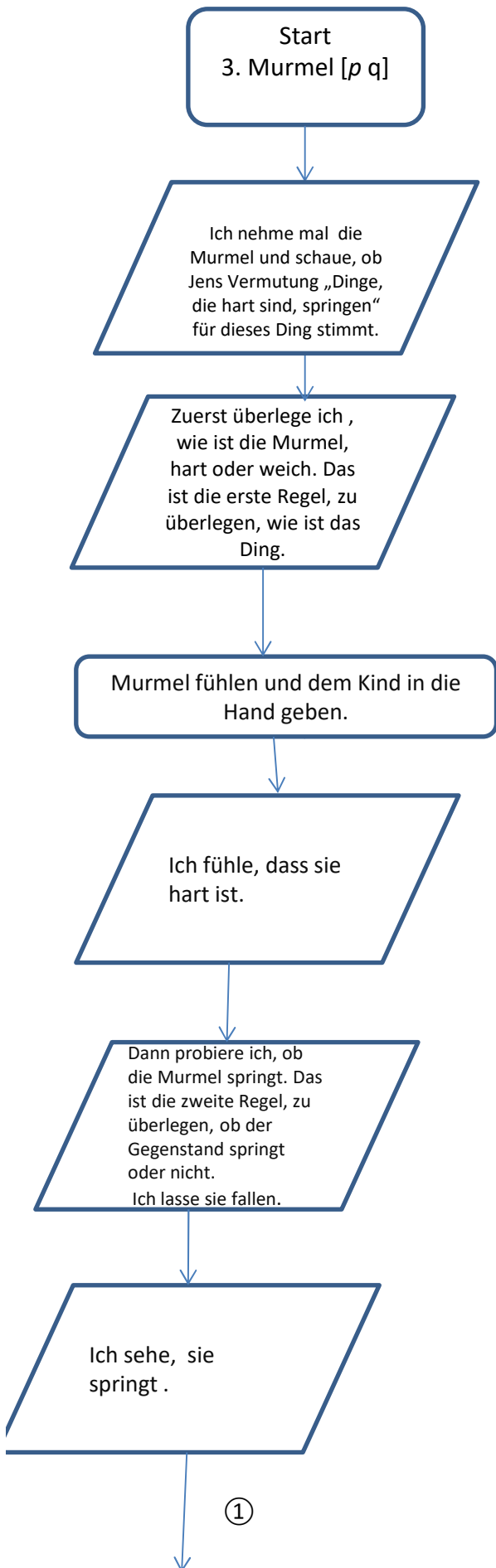
Protokoll 4c)



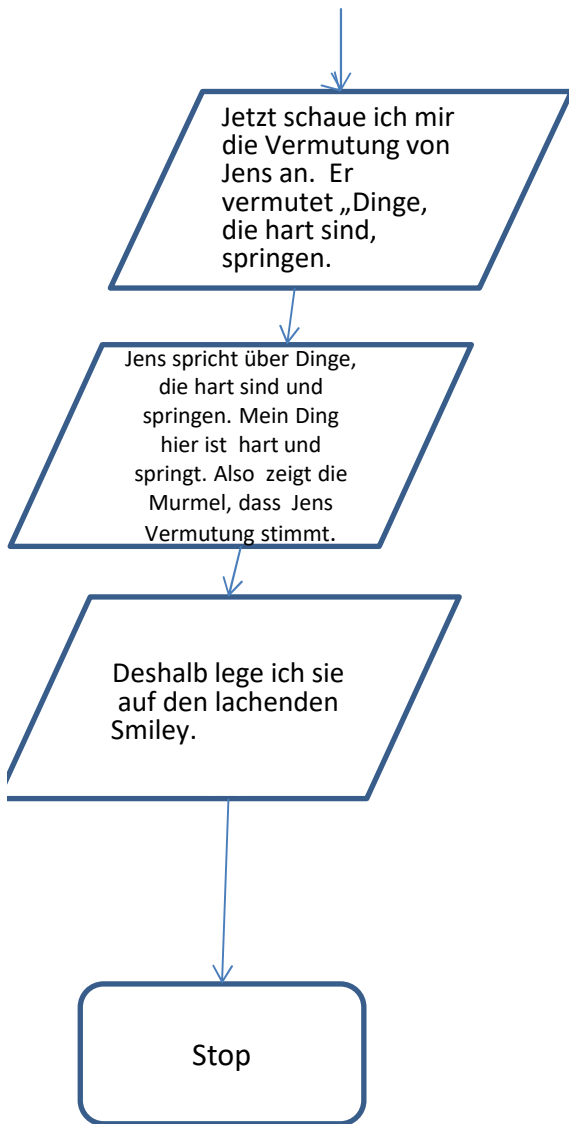


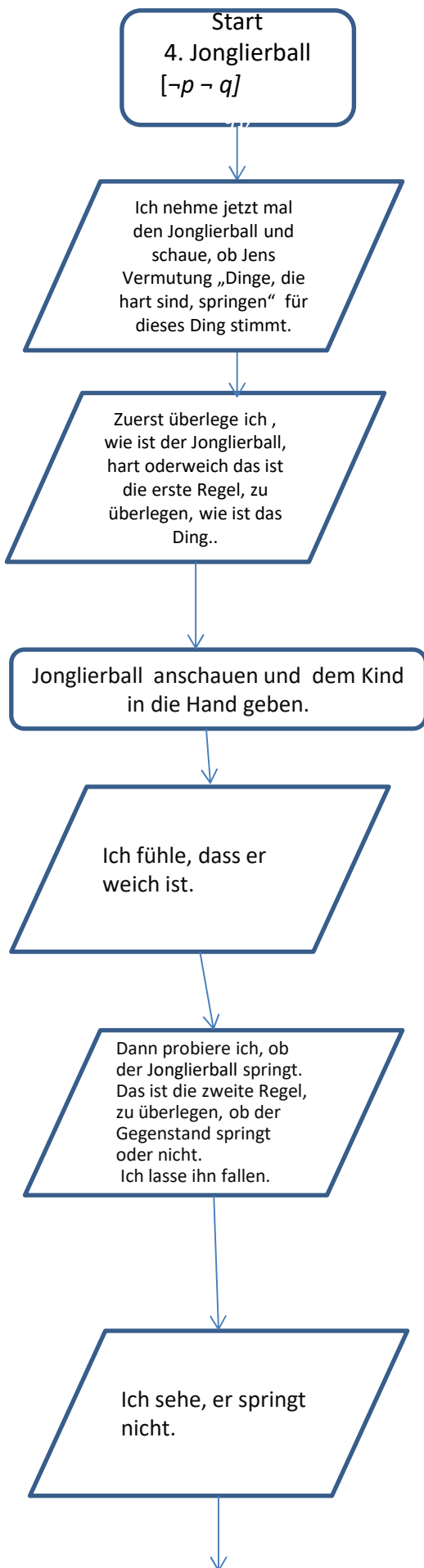


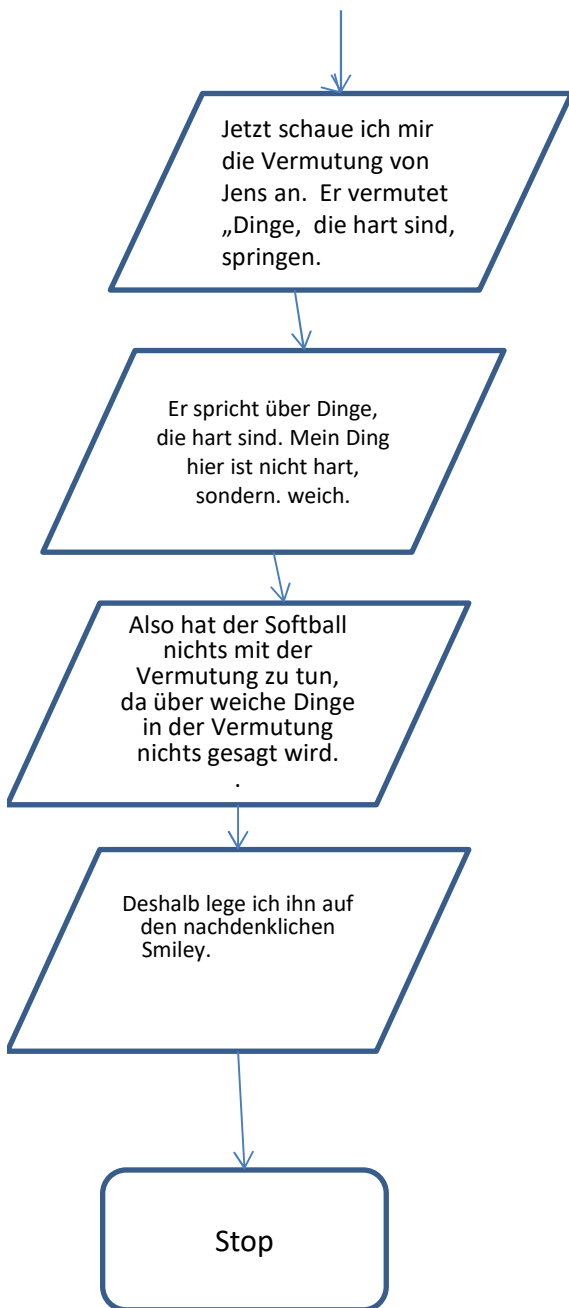




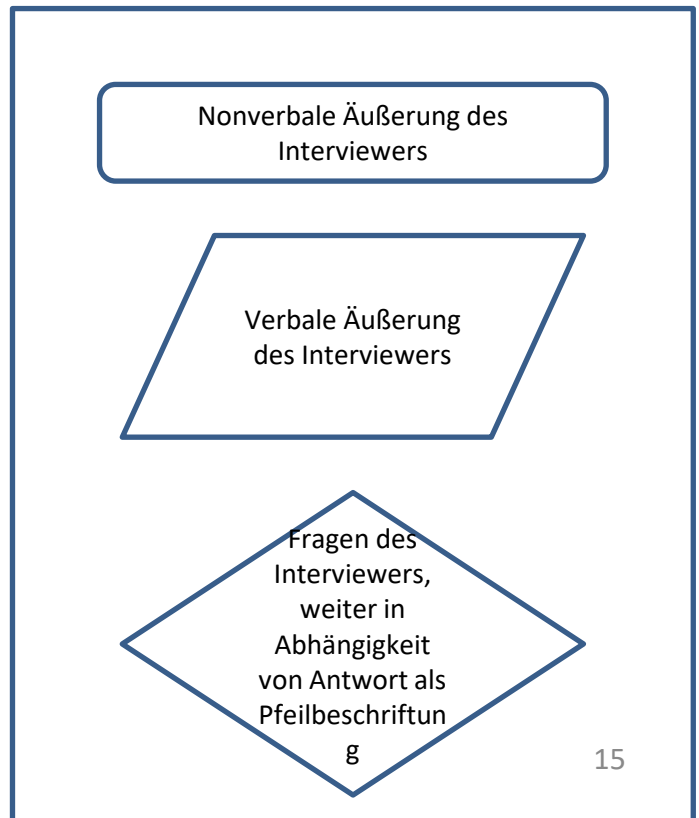








## Legende:



## Protokollbogen Laborstudie 1. Trainings Sitzung (Scaffolding mit Modellierung und adaptivem Nachfragen)

Geschlecht:    m    w    TN:       

In den einzelnen Spalten wird jeweils angekreuzt wie das Kind antwortet: **Richtig (r)**, **falsch (f)**; 3. Und 7. Wird in Stichpunkten protokolliert

Vermutung: „Dinge, die hart sind, springen“		1. Eigen- schaft			2. Sprung- ver- halten			3. Zusammen- fassung			4. Jens			5. Zuord- nung Ball			Kein Protokoll		
		a / b / c			a / b / c			a / b / c			a / b / c			a / b / c			a	b	
a) Knetkugel [p ¬q]  b) Softball [¬p q]  c) Murmel [p q]  d) Jonglierball [¬p ¬q]	Der Murmel ist hart und springt. [p q]																Modellierung	☺ ☹ ☹	☺ ☹ ☹
	Die Knetkugel ist hart und springt nicht. [p ¬q]																Modellierung	☺ ☹ ☹	☺ ☹ ☹
	Der Softball ist weich und springt [¬p q]																Modellierung	☺ ☹ ☹	☺ ☹ ☹
	Der Jonglierball ist weich und springt nicht. [¬p ¬q]																Modellierung	☺ ☹ ☹	☺ ☹ ☹
	Der Jonglierball ist weich und springt nicht. [¬p ¬q]																Modellierung	☺ ☹ ☹	☺ ☹ ☹



# Interviewleitfaden

## 1. Trainingsitzung (Scaffolding mit Modellierung und adaptivem Nachfragen)

Vermutung „Dinge, die hart sind, springen.“

(lila Knetkugel [ $p \neg q$ ]; Softball [ $\neg p q$ ]; Murmel [ $p q$ ]; Jonglierball)

Explizieren des Gedankenganges komplett

1. Sortieren der Gegenstände

→ nach Eigenschaften

→ nach Sprungverhalten

2. Präsentieren der Vermutung

→ Spielfigur einführen

3. Überprüfen der Vermutung

→ Modeling bis zum Schluss (Explizieren des kompletten Gedankenganges)

# Sortieren der Gegenstände nach Eigenschaften:

Jens Vermutung: „Dinge, die hart sind, springen.“

(trockene Knetkugel [p -q]; Softball [-p q]; Murmel [pq]; Jonglierball [-p -q])

„Schau mal, hier sind vier Dinge: ein Softball, eine Murmel, ein trockene Knetkugel und ein Jonglierball.“ (*Bälle zeigen*).

„Hier habe ich zwei Tafeln mitgebracht. Eine ist ganz hart (auf Metallplatte zeigen) und die andere ganz weich (*auf Platte mit Watte zeigen*).

„Nun sortiere doch mal die vier Dinge nach hart und weich. Die harten Dinge legst du hier hin (*auf Metallplatte zeigen*) und die weichen Dinge legst du hier hin (*auf Platte mit Watte zeigen*).

*Kind sortieren lassen.* [Richtige Lösung: Murmel: hart; trockener Knetkugel: hart; Softball: weich; Jonglierball: weich]

## Protokoll 1a)

Bei Fehler:

„Fühl doch noch einmal ganz genau. Ist das Ding hart oder weich?“

(*Kind anfassen lassen*).

„Und wo würdest du es dann hinlegen?“

*So lange fragen, bis alle Dinge auf der richtigen Platte liegen.*

## Protokoll 1b)

Protokoll 1c) *Bälle wegnehmen*



# Sortieren der Gegenstände nach Sprungverhalten

„Nun wollen wir mal sehen, welche dieser Dinge springen und welche nicht springen können.“

„Dafür habe ich hier noch etwas mitgebracht. Diese Karte bedeutet, dass das Ding springen kann. (*Karte für springt zeigen*) und diese bedeutet, dass das Ding nicht springen kann, also nur fällt und auf liegen bleibt. (*Karte für springt nicht zeigen*).“ (*Symbolkarten für springt und springt nicht auf jede Platte legen*).

„Nun lass die Dinge mal fallen, um zu sehen, ob sie springen können oder nicht.“  
Und sortiere die vier Dinge danach, ob sie springen oder nicht.

„Die Dinge, die hart sind und springen legst du hier hin.“ (Symbolkarte „springt“ im Feld „hart“). „Die Dinge, die hart sind und nicht springen legst du hier hin.“ (Symbolkarte „springt nicht“ im Feld „hart“).

die Dinge, die weich sind und springen kommen hier hin (Symbolkarte „springt“ im Feld „weich“).

„Die Dinge, die weich sind und nicht springen legst du hier hin (Symbolkarte „springt nicht“ im Feld „weich“).

(*Kind sortieren lassen*). [Richtige Lösung: Murmel hart, springt; Knetkugel hart, springt nicht; Softball: weich, springt; Jonglierball: weich, springt nicht]

## Protokoll 2a)

Bei Fehler:

„Wir haben ja eben schon gesehen, dass dieses Ding hart / weich ist. Und nun sag mir doch noch mal wo denn die harten / weichen Dinge hinkommen? (falls schon gesagt: Deswegen haben wir es ja auf das harte / weiche Feld gelegt.) und nun probiere doch noch einmal aus, ob das Ding springt oder nicht.“

(*Kind ausprobieren lassen*).

„Wo würdest du es jetzt hinlegen?“

*So lange fragen, bis alle Dinge auf der richtigen Platte liegen.*

(*Bälle liegen lassen*).

# Zusammenfassung

„Jetzt wissen wir, welche Dinge hart und welche weich sind.“  
(Bälle wegnehmen).

„Und jetzt sag mir doch nochmal warum du die Bälle auf die jeweiligen Felder gelegt hast. Also warum liegt zum Beispiel die Knetkugel dort?“

[auf Murmel bzw. Feld zeigen]

(Kind antworten lassen).

Protokoll 3a)

„Und warum liegt der Softball dort?“

(Kind antworten lassen).

Protokoll 3b)

„Was ist mit der Murmel?“

(Kind antworten lassen).

Protokoll 3c)

„Und warum liegt der Jonglierball auf diesem Feld?“

(Kind antworten lassen).

Protokoll 3d)

„Jetzt wissen wir, welche harten Dinge springen und welche nicht. Wir wissen welche weichen Dinge springen und welche nicht.“

(Kind antworten lassen).

# Erklärung: Was ist eine Vermutung?

„Wir wollen heute verschiedene Vermutungen überprüfen.“

„Eine Vermutung ist etwas, das ich zwar glaube, aber noch überprüfen muss.“  
(*Kästchen hervorholen und zeigen*).

„Ich glaube, dass in dem Säckchen Orangen versteckt sind. Was glaubst du?“ (*Kind antworten lassen*).

„Wie können wir herausbekommen, ob deine oder meine Vermutung stimmt?“  
(*Kind antworten lassen*).

Stummer Impuls: Säckchen mit Öffnung zum Kind halten (*Kind nachschauen lassen*).

„Es sind Äpfel drin.“

„Welche Vermutung stimmt, deine oder meine Vermutung? Oder hat keiner von uns beiden Recht?“ (*Kind antworten lassen*).

„Eine Vermutung kann stimmen, kann aber auch falsch sein. Das müssen wir immer überprüfen.“

# Präsentation der Vermutung:

„Nun haben wir ja schon einige Kinder gefragt, was sie denken, warum ein Ball springt. Jens meint, dass Dinge, die hart sind, springen.“

„Das hier ist Jens.“ (*auf männliche Playmobilfigur zeigen*).

„Wo würdest du Jens nun hinstellen? Jens meint ja, dass Dinge die hart sind springen. Die Dinge von denen er spricht, müssen also ‚hart sein‘ und ‚springen‘ können. [Richtige Antwort: „Auf die Symbolkarte „springt“ im Feld „hart“].

Protokoll 4a)

Bei Fehler:

„Spricht Jens von harten oder von weichen Dingen? Auf welches Feld würdest du Jens also stellen?“

*Kind antworten lassen* [Richtige Antwort: „hart“].

„Und, meint Jens, dass diese Dinge springen können oder, dass sie nicht springen können?“

*Kind antworten lassen* [Richtige Antwort : „springt“].

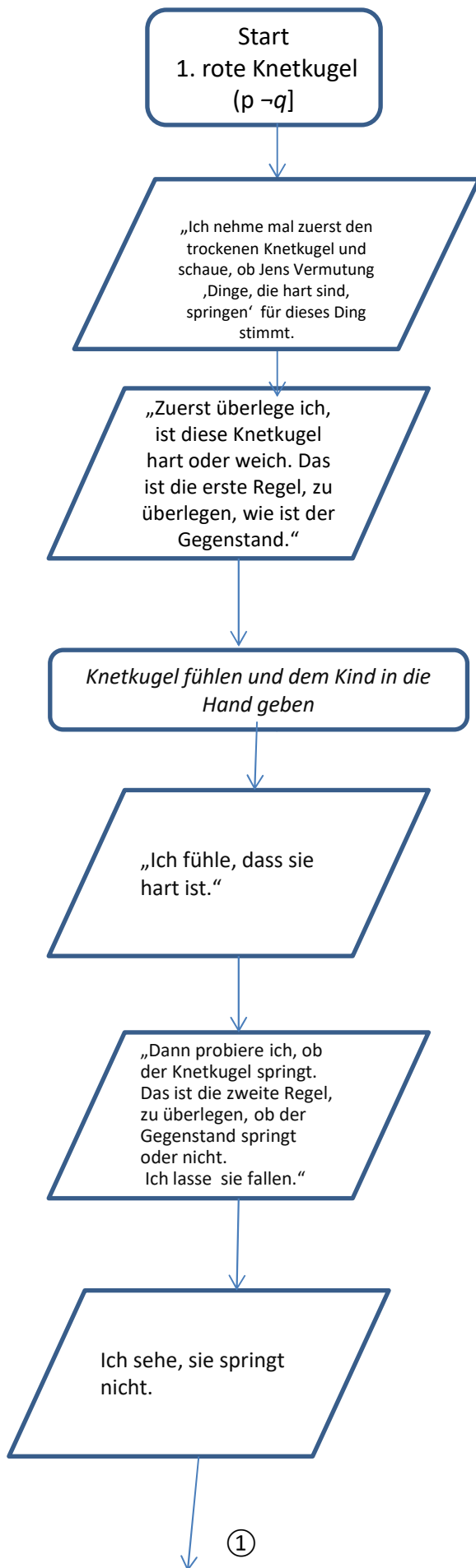
„Auf welches Feld musst du Jens also stellen?“

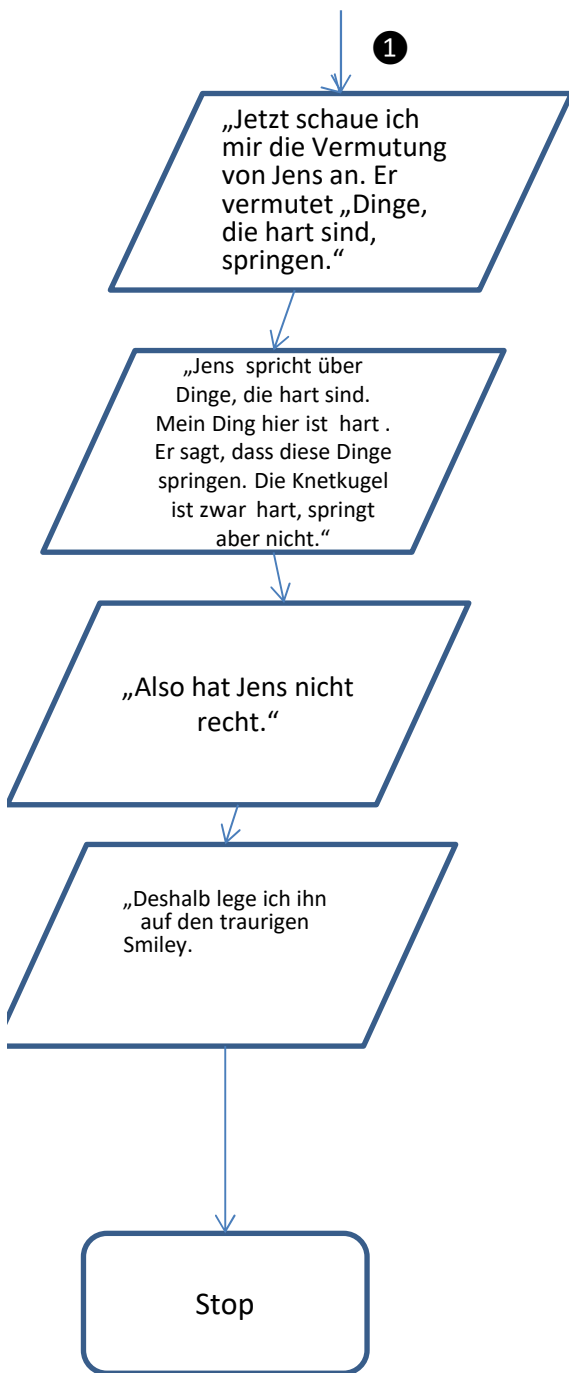
*Kind antworten lassen* [Richtige Antwort: Feld „hart“, Symbolkarte „springt“]

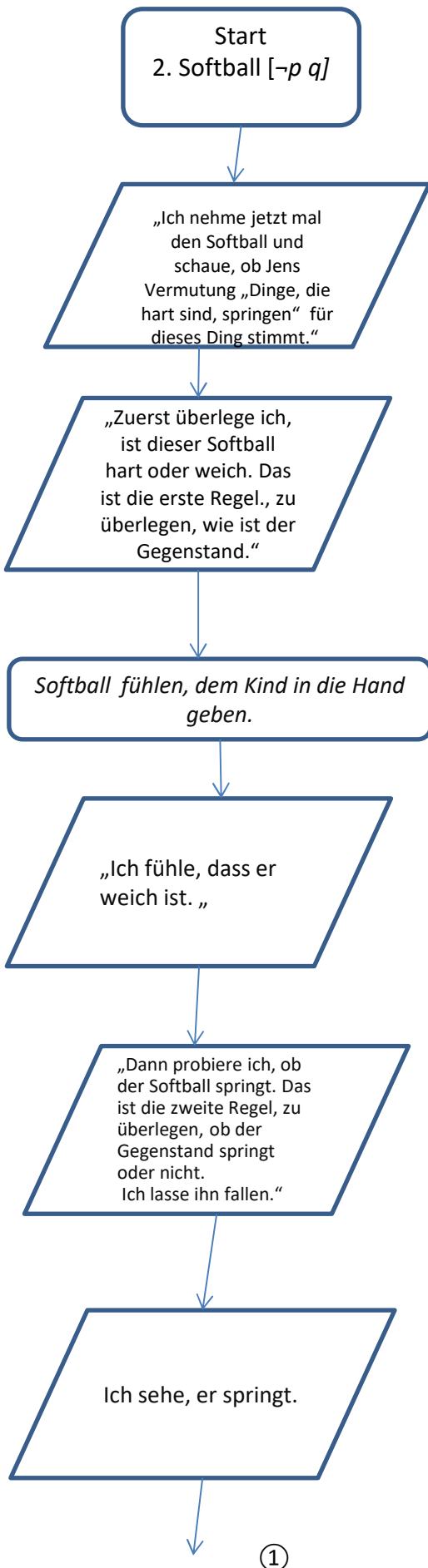
*Solange fragen, bis die Spielfigur auf dem richtigen Feld steht.*

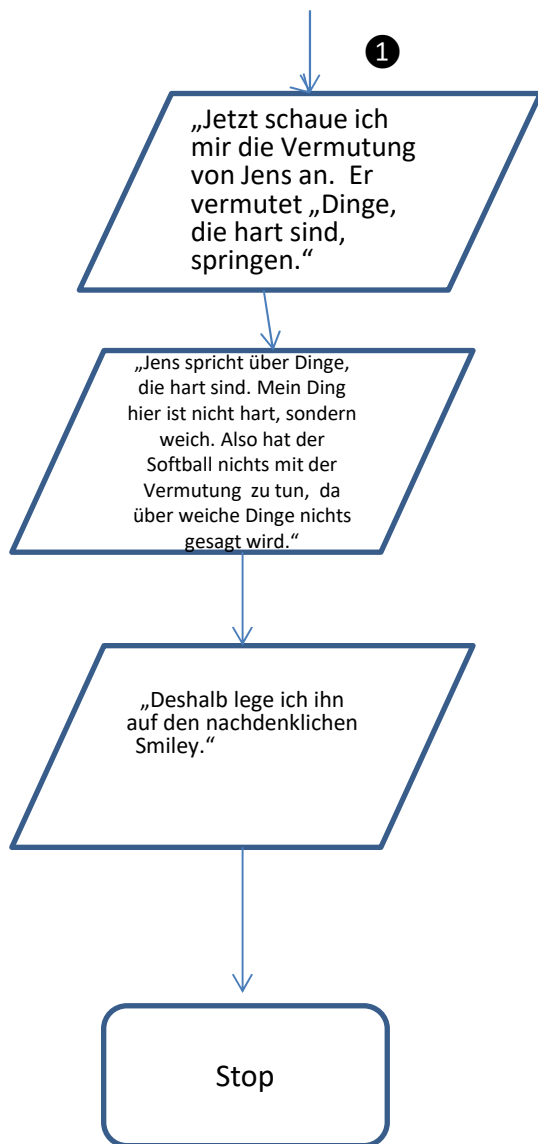
Protokoll 4b)

Protokoll 4c)

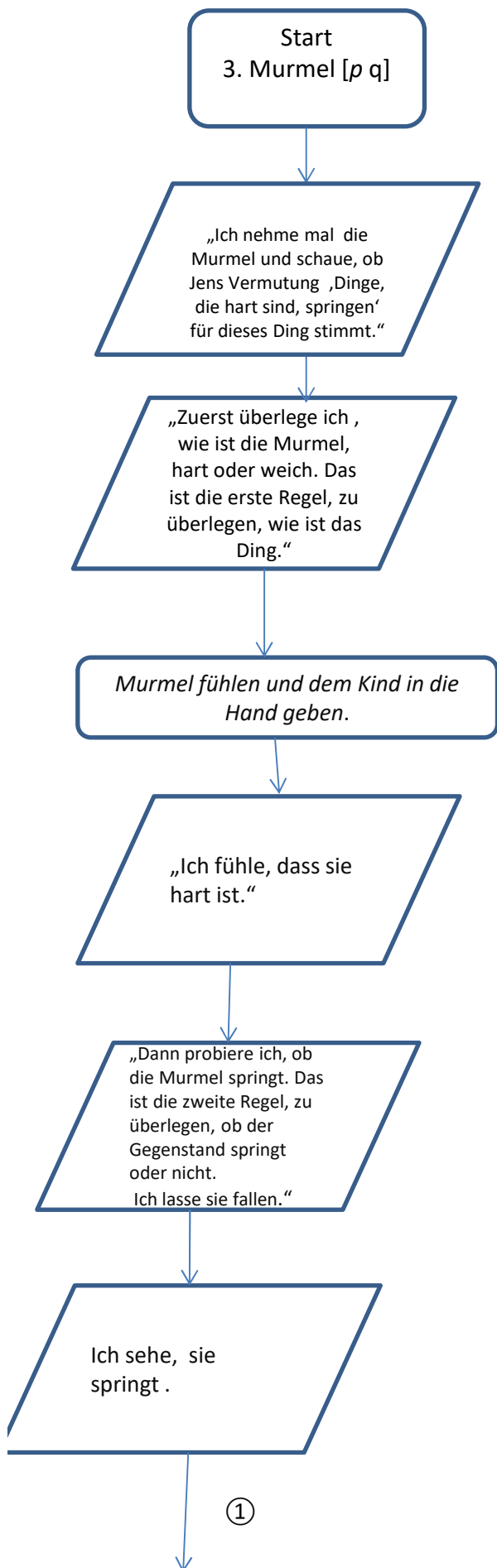


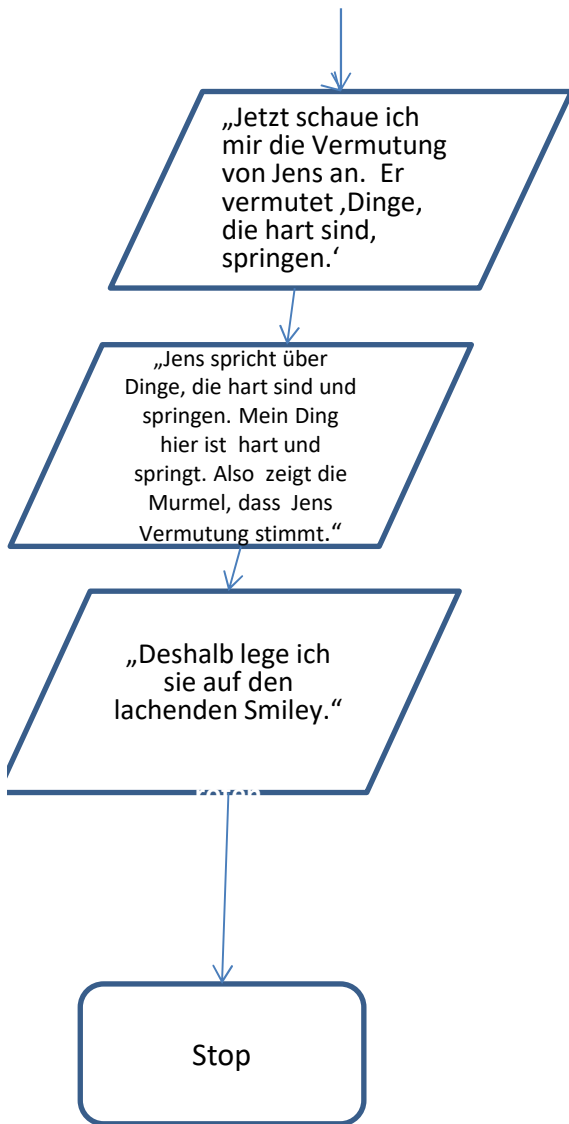


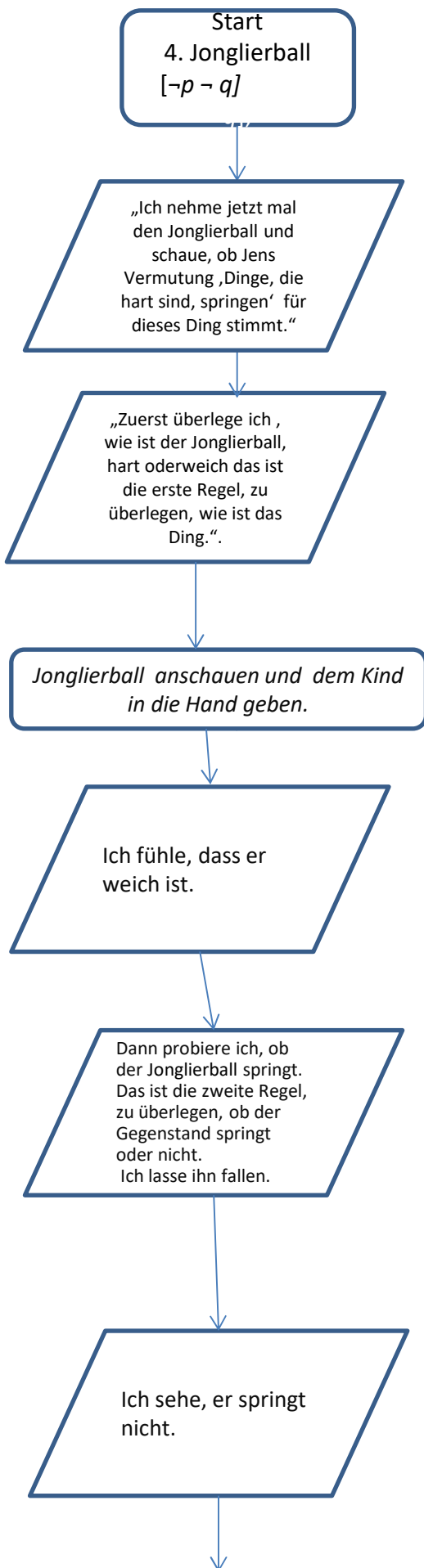


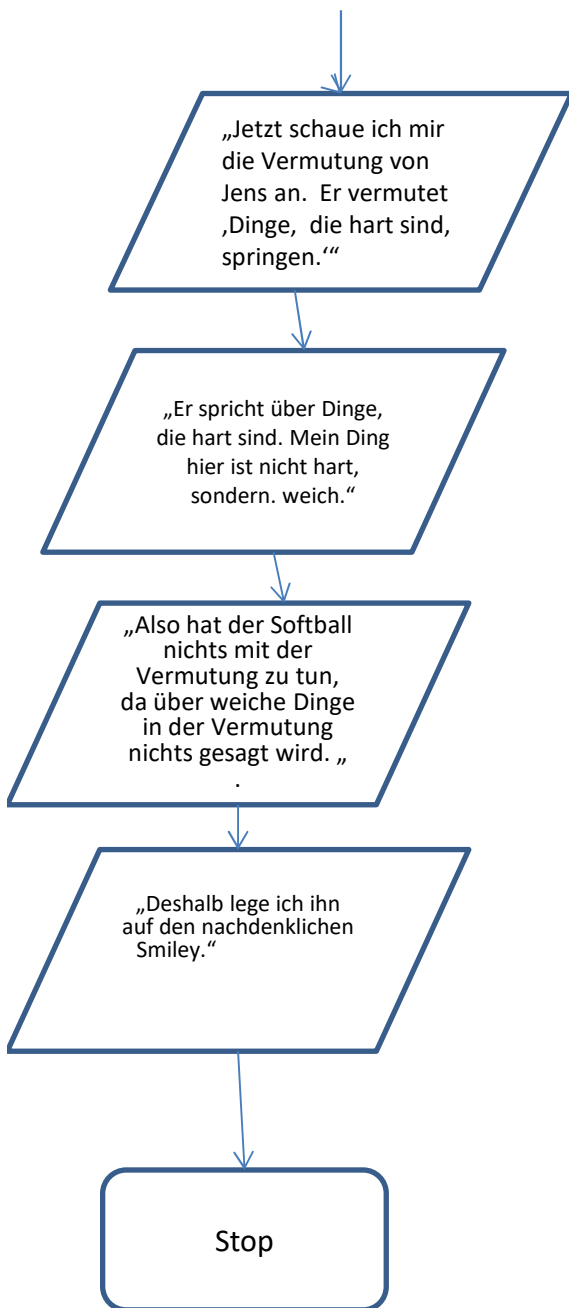




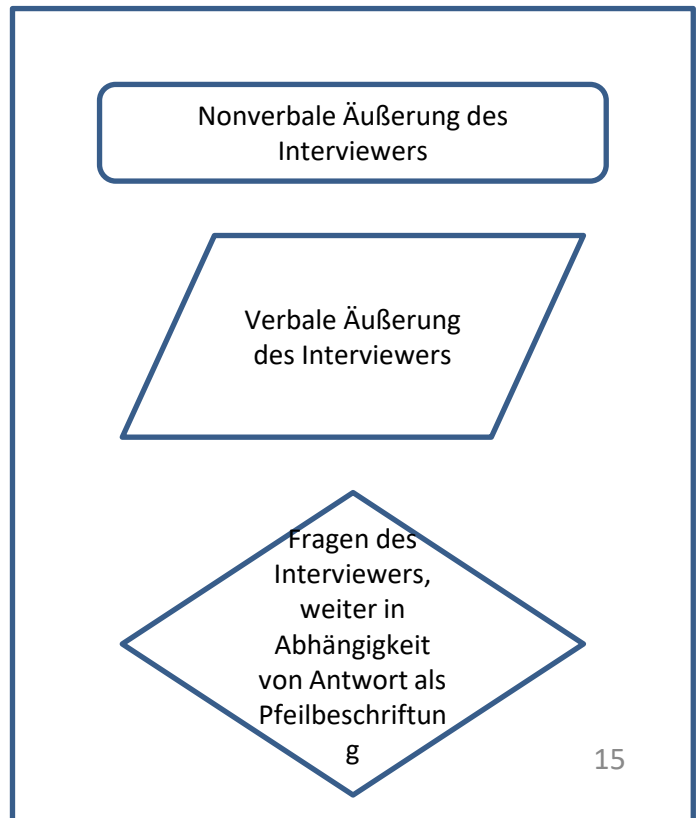








## Legende:



# Interviewleitfaden

## 1. Trainingsitzung (Scaffolding mit Modellierung und adaptivem Nachfragen)

Vermutung „Dinge, die hart sind, springen.“

(lila Knetkugel [ $p \neg q$ ]; Softball [ $\neg p q$ ]; Murmel [ $p q$ ]; Jonglierball)

Explizieren des Gedankenganges komplett

1. Sortieren der Gegenstände

→ nach Eigenschaften

→ nach Sprungverhalten

2. Präsentieren der Vermutung

→ Spielfigur einführen

3. Überprüfen der Vermutung

→ Modeling bis zum Schluss (Explizieren des kompletten Gedankenganges)

# Sortieren der Gegenstände nach Eigenschaften:

**Jens Vermutung: „Dinge, die hart sind, springen.“**

(trockener Knetkugel [p -q]; Softball [-p q]; Murmel [pq]; Jonglierball [-p -q])

**„Schau mal, hier sind vier Dinge: ein Softball, eine Murmel, ein trockener Knetball und ein Jonglierball.“**

(Bälle zeigen)

**„Hier habe ich zwei Tafeln mitgebracht. Eine ist ganz hart (auf Metallplatte zeigen)**

**und die andere ganz weich (auf Platte mit Watte zeigen).**

**„Nun sortier doch mal die vier Dinge nach hart und weich. Die harten Dinge legst du**

**hier hin (auf Metallplatte zeigen) und die weichen Dinge legst du hier hin (auf Platte**

mit Watte zeigen).

Kind sortieren lassen. [richtige Lösung: Murmel: hart; trockener Knetball: hart; Softball: weich; Jonglierball: weich]

Protokoll 1a)

Bei Fehler:

**I: „Fühl doch noch einmal ganz genau. Ist das Ding hart oder weich?“**

Kind anfassen lassen

**I: „Und wo würdest du es dann hinlegen?“**

So lange fragen, bis alle Dinge auf der richtigen Platte liegen

Protokoll 1b)

Protokoll 1c) Bälle wegnehmen

# Sortieren der Gegenstände nach Sprungverhalten

**„Nun wollen wir mal sehen, welche dieser Dinge springen und welche nicht springen können.“**

**„Dafür habe ich hier noch etwas mitgebracht. Diese Karte bedeutet, dass das Ding springen kann. (Karte für springt zeigen) und diese bedeutet, dass das Ding nicht springen kann, also nur fällt und auf liegen bleibt. (Karte für springt nicht zeigen).“**  
(Symbolkarten für springt und springt nicht auf jede Platte legen.)

**„Nun lass die Dinge mal fallen, um zu sehen, ob sie springen können oder nicht.“**  
**Und sortiere die vier Dinge danach, ob sie springen oder nicht.**

**Die Dinge, die hart sind und springen legst du hier hin** (Symbolkarte „springt“ im Feld „hart“), **die Dinge, die hart sind und nicht springen legst du hier hin** (Symbolkarte „springt nicht“ im Feld „hart“),

**die Dinge, die weich sind und springen kommen hier hin** (Symbolkarte „springt“ im Feld „weich“)

**und die Dinge, die weich sind und nicht springen legst du hier hin** (Symbolkarte „springt nicht“ im Feld „weich“).

Kind sortieren lassen. [richtige Lösung: Murmel hart, springt; Knetkugel hart, springt nicht; Softball: weich, springt; Jonglierball: weich, springt nicht]

Protokoll 2a)

Bei Fehler:

**I: „Wir haben ja eben schon gesehen, dass dieses Ding hart / weich ist. Und nun sag mir doch noch mal wo denn die harten / weichen Dinge hinkommen? (falls schon gesagt: Deswegen haben wir es ja auf das harte / weiche Feld gelegt.) und nun probiere doch noch einmal aus, ob das Ding springt oder nicht.“**

Kind ausprobieren lassen (springt)

**Wo würdest du es jetzt hinlegen?“**

So lange fragen, bis alle Dinge auf der richtigen Platte liegen

Bälle liegen lassen



# Zusammenfassung

**„Jetzt wissen wir, welche Dinge hart und welche weich sind.“**

Bälle wegnehmen.

**„Und jetzt sag mir doch nochmal warum du die Bälle auf die jeweiligen Felder gelegt hast. Also warum liegt zum Beispiel die Knetkugel dort?“**

[auf Murmel bzw. Feld zeigen]

Kind antworten lassen

Protokoll 3a)

**„Und warum liegt der Softball dort?“**

Kind antworten lassen

Protokoll 3b)

**„Was ist mit der Murmel?“**

Kind antworten lassen

Protokoll 3c)

**„Und warum liegt der Jonglierball auf diesem Feld?“**

Kind antworten lassen

Protokoll 3d)

**„Jetzt wissen wir, welche harten Dinge springen und welche nicht. Wir wissen welche weichen Dinge springen und welche nicht.“**

Bälle wegnehmen.

# Erklärung: Was ist eine Vermutung?

**Wir wollen heute verschiedene Vermutungen überprüfen.**

**Eine Vermutung ist etwas, das ich zwar glaube, aber noch überprüfen muss.**

(Kästchen hervorholen und zeigen)

**Ich glaube, dass in dem Säckchen Orangen versteckt sind. Was glaubst du?** (Kind antworten lassen)

**Wie können wir herausbekommen, ob deine oder meine Vermutung stimmt?** (Kind antworten lassen)

Stummer Impuls: Säckchen mit Öffnung zum Kind halten (Kind nachschauen lassen).

Es sind Äpfel drin.

Welche Vermutung stimmt, deine oder meine Vermutung? Oder hat keiner von uns beiden Recht? (Kind antworten lassen)

Eine Vermutung kann stimmen, kann aber auch falsch sein. Das müssen wir immer überprüfen.

# Präsentation der Vermutung:

**„Nun haben wir ja schon einige Kinder gefragt, was sie denken, warum ein Ball springt. Jens meint, dass Dinge, die hart sind, springen.“**

**„Das hier ist Jens. (auf männliche Playmobilfigur zeigen)**

**Wo würdest du Jens nun hinstellen? Jens meint ja, dass Dinge die hart sind springen. Die Dinge von denen er spricht müssen also hart sein und springen können. (Richtige Antwort: „Auf die Symbolkarte „springt“ im Feld „hart“.)**

Protokoll 4a)

Bei Fehler:

**„Spricht Jens von harten oder von weichen Dingen? Auf welches Feld würdest du Jens also stellen?“**

Kind antworten lassen (richtige Antwort: hart)

**„Und, meint Jens, dass diese Dinge springen können oder, dass sie nicht springen können?“**

Kind antworten lassen (richtige Antwort: springen)

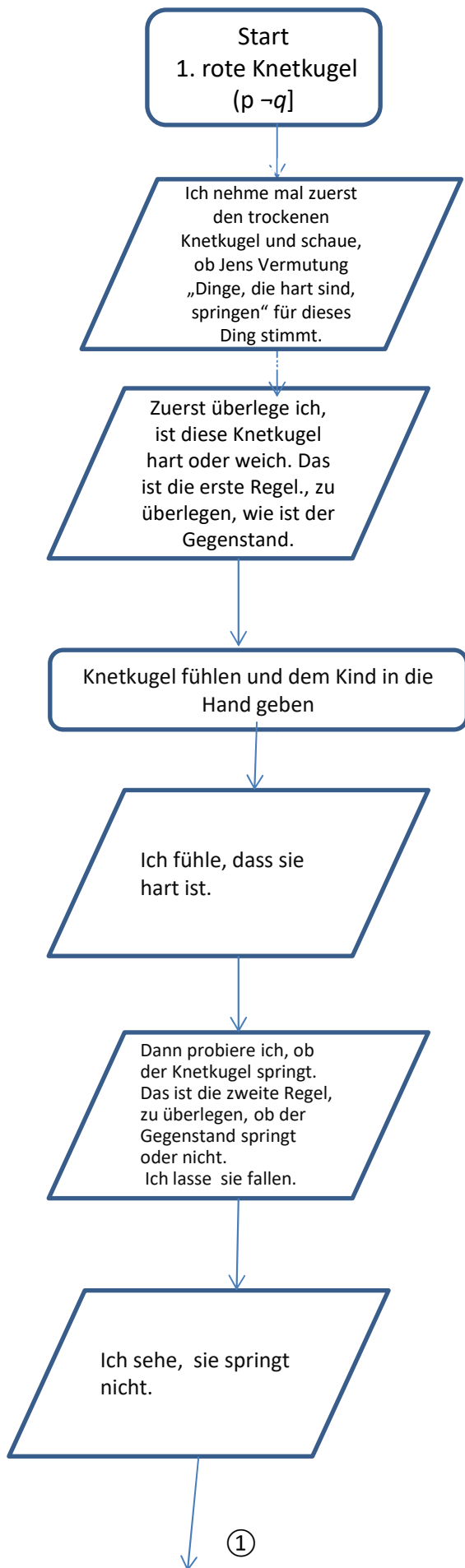
**„Auf welches Feld musst du Jens also stellen?“**

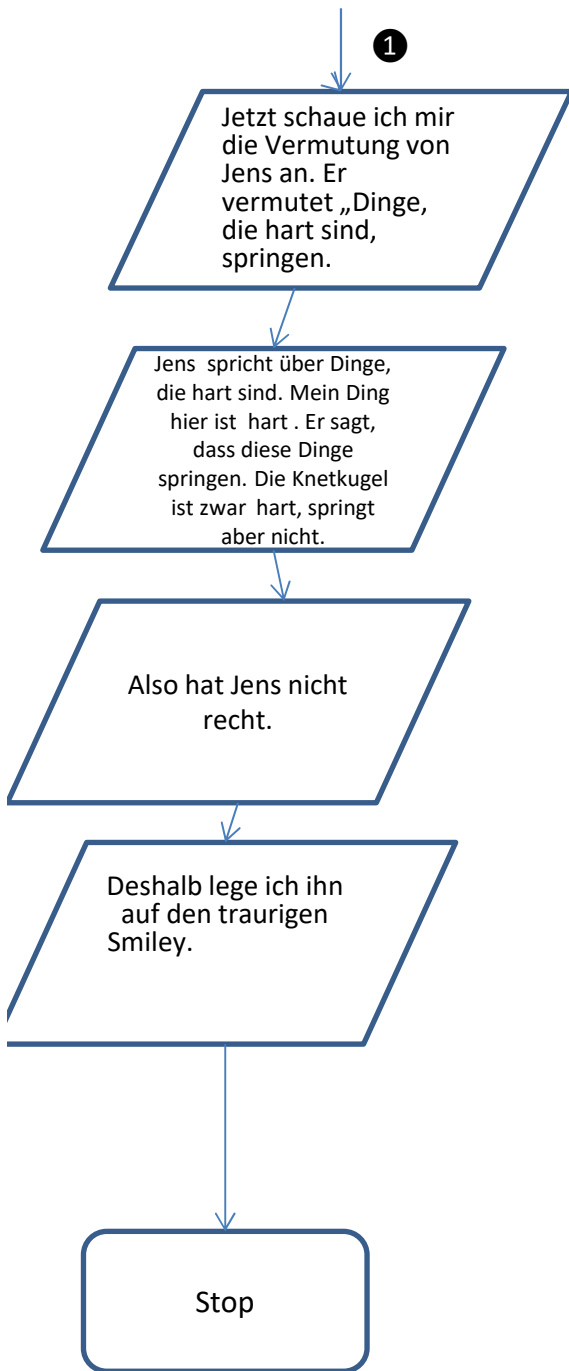
Kind antworten lassen (richtige Antwort: Feld „hart“, Symbolkarte „springt“)

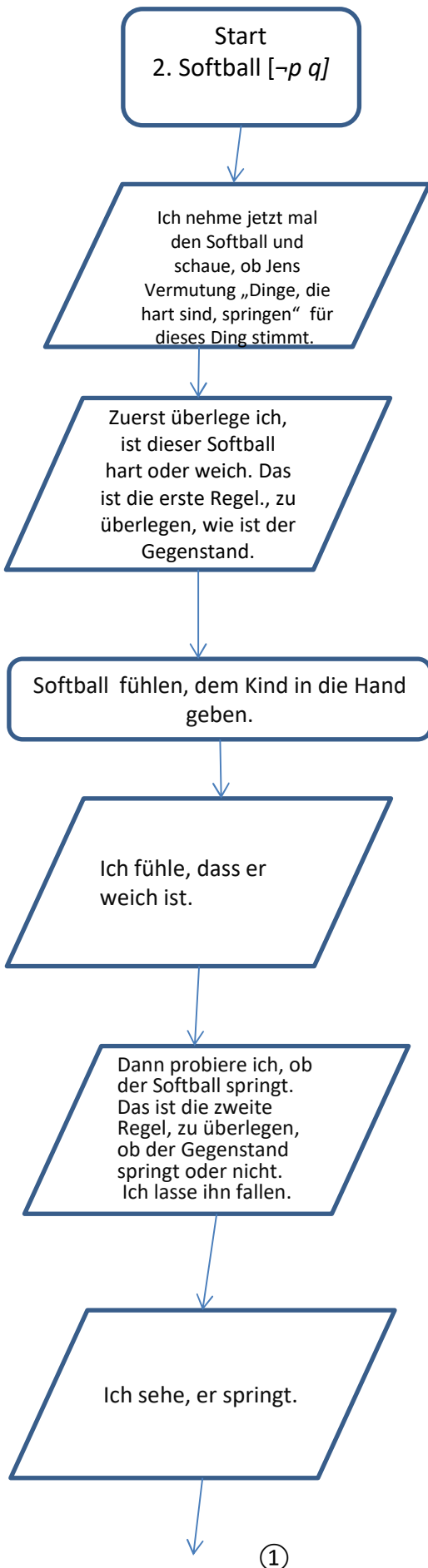
Solange fragen, bis die Spielfigur auf dem richtigen Feld steht.

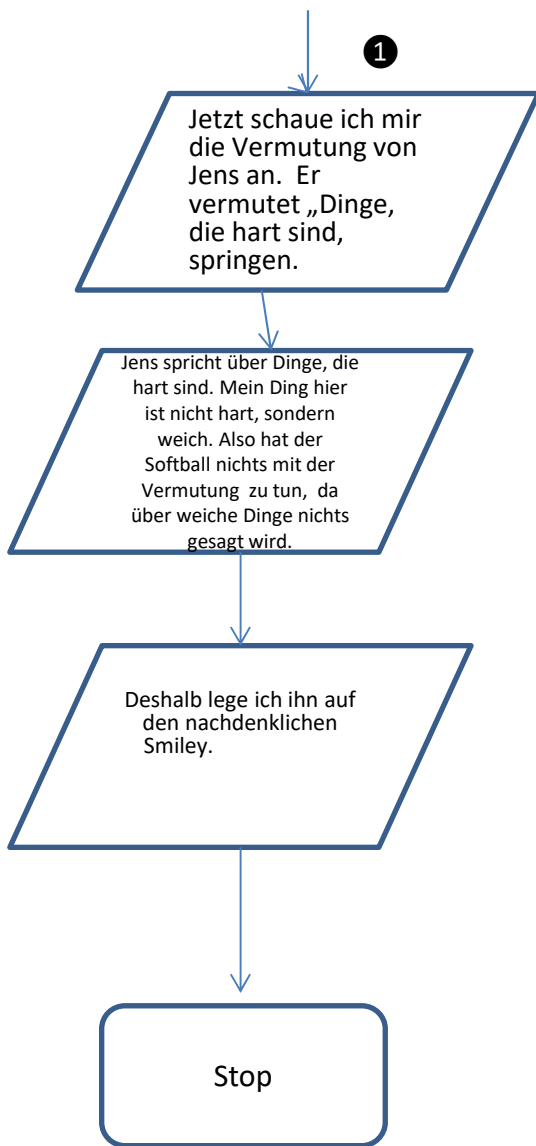
Protokoll 4b)

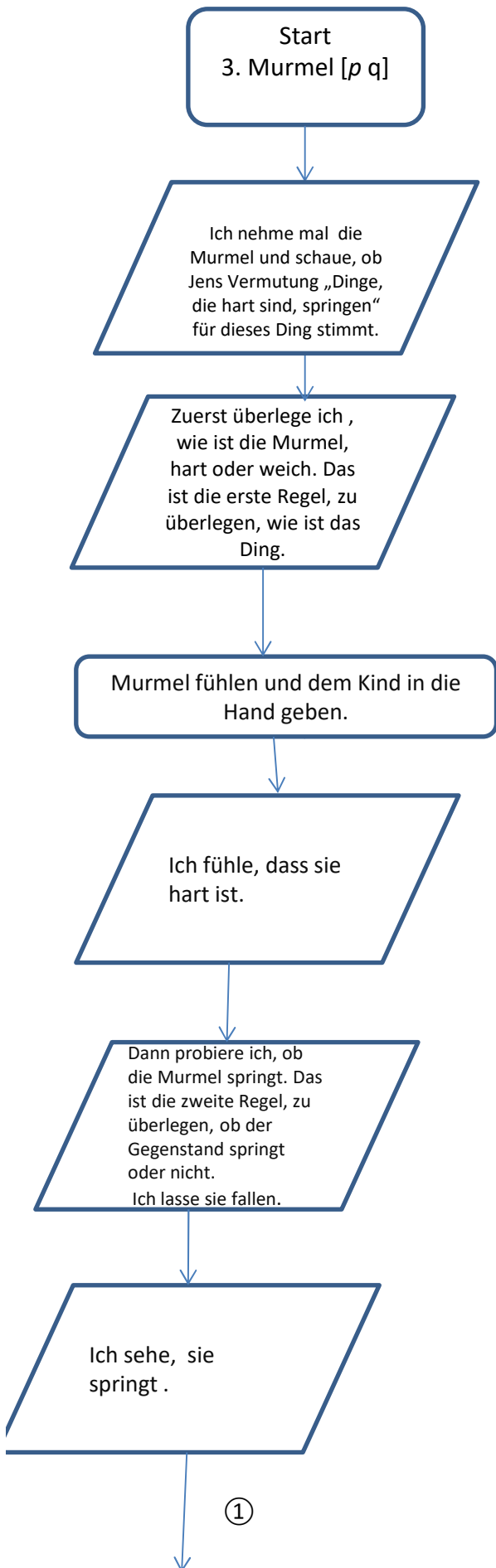
Protokoll 4c)



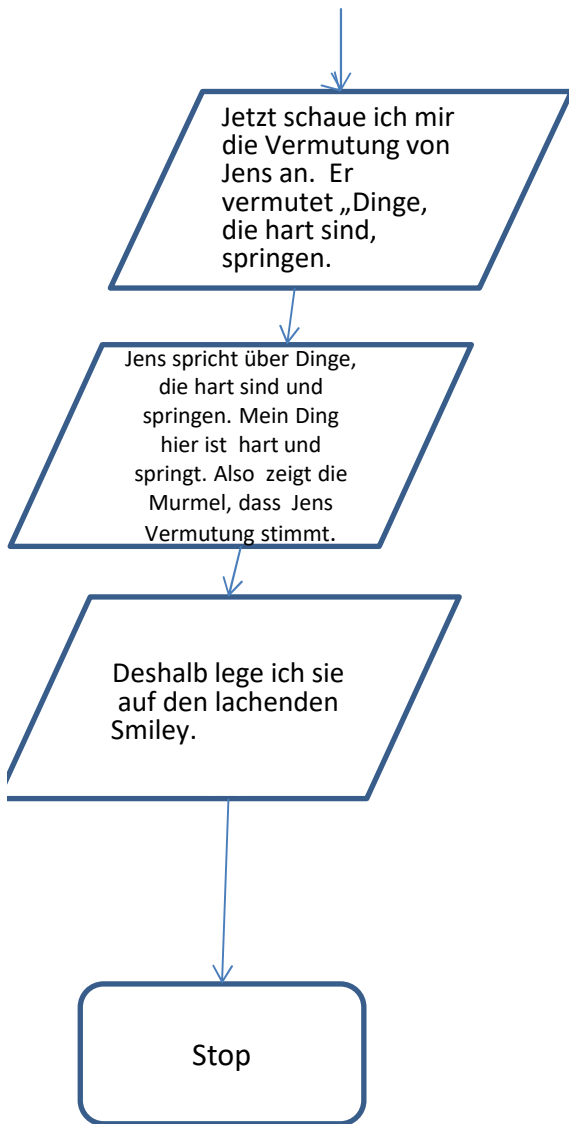


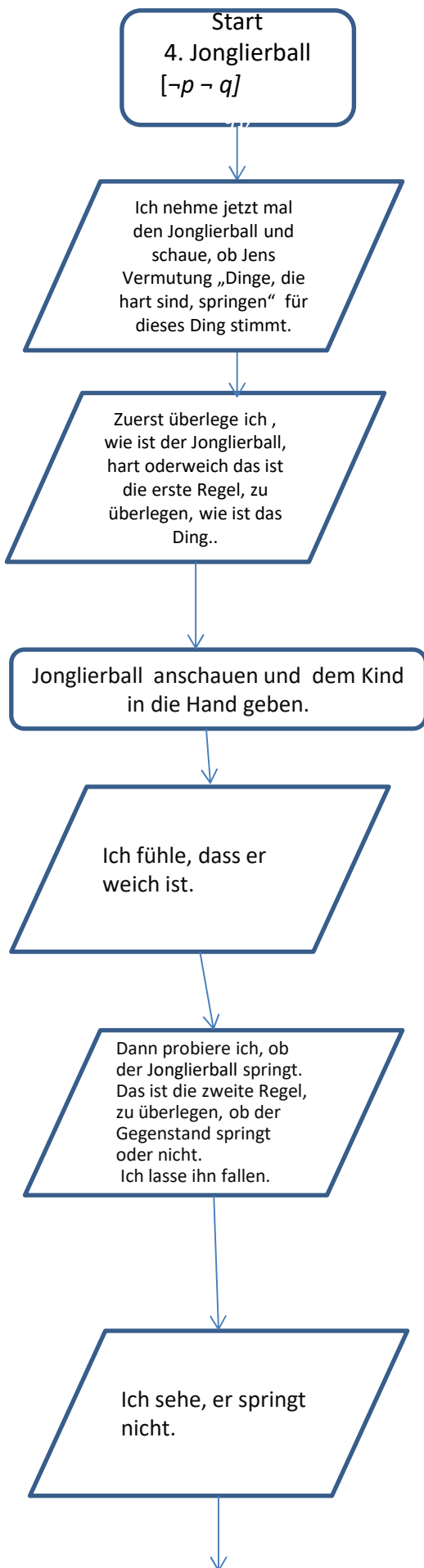


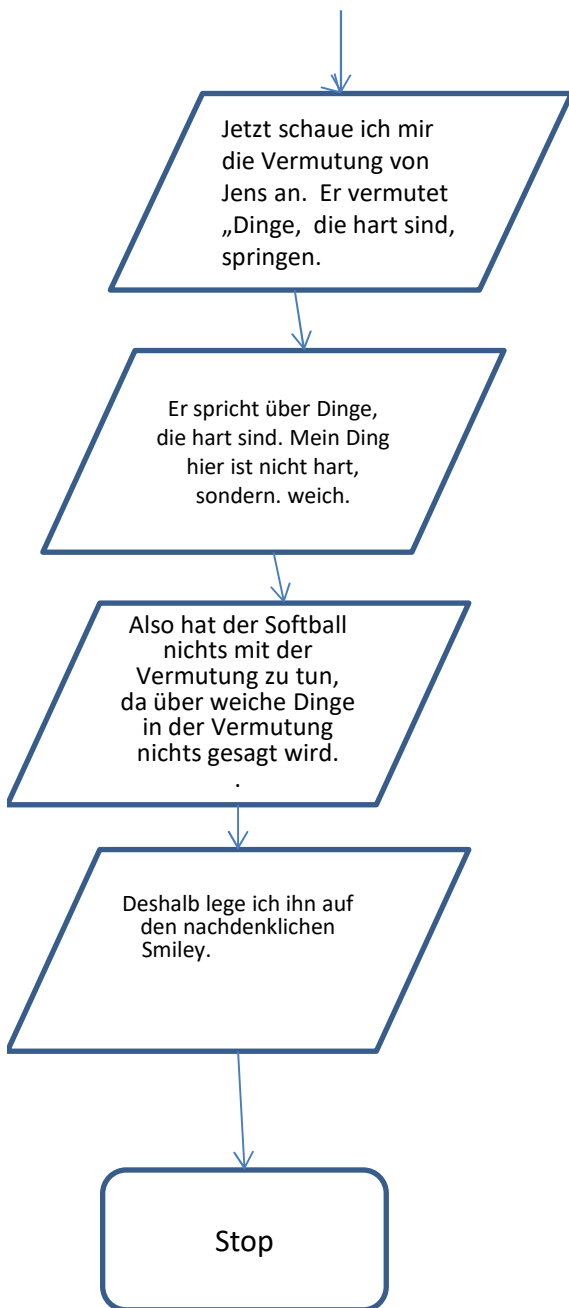




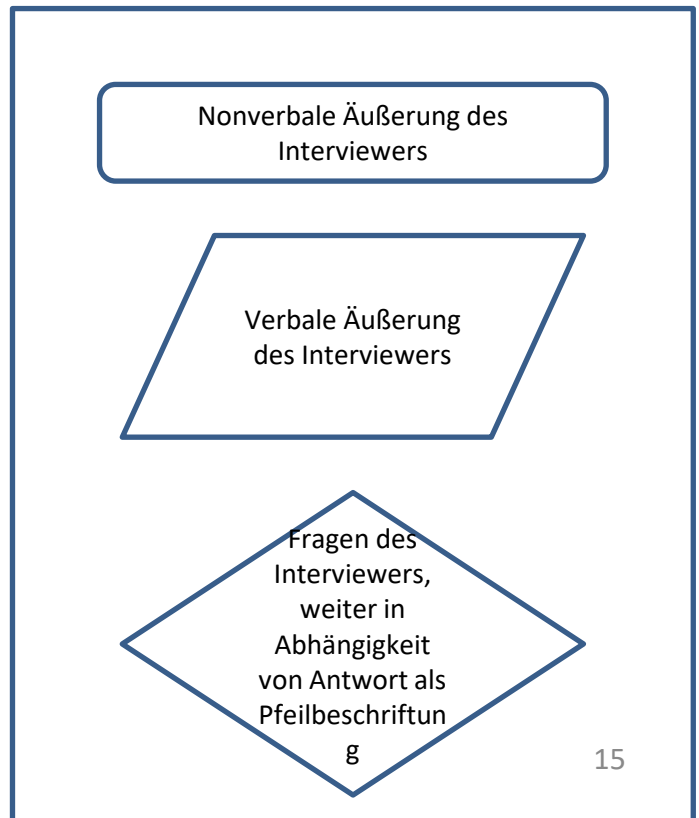








## Legende:



## Protokollbogen 2. Trainingssitzung (Scaffolding mit Modellierung und adaptivem Nachfragen)

Geschlecht: m w

TN:

In den einzelnen Spalten wird jeweils angekreuzt wie das Kind antwortet: **Richtig (r)**, **falsch (f)**; 3. Und 7. Wird in Stichpunkten protokolliert

Vermutung: „Dinge, die mit Luft sind, springen“		1. Eigen- schaft			2. Sprung- ver- halten			3. Zu- sammen- fassung			4. Jens			5. Zuord- nung Ball			6. Zusammen- fassung			7. Wahl des Kindes		
		a / b / c			a / b / c						a / b / c			a / b / c						a	b	
a) Wasserball [p -q]  b) Spielball [p q]  c) Würfel [-p q]  d) Aufg. Luftballon [-p -q]	Der Tennisball ist mit Luft und springt. [p q]																			Kein Protokoll	☺ ☹ ☹	☺ ☹ ☹
	Der Wasserball ist mit Luft und springt nicht. [p -q]																				☺ ☹ ☹	☺ ☹ ☹
	Der Würfel ist ohne Luft und springt [-p q]																				☺ ☹ ☹	☺ ☹ ☹
	Der aufgeschnittene Luftballon ist ohne Luft und springt nicht. [-p -q]																				☺ ☹ ☹	☺ ☹ ☹

## Interviewleitfaden 2. Trainingsitzung (Scaffolding mit Modellierung und adaptivem Nachfragen)

Vermutung „Dinge, die mit Luft gefüllt sind, springen.“

(Wasserball mit wenig Luft [p -q]; Spielball [p q]; Würfel [-p q]; aufgeschnittener Luftballon [-p -q];

Gegenstand Explizieren des Gedankenganges, 2.+3. Gegenstand zeigen der Gedanken bis zur Entscheidung, 4. Gegenstand adaptives Nachfragen

1. Sortieren der Gegenstände  
→ nach Eigenschaften  
Nach Sprungverhalten

2. Präsentieren der Vermutung  
→ Spielfigur einführen

3. Überprüfen der Vermutung  
Modeling /Fading Adaptives/  
Nachfragen  
Bei erstem Gegenstand  
Wasserball kein Protokoll bei 5  
6 und 7

# Sortieren der Gegenstände nach Eigenschaften

„Schau´ mal, hier sind vier Dinge: ein aufgeschnittener Luftballon, ein Spielball, ein aufblasbarer Wasserball und einen Würfel.“

*(Bälle zeigen)*

„Hier habe ich zwei kleine Luftkissen mitgebracht. Eins ist nicht aufgeblasen und flach, also ohne Luft.“*(auf nicht aufgeblasenes Luftkissen zeigen)*

und eins aufgeblasen, also mit Luft“*(auf aufgeblasenes Luftkissen zeigen)*.

„Nun sortiere doch mal die vier Dinge danach, ob sie mit Luft oder ohne Luft sind.

Die Dinge ohne Luft legst du hier hin *(auf nicht aufgeblasenes Luftkissen zeigen)*

und die Dinge mit Luft legst du hier hin *(auf aufgeblasenes Luftkissen zeigen)*.

Kind sortieren lassen. [richtige Lösung: Spielball, aufblasbarer Wasserball: mit Luft; aufgeschnittener Luftballon, Würfel: ohne Luft]

Protokoll 1a)

Bei Fehler:

„Schau doch noch einmal genau hin. Ist das Ding mit Luft oder ohne Luft?“

Kind anfassen lassen.

„Und wo würdest du es dann hinlegen?“

So lange fragen, bis alle Dinge auf dem richtigen Kissen liegen

Protokoll 1b)

Protokoll 1c)

„Jetzt wissen wir, welche Dinge mit Luft und welche ohne Luft sind.“

Bälle wegnehmen

(

# Sortieren der Gegenstände nach Sprungverhalten

„Nun wollen wir mal sehen, welche dieser Dinge springen und welche nicht springen können.“

„Dafür habe ich hier noch etwas mitgebracht.“

*(Symbolkarten für springt und springt nicht auf jedes Luftkissen legen.)*

„Nun lass die Dinge mal fallen um zu sehen, ob sie springen können oder nicht.“  
Kind ausprobieren lassen.

„Und nun sortiere die vier Dinge danach, ob sie springen oder nicht.“

Die Dinge, die mit Luft sind und springen legst du hier hin. *(Symbolkarte „springt“ auf Kissen mit Luft)*

die Dinge, die mit Luft sind und nicht springen legst du hier hin. *(Symbolkarte „springt nicht“ auf Kissen mit Luft)*

die Dinge, die ohne Luft sind und springen kommen hier hin. *(Symbolkarte „springt“ auf nicht- Kissen mit Luft)*

und die Dinge, die ohne Luft sind und nicht springen legst du hier hin. *(Symbolkarte „springt nicht“ auf Kissen mit Luft)*

Kind sortieren lassen. [richtige Lösung: Tennisball: mit Luft, springt; Wasserball: mit Luft, springt nicht; aufgeschnittener Luftballon: ohne Luft, springt nicht; Würfel: ohne Luft, springt]

Protokoll 2a)

Bei Fehler:

„Schau doch noch einmal genau hin. Ist das Ding mit Luft oder ohne Luft?“

Kind antworten lassen

und nun probiere doch noch einmal aus, ob das Ding springt oder nicht.

Kind ausprobieren lassen

Wo würdest du es jetzt hinlegen?“ So lange fragen, bis alle Dinge auf der richtigen Platte liegen

Protokoll 2b)

Protokoll 2c)

Bälle liegen lassen



# Zusammenfassung

„Und jetzt sag mir doch nochmal warum du die Bälle auf die jeweiligen Felder gelegt hast. Also warum liegt zum Beispiel der Wasserball dort?“

*(auf Wasserball bzw. Feld zeigen)*

Kind antworten lassen

Protokoll 3a)

„Und warum liegt der Spielball dort?“

*(Kind antworten lassen)*

Protokoll 3b)

„Was ist mit dem aufgeschnittene Luftballon?“

*(Kind antworten lassen)*

Protokoll 3c)

„Und warum liegt der Würfel auf diesem Feld?“

*(Kind antworten lassen)*

Protokoll 3d)

„Jetzt wissen wir, welche Dinge ohne Luft springen und welche nicht. Wir wissen welche Dinge mit Luft springen und welche nicht.“

## Erklärung: Was ist eine Vermutung?

„Wir wollen heute verschiedene Vermutungen überprüfen. Eine Vermutung ist etwas, das ich zwar glaube, aber noch überprüfen muss.“

*(Kästchen hervorholen und zeigen)*

„Ich vermute, dass in dem Kästchen ein Boxer versteckt ist. Was glaubst du?“ *(Kind antworten lassen)*

„Wie können wir herausbekommen, ob deine oder meine Vermutung stimmt?“ *(Kind antworten lassen)*

„Wir können doch einfach nachschauen, um festzustellen, ob deine oder meine Vermutung stimmt.“ *(Kind nachschauen lassen).*

„Ja, richtig, da ist eine Tänzerin drin.“

„Welche Vermutung stimmt, deine oder meine Vermutung? Oder hat keiner von uns beiden Recht?“ *(Kind antworten lassen).*

„Eine Vermutung kann stimmen, muss aber sie kann auch falsch sein. Das müssen wir immer überprüfen.“

# Präsentation der Vermutung

„Nun haben wir ja schon einige Kinder gefragt, was sie denken, warum ein Ball springt. Mia meint, dass Dinge die mit Luft sind, springen.“

„Das hier ist Mia. (*auf weibliche Playmobilfigur zeigen*)

Wo würdest du Mia nun hinstellen? Mia meint ja, dass Dinge die mit Luft sind springen. Die Dinge, von denen sie spricht, müssen also mit Luft sein und springen können. (Richtige Antwort: „Auf die Symbolkarte „springt“ auf dem Kissen mit Luft“.)

## Protokoll 4a)

Bei Fehler:

„Spricht Mia von Dingen mit Luft oder ohne Luft? Auf welches Kissen würdest du Mia also stellen?“

Kind antworten lassen (richtige Antwort: auf das Kissen mit Luft)

„Und, meint Mia, dass diese Dinge springen können oder, dass sie nicht springen können?“

Kind antworten lassen (richtige Antwort: springen)

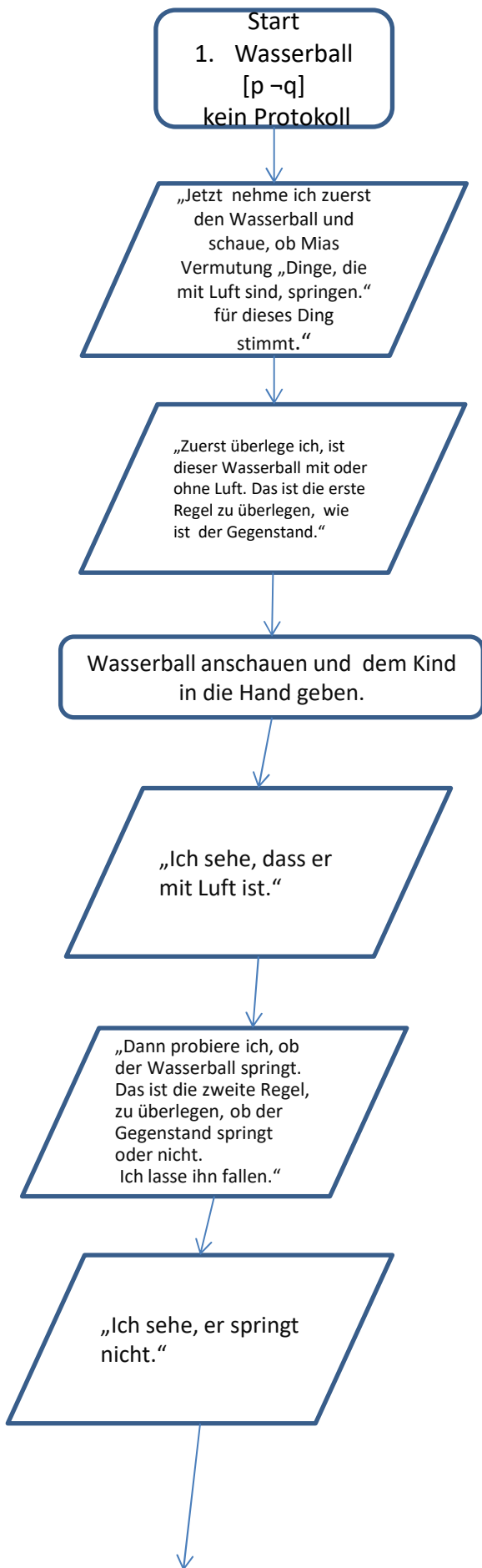
„Auf welches Feld musst du Mia also stellen?“

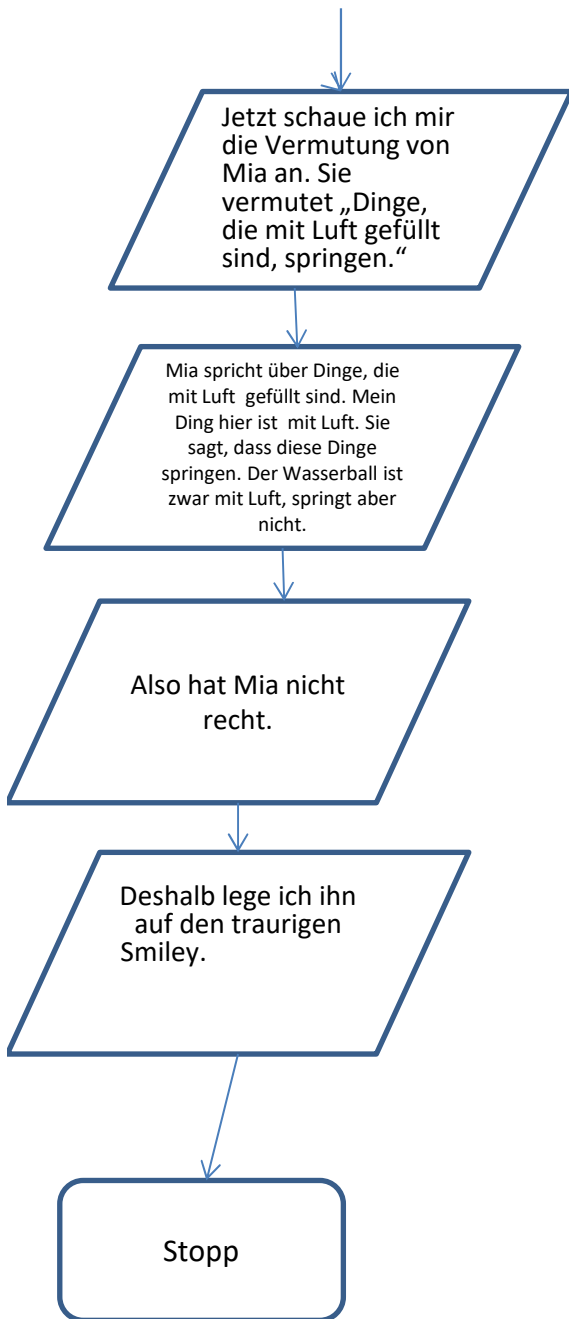
(*Kind antworten lassen*) (richtige Antwort: auf das Kissen mit Luft, Symbolkarte „springt“)

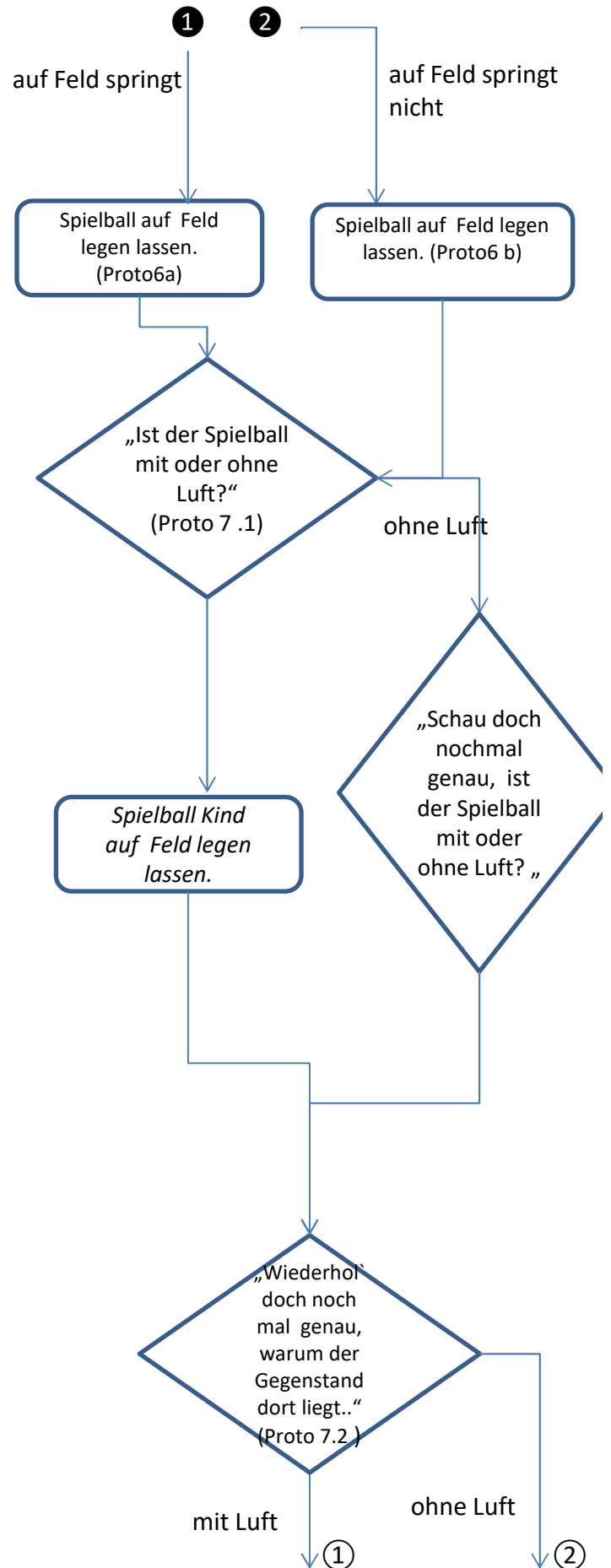
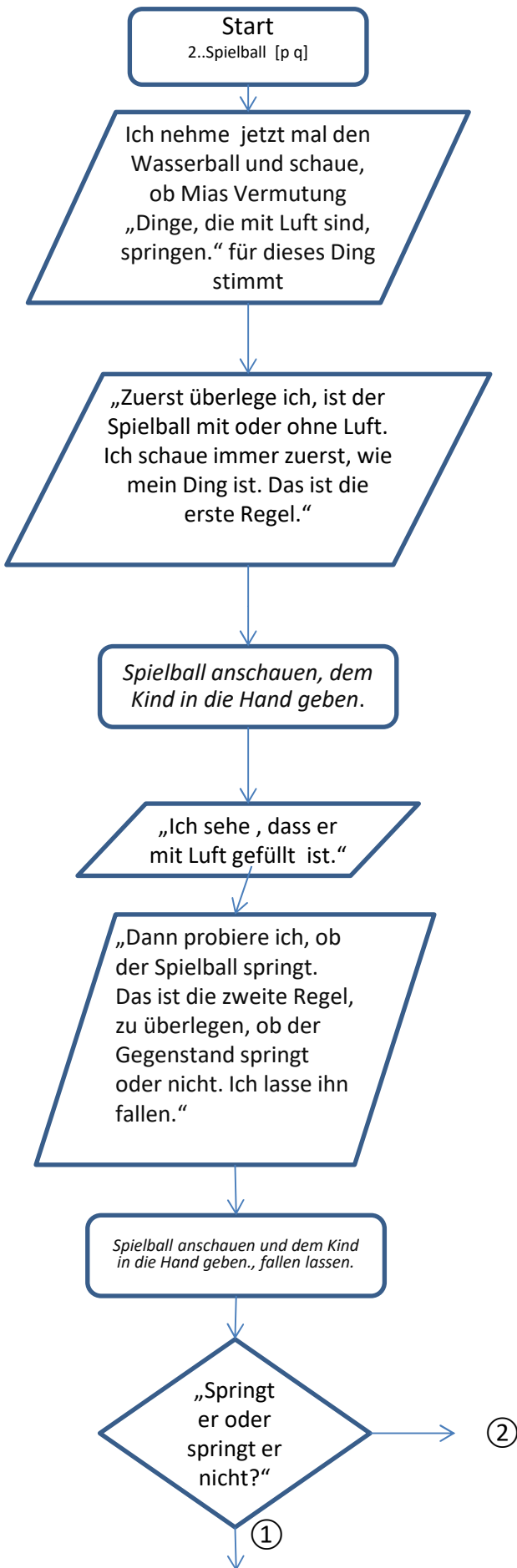
Solange fragen, bis die Spielfigur auf dem richtigen Feld steht.

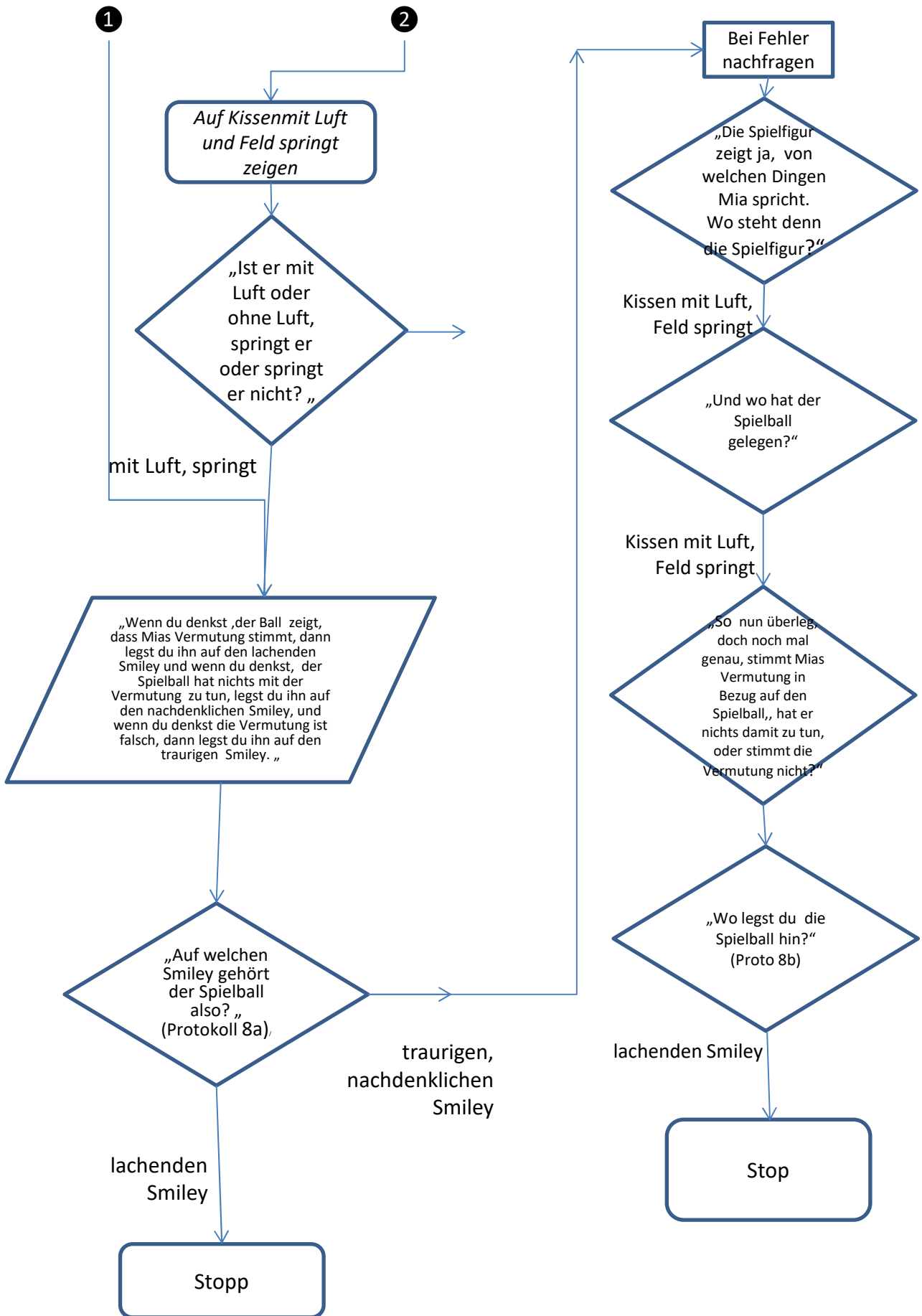
## Protokoll 4b)

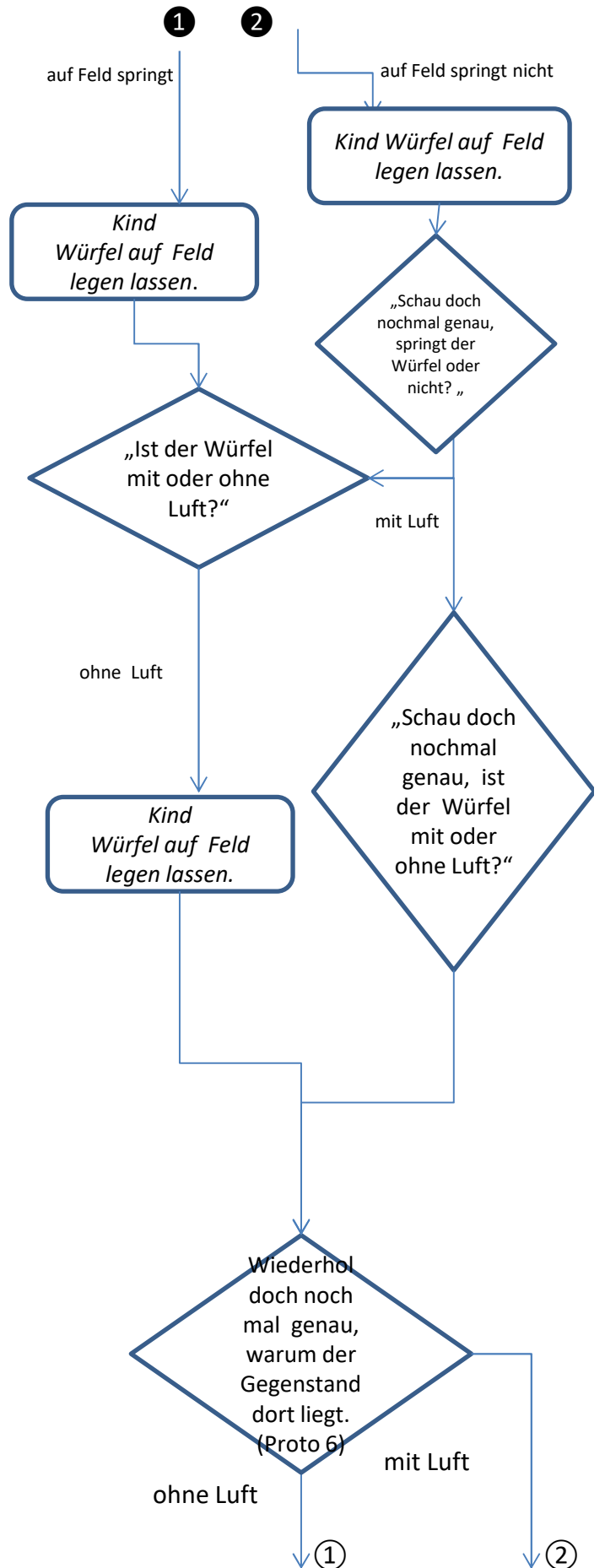
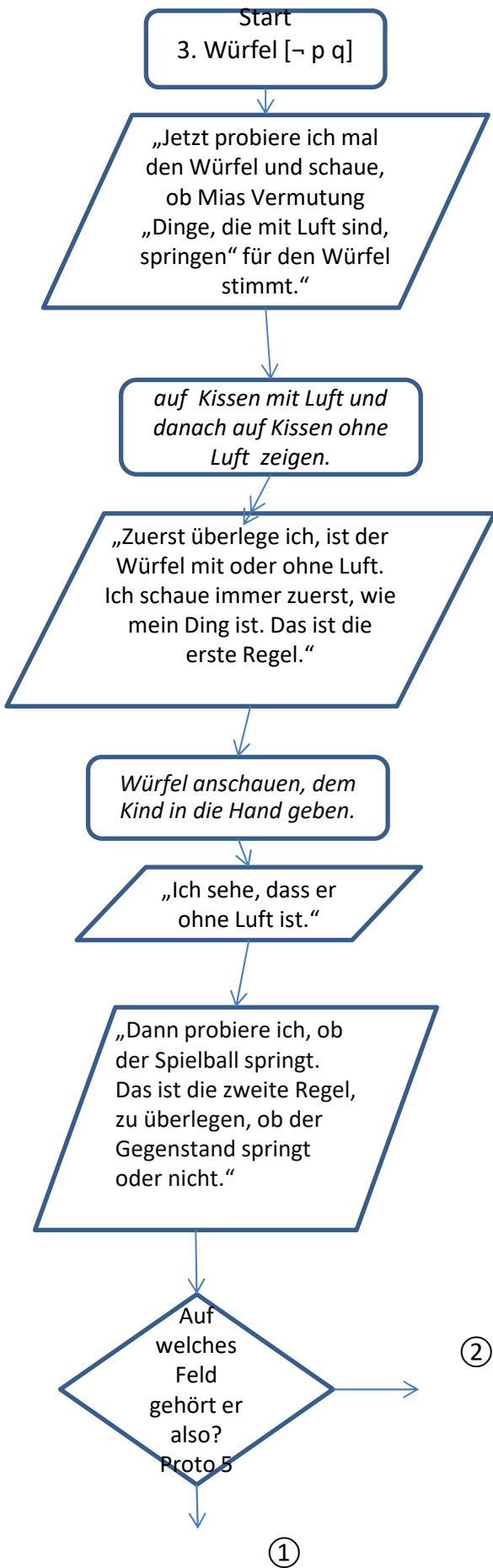
## Protokoll 4c)



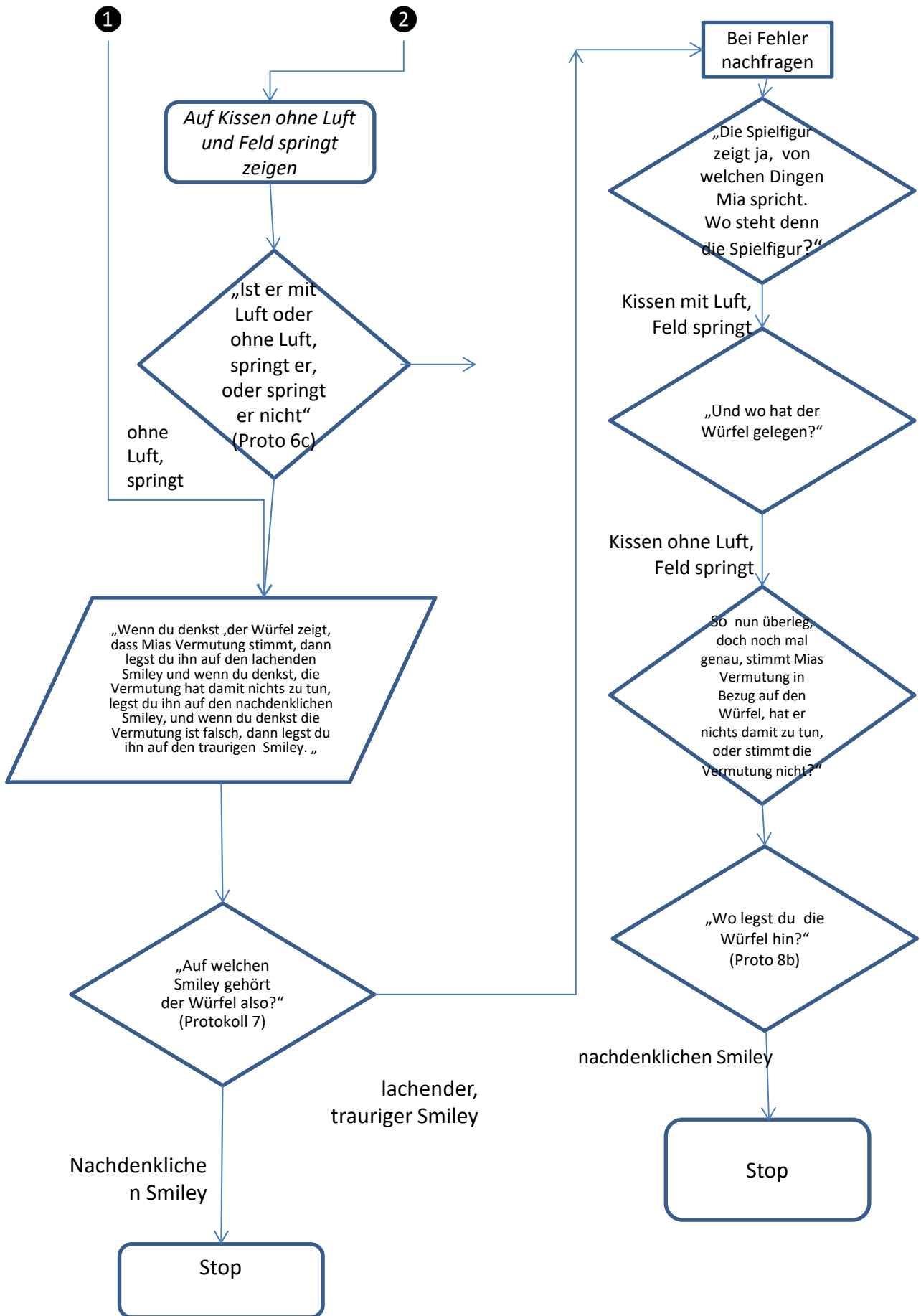


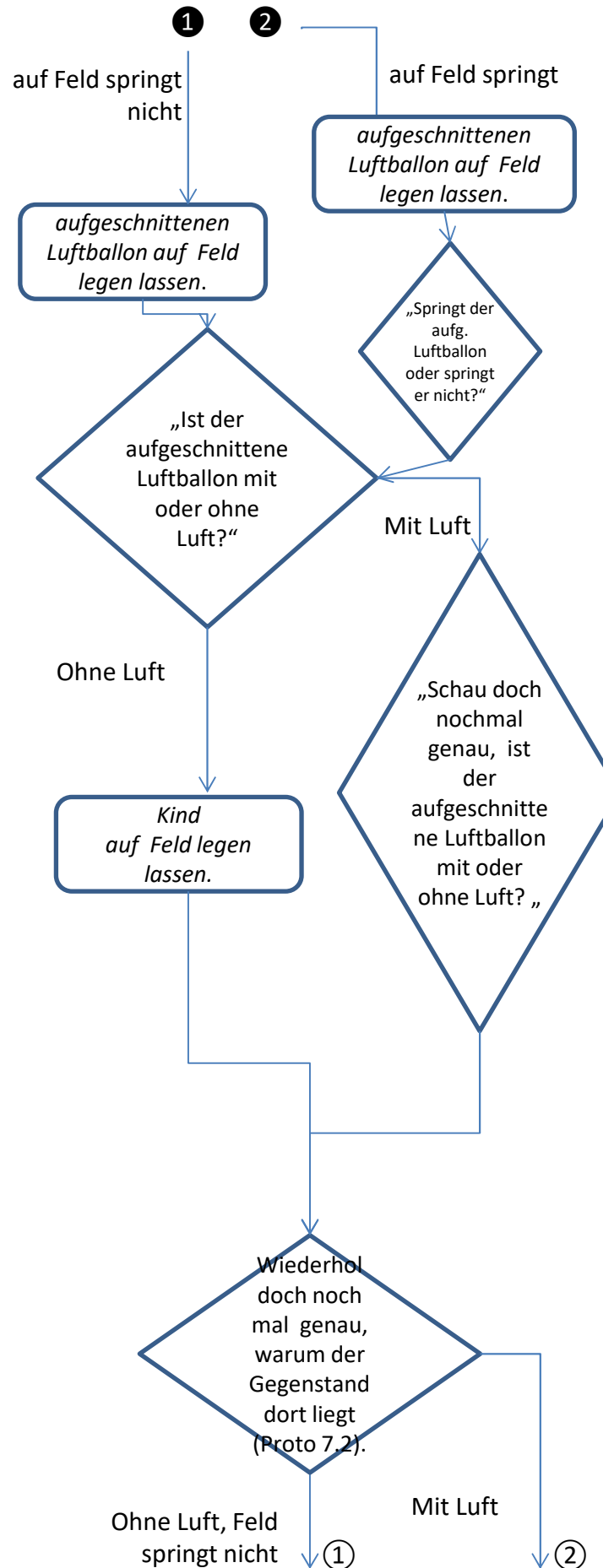
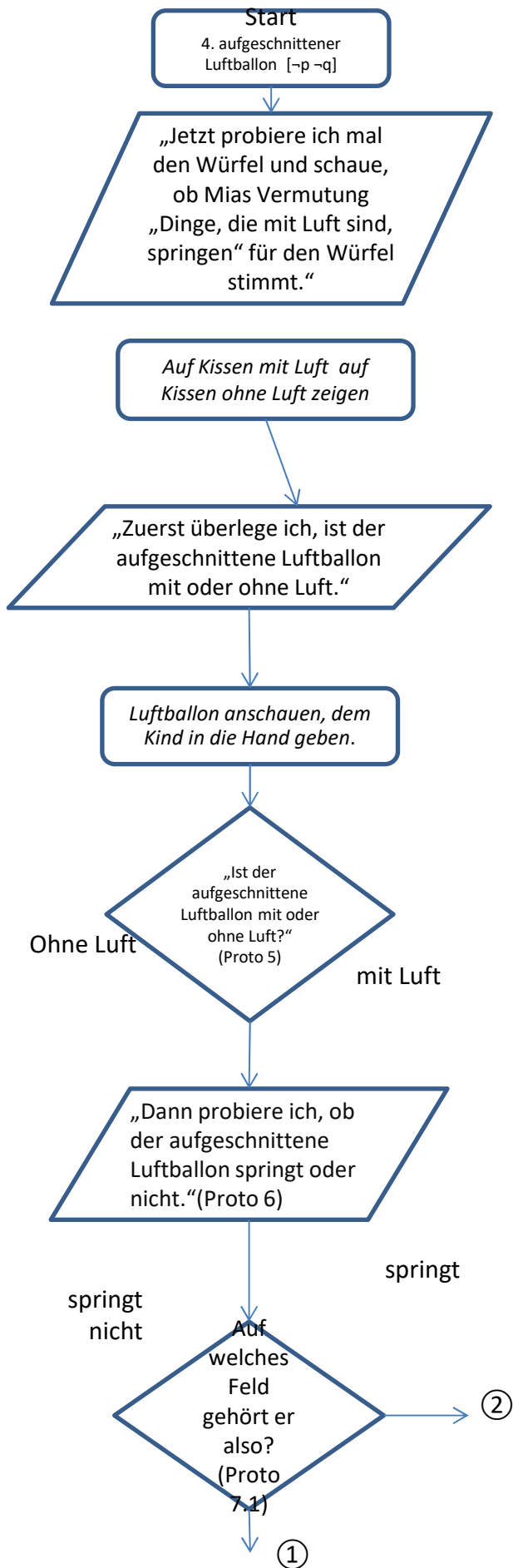


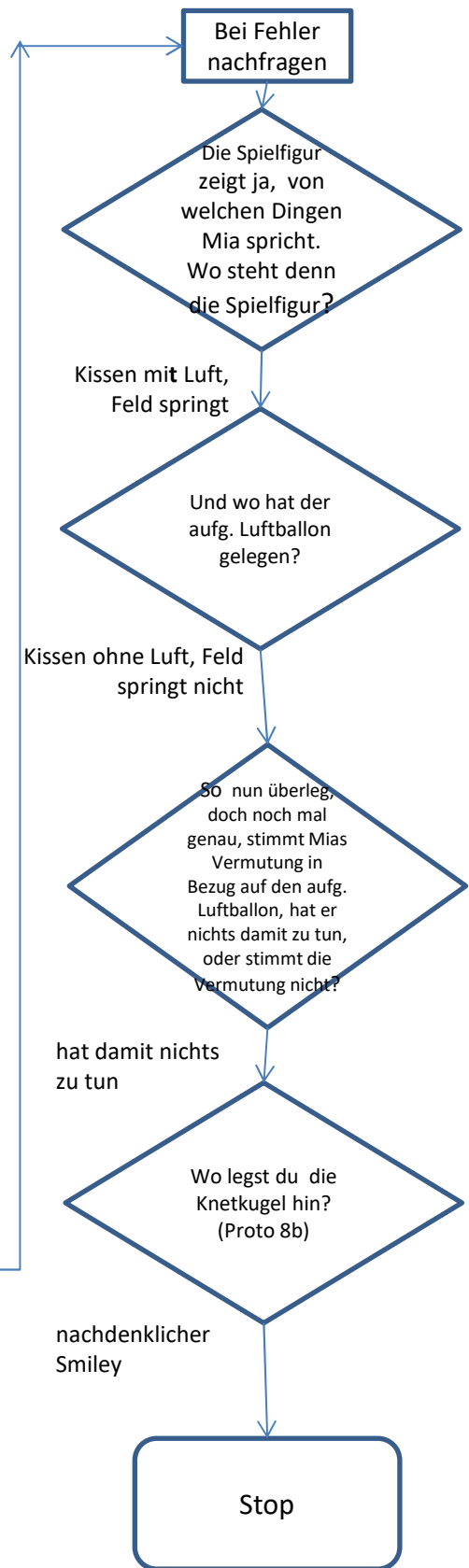
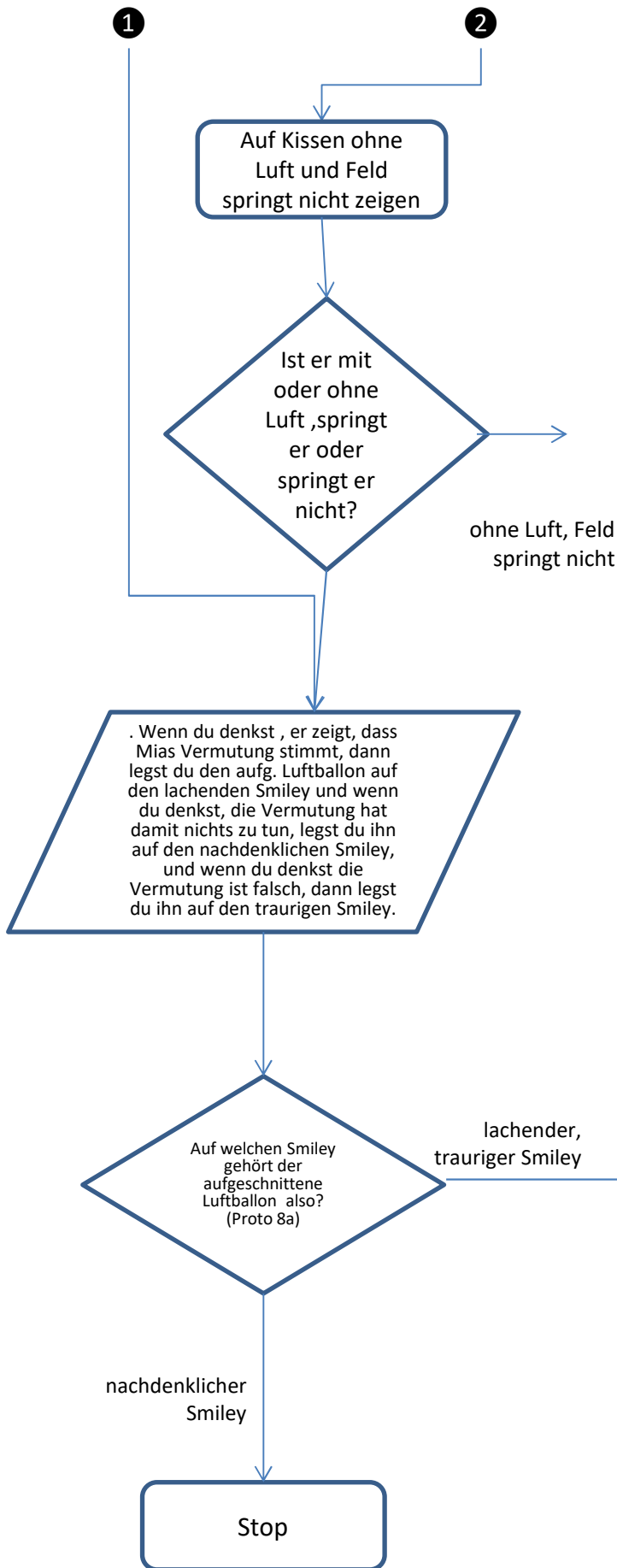




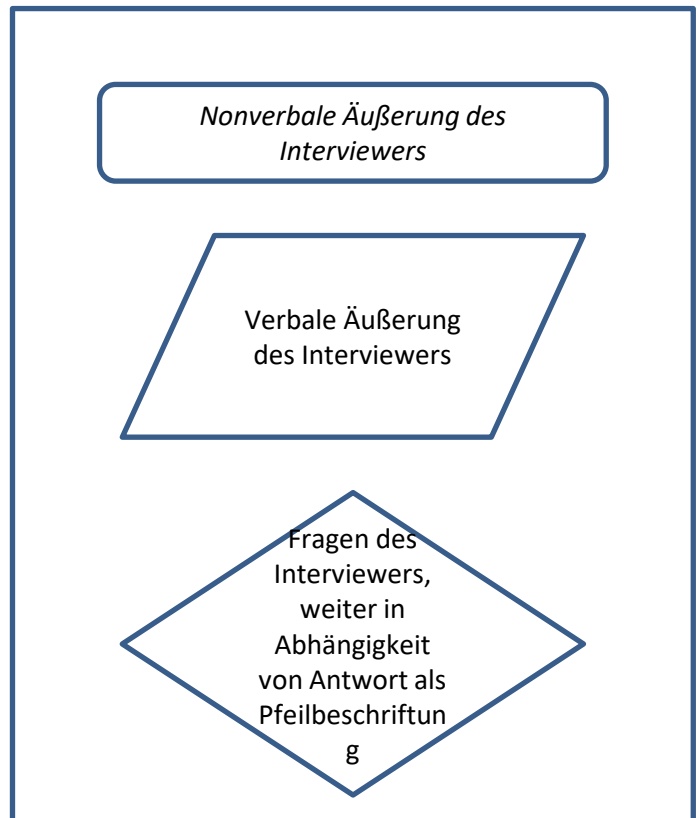








## Legende:



Geschlecht: m w

Geburtsmonat: \_\_\_\_\_

Geburtsjahr: 20

TN:

Naive Theorie	Formen der Evidenz	Sortierung			Soft Scaffolds [S]	
Beispiel	Der Tischtennisball ist mit Luft und schwimmt. [p q]				Scaff.	Nachfrage Scaffold
„Dinge, die mit Luft sind, schwimmen.“ [p q]; Bestantwort: konditionales Antwortmuster]  Spielball Fensterkitt-Ball Styroporstück Stein	Der Spielball ist mit Luft und schwimmt. [p q]		☺ ☹ ☹	☺ ☹ ☹	S	
	Der Fensterkitt Ball ist mit Luft und geht unter.[p -q]		☺ ☹ ☹	☺ ☹ ☹	S	
	Das Styroporstück ist ohne Luft und schwimmt. [-p q]		☺ ☹ ☹	☺ ☹ ☹	S	
	Der Stein ist ohne Luft und geht unter. [-p -q]		☺ ☹ ☹	☺ ☹ ☹	S	
„Dinge, klein sind, schwimmen.“ [p q]; Bestantwort: konditionales Antwortmuster]  5-Centstück Würfel aus Styropor Marmorstein Bimsstein	Ein Bimsstein ist klein und schwimmt [p q]		☺ ☹ ☹	☺ ☹ ☹		
	Das 5-Cent-Stück ist klein und geht unter. [p -q]		☺ ☹ ☹	☺ ☹ ☹	S	
	Der Würfel ist groß und schwimmt.[-p q]		☺ ☹ ☹	☺ ☹ ☹		
	Der Marmorstein ist groß und geht unter.[-p -q]		☺ ☹ ☹	☺ ☹ ☹		
„Dinge, die schwer sind, gehen unter.“ [p q] Bestantwort: konditionales Antwortmuster  Schildkröte Eisenstück Büroklammer Holzklotz	Der Eisenstück ist schwer und geht unter. [p q]		☺ ☹ ☹	☺ ☹ ☹		
	Der Holzklotz ist schwer und schwimmt [p -q]		☺ ☹ ☹	☺ ☹ ☹		
	Die Büroklammer ist leicht und geht unter.[-p q]		☺ ☹ ☹	☺ ☹ ☹		
	Die Schildkröte ist leicht und schwimmt. [-p -q]		☺ ☹ ☹	☺ ☹ ☹	S	

## Interviewleitfaden Transfer-Test

### „Schwimmen und Sinken“

„Guten Morgen.“

„Ich bin \_\_\_\_\_“ (*eigenen Namen nennen*).

„Das ist \_\_\_\_\_.“ (*Name der Begleitperson, wissenschaftliche Hilfskraft nennen*).

„Wie heißt du? (*Kind nennt seinen Namen*).

„Heute möchte ich herausfinden, wie Kinder in deinem Alter über bestimmte Dinge denken.“

„Das ist kein Test und es gibt kein „richtig“ oder „falsch“. Wichtig ist, dass du die Antwort nennst, die du am besten findest.“

Schau mal.“ Tischtennisball und Knetkugel ins Wasser tun.

„Was hast du beobachtet?“ (*Antwort abwarten*).

„Genau, der Tischtennisball schwimmt, die Knetkugel geht unter.“

„Wir sprechen nur von schwimmen, wenn das Ding sich auf dem Wasser hält und nicht sinkt. Wenn das Ding auf den Boden der Kiste sinkt, sagen wir, dass es untergeht. Hast du das verstanden? (ggf. noch einmal erklären).“

„Für die Denkaufgaben sind wir zu einem Kindergarten gefahren und haben Kinder gefragt, welche Dinge schwimmen. Diese unterschiedlichen Vermutungen haben wir gesammelt.“

„Eine Vermutung ist etwas, das ich zwar glaube, was ich aber noch überprüfen muss.“

## 1. Vermutung: „Dinge, die leicht sind, schwimmen.“

Styroporstück [ p q]; Nadel [p ¬q]; Holzstück [¬ p q]; Tonklumpen [¬ p ¬q]

Sortieren der Gegenstände nach Eigenschaften:

„Schau mal, hier sind vier Dinge: ein Styroporstück, eine Nadel, eine Kerze, ein Tonklumpen  
(Gegenstände zeigen)

„Sortiere doch mal die leichten Dinge hier hin (auf linke Seite zeigen) und die schweren Dinge  
hierhin.“ (auf rechte Seite zeigen).

Kind sortieren lassen. (Richtige Lösung: Styroporstück und Nadel leicht, Kerze, Tonklumpen schwer)

Bei Fehler:

„Überleg´ doch mal, ist das Ding leicht oder schwer?“

(Kind antworten lassen.)

(Maximal 2x helfen).

„Das ist Tim.“

„Ich lese Tims Vermutung, welche Dinge springen, einmal vor:

„Dinge, die leicht sind, schwimmen.“

„Die Vermutung wollen wir jetzt überprüfen.“

1. Styroporkugel [ p q]

„Dazu nehme ich das Styroporstück hier aus der Kiste“ (Styroporstück zeigen). „Das Besondere an  
ihm ist, dass es leicht ist. Gut. Jetzt probieren wir aus, ob das Styroporstück schwimmt.“

„Schau genau hin.“ (Styroporstück fallen lassen). „Hast du gesehen, was passiert ist?“ (Antwort des  
Kindes abwarten). „Ja, richtig. Das Styroporstück schwimmt, weil es sich oben auf dem Wasser  
bewegt. Jetzt denke nochmal an Tim und seine Vermutung.“ „Tim vermutet ja: „Dinge, die leicht  
sind, schwimmen.“ „Was zeigt dir das Styroporstück über Tims Vermutung?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du das Styroporstück auf den lachenden Smiley.“  
(Auf den Smiley zeigen).

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun legst du Styroporstück auf den nachdenklichen Smiley.“ (Auf den Smiley zeigen).

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du Styroporstück auf den traurigen, Smiley.  
(Auf den Smiley zeigen).

*Kind legt Styroporstück auf einen Smiley.*

Bei Fehler:

„Überleg´ doch mal, ist das Ding leicht oder schwer?“

*(Kind antworten lassen).*

„Überleg´ doch noch mal, schwimmt das Ding oder nicht?“

*(Kind antworten lassen).*

*(Maximal 2x helfen)*

2. Nadel [p -q];

„Jetzt nehme ich die Nadel, um Tims Vermutung zu überprüfen.“

„Tim vermutet:„ Dinge, die leicht sind, schwimmen.“

„Die Vermutung wollen wir jetzt überprüfen.“

„Das Besondere an ihr ist, dass sie leicht ist. Gut. Jetzt probieren wir aus, was probieren wir aus, ob sie schwimmt.“

„Schau genau hin.“ (Nadel fallen lassen). „Hast du gesehen, was passiert ist?“ (Antwort des Kindes abwarten). „Ja, richtig. Die Nadel schwimmt nicht.“



„Die Nadel schwimmt nicht, sie sinkt nach unten auf den Boden Jetzt denke nochmal an Tim und seine Vermutung.“ „Tim vermutet ja: „Dinge, die leicht sind, schwimmen.“ „Was zeigt dir die Nadel über Tims Vermutung?“

„Ich fasse nochmal zusammen: Die Nadel ist leicht und schwimmt nicht.“

„Wenn du glaubst, Tims Vermutung stimmt, legst du die Nadel hierhin.“ (*Auf den lachenden Smiley zeigen*).

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun legst du die Nadel hierhin.“ (Auf den nachdenklichen Smiley zeigen.)

Wenn du glaubst, Tims Vermutung stimmt nicht, legst du die Nadel hierhin. (*Auf den traurigen Smiley zeigen*).

Bei Fehler:

„Überleg' doch mal, ist das Ding leicht oder schwer?“

(*Kind antworten lassen*).

„Überleg' doch noch mal, schwimmt das Ding oder nicht?“

(Kind antworten lassen).

(*Maximal 2x helfen*).

3. Holzstück[~ p q];

„Jetzt nehme ich ein Holzstück, um Tims Vermutung zu überprüfen.“

„Tim vermutet: ‚Dinge, die leicht sind, schwimmen.‘

„Die Vermutung wollen wir jetzt überprüfen.“

„Das Besondere am Holzstück ist, dass es schwer ist. Gut. Jetzt probieren wir aus, ob es schwimmt.“

„Schau genau hin.“ (*Holzstück fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig. Das Holzstück schwimmt nicht.“

„Das Holzstück schwimmt nicht, es sinkt nach unten. Jetzt denke nochmal an Tim und seine Vermutung.“ ,Tim vermutet ja: „Dinge, die leicht sind, schwimmen.“ „Was zeigt dir das Holzstück über Tims Vermutung?“

„Ich fasse nochmal zusammen: Das Holzstück ist schwer und schwimmt nicht.“

„Wenn du glaubst, Tims Vermutung stimmt, legst du das Holzstück hierhin.“ (*Auf den lachenden Smiley zeigen*).

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun legst du das Holzstück hierhin.“ (*Auf den nachdenklichen Smiley zeigen*).

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du das Holzstück hierhin.“ (*Auf den Smiley zeigen*).

Bei Fehler:

„Überleg´ doch mal, ist das Ding leicht oder schwer?“

(*Kind antworten lassen*).

„Überleg´ doch noch mal, schwimmt das Ding oder nicht?“

(*Kind antworten lassen*).

(Maximal 2x helfen)

Tonklumpen [-p -q];

„Jetzt überprüfen wir noch einmal Tims Vermutung.“

„Tim vermutet: ‚Dinge, die leicht sind, schwimmen. ‘“

„Dazu nehme ich den großen Tonklumpen hier aus der Kiste.“ (*Großen Tonklumpen zeigen*). „Das Besondere an ihm ist, dass er schwer ist. Gut. Jetzt probieren wir aus, ob der Tonklumpen schwimmt.“ (*Großen Tonklumpen fallen lassen*).

„Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig. Der große Tonklumpen schwimmt nicht.“

„Jetzt denk´ nochmal an Tim und seine Vermutung.“ „Tim vermutet ja: ‚Dinge, die leicht sind, schwimmen. ‘ Was zeigt dir der große Tonklumpen über Tims Vermutung?“

„Wenn du glaubst, Tims Vermutung stimmt, legst du den großen Tonklumpen hierhin.“ (*Auf den lachenden Smiley zeigen*).

„Wenn du glaubst, Tims Vermutung hat damit nichts zu tun, legst du den großen Tonklumpen auf hierhin.“ (*Auf den nachdenklichen Smiley zeigen*).

„Wenn du glaubst, Tims Vermutung stimmt nicht, legst du den großen Tonklumpen hierhin.“ (*Auf den traurigen Smiley zeigen*).

*Kind legt den großen Tonklumpen auf einen Smiley.*

Bei Fehler:

„Überleg´ doch mal, ist das Ding leicht oder schwer?“

(*Kind antworten lassen*).

„Überleg´ doch noch mal, schwimmt das Ding oder nicht?“

(*Kind antworten lassen*).

(*Maximal 2x helfen*)

## 2. Vermutung: „Dinge, die klein sind, schwimmen.“

5-Centstück [p  $\neg$ q] ; Würfel aus Styropor [ $\neg$ p q] ; Marmorstein [ $\neg$ p  $\neg$ q];

Bimsstein [p q]

Sortieren der Gegenstände nach Eigenschaften:

„Schau mal, hier sind vier Dinge: Bimsstein, Würfel aus Styropor, Marmorstein 5-Centstück.“  
(Gegenstände zeigen).

„Sortiere doch mal die kleinen Dinge hier hin (auf linke Seite zeigen) und die großen Dinge hierhin (auf rechte Seite zeigen).

Kind sortieren lassen. (richtige Lösung: Bimsstein und 5-Centstück klein; Marmorstein und Würfel aus Styropor groß).

Bei Fehler:

„Überleg' doch mal, ist das Ding leicht oder schwer?“

(Kind antworten lassen).

(Maximal 2x helfen)

„Das ist Paul.“

„Er vermutet: ‚Dinge, die klein sind, schwimmen.‘“

1. 5-Centstück [p  $\neg$ q]

„Dazu nehme ich das 5-Centstück hier aus der Kiste.“ (5-Centstück zeigen). „Das Besondere an ihm ist, dass es klein ist. Gut. Jetzt probieren wir aus, ob das 5-Centstück schwimmt. (5-Centstück fallen lassen).

„Hast du gesehen, was passiert ist?“ (Antwort des Kindes abwarten). „Ja, richtig. Das 5-Centstück ist klein und schwimmt nicht.“

Jetzt denke nochmal an Paul und seine Vermutung. Paul vermutet ja: ‚Dinge, die leicht sind, schwimmen.‘ Was zeigt dir 5-Centstück über Pauls Vermutung? “

„Wenn du glaubst, Paul Vermutung stimmt, legst du das 5-Centstück auf den lachenden Smiley.“  
(Auf den Smiley zeigen).

„Wenn du glaubst, Paul Vermutung hat damit nichts zu tun, legst du das 5-Centstück hierhin. *(Auf den nachdenklichen Smiley zeigen).*

„Wenn du glaubst, Paul Vermutung stimmt nicht, legst du das 5-Centstück hierhin.“ *(Auf den traurigen Smiley zeigen.)*

*Kind legt das 5-Centstück auf einen Smiley.*

„Jetzt überprüfen wir Pauls Vermutung ‚Dinge, die klein sind, schwimmen‘ mit einem anderen Ding.“

## 2. Würfel aus Styropor [-p q]

„Dazu nehme ich den Würfel hier aus der Kiste.“ *(Würfel zeigen).* „Das Besondere an ihm ist, dass es groß ist. Gut. Jetzt probieren wir aus, ob der Würfel schwimmt.“ *(Würfel fallen lassen).* „Hast du gesehen, was passiert ist?“ *(Antwort des Kindes abwarten).* „Ja, richtig. Der Würfel ist groß und schwimmt.“

„Jetzt denke nochmal an Paul und seine Vermutung.“

„Paul vermutet ja: ‚Dinge, die klein sind, schwimmen.‘ „Was zeigt dir der Würfel über Pauls Vermutung?“

„Wenn du glaubst, Paul Vermutung stimmt, legst du den Würfel hierhin.“ *(Auf den lachenden Smiley zeigen).*

Wenn du glaubst, Paul Vermutung hat damit nichts zu tun, legst du den Würfel hierhin. *(Auf den nachdenklichen Smiley zeigen.)*

Wenn du glaubst, Paul Vermutung stimmt nicht, legst du den Würfel auf den traurigen Smiley. *(Auf den Smiley zeigen.)*

*Kind legt den Würfel auf einen Smiley.*

Bei Fehler:

„Überleg´ doch mal, ist das Ding leicht oder schwer?“

*(Kind antworten lassen).*

„Überleg´ doch noch mal, schwimmt das Ding oder nicht?“

*(Kind antworten lassen).*

(Maximal 2x helfen)

„Jetzt überprüfen wir Pauls Vermutung. Er vermutet: ‚Dinge, die klein sind, schwimmen.‘“

### 3. Gegenstand [-p -q]

„Dazu nehme ich den Marmorstein hier aus der Kiste.“ (*Marmorstein zeigen*). „Das Besondere an ihm ist, dass er groß ist. Gut. Jetzt probieren wir aus, ob der Marmorstein schwimmt.“

„Schau genau hin.“ (*Marmorstein fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig. Der Marmorstein ist groß und schwimmt nicht.“

„Jetzt denke nochmal an Paul und seine Vermutung.“

„Paul vermutet ja: ‚Dinge, die klein sind, schwimmen.‘ Was zeigt dir der Marmorstein über Pauls Vermutung?“

„Wenn du glaubst, Pauls Vermutung stimmt, legst du den Marmorstein hierhin.“ (*Auf den lachenden Smiley zeigen*).

Wenn du glaubst, Pauls Vermutung hat damit nichts zu tun, legst du den Marmorstein hierhin. (*Auf den nachdenklichen Smiley zeigen*).

Wenn du glaubst, Pauls Vermutung stimmt nicht, legst du den Marmorstein auf den traurigen Smiley. (*Auf den traurigen Smiley zeigen*).

*Kind legt den Marmorstein auf einen Smiley.*

Bei Fehler:

„Überleg´ doch mal, ist das Ding leicht oder schwer?“

(*Kind antworten lassen*).

„Überleg´ doch noch mal, schwimmt das Ding oder nicht?“

(*Kind antworten lassen*).

(*Maximal 2x helfen*)

„Jetzt nehmen wir noch ein Ding, um Pauls Vermutung zu überprüfen.“

Er vermutet ‚Dinge, die klein sind, schwimmen.‘“

#### 4.. Gegenstand [p q]

„Dazu nehme ich den Bimsstein hier aus der Kiste.“ (*Bimsstein zeigen*). „Das Besondere an ihm ist, dass er klein ist. Gut. Jetzt probieren wir aus, ob der Bimsstein schwimmt.“ (*Bimsstein fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig. Der Bimsstein ist klein und schwimmt.“

„Jetzt denke nochmal an Paul und seine Vermutung. „Paul vermutet ja: ‚Dinge, die klein sind, schwimmen.‘ Was zeigt dir der Bimsstein über Pauls Vermutung?“

„Wenn du glaubst, Paul Vermutung stimmt, legst du den Bimsstein hierhin.“ (*Auf den lachenden Smiley zeigen.*)

Wenn du glaubst, Paul Vermutung hat damit nichts zu tun, legst du den Bimsstein hierhin. (*Auf den nachdenklichen Smiley zeigen.*)

Wenn du glaubst, Paul Vermutung stimmt nicht, legst du den Bimsstein auf den traurigen Smiley. (*Auf den traurigen Smiley zeigen.*)

Kind legt den Bimsstein auf einen Smiley.

Bei Fehler:

„Überleg´ doch mal, ist das Ding leicht oder schwer?“

(*Kind antworten lassen*).

„Überleg´ doch noch mal, schwimmt das Ding oder nicht?“

(*Kind antworten lassen*).

(*Maximal 2x helfen*)

### 3. Vermutung: „Dinge, die schwer sind, schwimmen nicht.“

Holzklötz [p -q]; Büroklammer [-p q]; Stein [ p q]; Wasserball [-p -q];

Sortieren der Gegenstände nach Eigenschaften:

„Schau mal, hier sind vier Dinge: Stein, Holzklötz, Büroklammer, Wasserball (*Gegenstände zeigen*).

„Sortiere doch mal die schweren Dinge hier hin (*auf linke Seite zeigen*) und die leichten Dinge hierhin (*auf rechte Seite zeigen*).

*Kind sortieren lassen.* (Richtige Lösung: Holzwürfel, Stein schwer, Büroklammer, Wasserball Bimsstein leicht)

Bei Fehler:

„Überleg´ doch mal, ist das Ding schwer oder leicht?“

*(Kind antworten lassen.)*

*(Maximal 2x helfen)* „Das ist Petra. Sie vermutet: ‚Dinge, die schwer sind, schwimmen nicht. ‘“

1. Gegenstand [p -q];

„Dazu nehme ich den Holzklötz hier aus der Kiste.“ (*Bimsstein zeigen*). „Das Besondere an ihm ist, dass er schwer ist. Gut. Jetzt probieren wir aus, ob der Holzklötz schwimmt.“

„Schau genau hin (*Holzwürfel fallen lassen*). Hast du gesehen, was passiert ist? (*Antwort des Kindes abwarten*). „Ja, richtig. Der Holzklötz ist schwer und schwimmt.“

„Jetzt denke nochmal an Petra und ihre Vermutung. Petra vermutet ja: ‚Dinge, die schwer sind, schwimmen nicht.‘ Was zeigt dir der Holzklötz über Petra Vermutung?“

„Wenn du glaubst, Petra Vermutung stimmt, legst du den Holzklötz hierhin.“ (Auf den lachenden Smiley zeigen.)

Wenn du glaubst, Petra Vermutung hat damit nichts zu tun, legst du den Holzklötz hierhin (Auf den nachdenklichen Smiley zeigen.)



Wenn du glaubst, Petra Vermutung stimmt nicht, legst du den Holzklötz auf den traurigen Smiley.  
(Auf den Smiley zeigen.)

Bei Fehler:

Überleg' doch mal, ist das Ding leicht oder schwer?

Kind antworten lassen.

Überleg' doch noch mal, schwimmt das Ding oder nicht?

Kind antworten lassen.

(Maximal 2x helfen).

Petras Vermutung wollen wir jetzt noch mit einem weiteren Ding überprüfen.

Sie vermutet. ‚Dinge, die schwer sind, schwimmen nicht.‘

2. Gegenstand [¬p q];

„Dazu nehme ich die Büroklammer hier aus der Kiste.“ (*Büroklammer zeigen*). „Das Besondere an ihr ist, dass sie leicht ist. Gut. Jetzt probieren wir aus, ob die Büroklammer schwimmt.“

„Schau genau hin.“ (*Büroklammer fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig. Die Büroklammer ist leicht und schwimmt nicht. Jetzt denke nochmal an Petra und ihre Vermutung.“

„Petra vermutet ja: ‚Dinge, die schwer sind, schwimmen nicht.‘ Was zeigt dir die Büroklammer über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du die Büroklammer hierhin.“ (*Auf den lachenden Smiley zeigen*.)

„Wenn du glaubst, Petra Vermutung hat damit nichts zu tun, legst du die Büroklammer hierhin.“ (*Auf den nachdenklichen Smiley zeigen*.)

„Wenn du glaubst, Petra Vermutung stimmt nicht, legst du die Büroklammer hierhin.“ (*Auf den traurigen Smiley zeigen*.)

Bei Fehler:

„Überleg´ doch mal, ist das Ding leicht oder schwer?“

*(Kind antworten lassen).*

„Überleg´ doch noch mal, schwimmt das Ding oder nicht?“

*(Kind antworten lassen).*

*(Maximal 2x helfen).*

„Die Vermutung wollen wir jetzt noch mit einem anderen Ding überprüfen.“

Petra vermutet: ‚Dinge, die schwer sind, schwimmen nicht.‘

3 .Gegenstand [p q]

„Dazu nehme ich den Stein hier aus der Kiste.“ *(Stein zeigen)*. „Das Besondere an ihm ist, dass er schwer ist. Gut. Jetzt probieren wir aus, ob der Stein schwimmt.“ *(Stein fallen lassen)*. „Hast du gesehen, was passiert ist?“ *(Antwort des Kindes abwarten)*. „Ja, richtig, der Stein ist schwer und schwimmt nicht.“

„Jetzt denke nochmal an Petras und ihre Vermutung.“

„Sie vermutet ja: ‚Dinge, die schwer sind, schwimmen nicht. ‘Was zeigt dir der Stein über ihre Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du den Stein hierhin.“ *(Auf den lachenden Smiley zeigen.)*

Wenn du glaubst, Petras Vermutung hat damit nichts zu tun, legst du den Stein hierhin. *(Auf den nachdenklichen Smiley zeigen.)*

Wenn du glaubst, Petras Vermutung stimmt nicht, legst du den Stein auf den Smiley. *(Auf den traurigen Smiley zeigen.)*

Bei Fehler:

„Überleg´ doch mal, ist das Ding leicht oder schwer?“

*(Kind antworten lassen).*

„Überleg´ doch noch mal, schwimmt das Ding oder nicht?“

*(Kind antworten lassen).*

*(Maximal 2x helfen)*

„Jetzt nehmen wir noch ein Ding, um Petras Vermutung ‚Dinge, die schwer sind, schwimmen nicht‘ zu überprüfen“.

Wasserball [-p -q]

„Dazu nehme ich den Wasserball hier aus der Kiste.“ (*Wasserball zeigen*). „Das Besondere an ihm ist, dass er leicht ist. Gut. Jetzt probieren wir aus, ob der Wasserball schwimmt.“ (*Wasserball fallen lassen*). „Hast du gesehen, was passiert ist?“ (*Antwort des Kindes abwarten*). „Ja, richtig, der Wasserball ist leicht und schwimmt nicht.“

„ Jetzt denke nochmal an Petras und ihre Vermutung. Petra vermutet ja: ‚Dinge, die schwer sind, schwimmen nicht. ‘ Was zeigt dir der Wasserball über Petras Vermutung?“

„Wenn du glaubst, Petras Vermutung stimmt, legst du den Wasserball hierhin. (*Auf den lachenden Smiley zeigen.*)

Wenn du glaubst, Petra Vermutung hat damit nichts zu tun, legst du den Wasserball hierhin. (*Auf den nachdenklichen Smiley zeigen.*)

Wenn du glaubst, Petra Vermutung stimmt nicht, legst du den Wasserball hierhin. (*Auf den traurigen Smiley zeigen.*)

Bei Fehler:

„Überleg´ doch mal, ist das Ding leicht oder schwer?“

(*Kind antworten lassen*).

„Überleg´ doch noch mal, schwimmt das Ding oder nicht?“

(*Kind antworten lassen*).

(*maximal 2x helfen*)

# Protokollbogen Argumentationstest

Geschlecht: m w

Geburtsmonat: \_\_\_\_\_

Geburtsjahr: 20

TN:

„Schau das ist Tim. Er denkt: „Dinge, die leicht sind, springen.“ (Spielfigur Tim zeigen), „ Das ist Anna, sie glaubt, Tim hat nicht Recht.“ (Spielfigur Anna zeigen).

I: (Tim) „Schau, ich habe hier eine Styroporkugel. Sie ist leicht und springt.“ (Dabei eine Styroporkugel fallen lassen). Also hat Tim Recht.

I: (Anna) „ Das stimmt nicht. Wenn du jetzt einen Wattebausch fallen lässt, siehst du, dass er nicht springt.“ (Dabei Wattebausch fallen lassen.)

„Denkst du, Anna kann mit dem Wattebausch zeigen, dass Tims Vermutung nicht stimmt?“ (Kind antworten lassen).

Ja nein

AGT 1.1	Anna	Anna hat ein Gegenbeispiel gefunden WEIL: der Wattebausch leicht ist, aber nicht springt. [p $\neg$ q]	<input type="checkbox"/>	<input type="checkbox"/>	(Begründung erfragen)
------------	------	--	--------------------------	--------------------------	-----------------------

„Warum denkst du, dass Annas Beispiel mit dem Wattebausch gut/schlecht ist?“ (Kind antworten lassen).

Antwort:

---



---

„Ich habe dir noch ein paar Gegenstände mitgebracht. (Dinge vor das Kind legen)

- (a) ein Luftballon, der leicht ist und springt,
- (b) eine Tüte, die leicht ist und nicht springt und
- (c) einen Luftballon mit Sand, der schwer ist und nicht springt.

„Ist ein Gegenstand dabei, mit dem du zeigen kannst, dass die Vermutung „Dinge, die leicht sind, springen“ nicht stimmt?

AGT 1.2	ja	nein
	<input type="checkbox"/>	<input type="checkbox"/>

Welcher?

# Interviewleitfaden

## Intervention I mit Modeling

Vermutung „Dinge, die hart sind, springen.“

(lila Knetkugel [ $p \neg q$ ]; Softball [ $\neg p q$ ]; Murmel [ $p q$ ]; Jonglierball)

Explizieren des Gedankenganges komplett

1. Sortieren der Gegenstände

→ nach Eigenschaften

→ nach Sprungverhalten

2. Präsentieren der Vermutung

→ Spielfigur einführen

3. Überprüfen der Vermutung

→ Modeling bis zum Schluss (Explizieren des kompletten Gedankenganges)

# Sortieren der Gegenstände nach Eigenschaften:

**Jens Vermutung: „Dinge, die hart sind, springen.“**

(trockener Knetkugel [p -q]; Softball [-p q]; Murmel [pq]; Jonglierball [-p -q])

**„Schau mal, hier sind vier Dinge: ein Softball, eine Murmel, ein trockener Knetball und ein Jonglierball.“**

(Bälle zeigen)

**„Hier habe ich zwei Tafeln mitgebracht. Eine ist ganz hart (auf Metallplatte zeigen)**

**und die andere ganz weich (auf Platte mit Watte zeigen).**

**„Nun sortier doch mal die vier Dinge nach hart und weich. Die harten Dinge legst du**

**hier hin (auf Metallplatte zeigen) und die weichen Dinge legst du hier hin (auf Platte**

mit Watte zeigen).

Kind sortieren lassen. [richtige Lösung: Murmel: hart; trockener Knetball: hart; Softball: weich; Jonglierball: weich]

Protokoll 1a)

Bei Fehler:

**I: „Fühl doch noch einmal ganz genau. Ist das Ding hart oder weich?“**

Kind anfassen lassen

**I: „Und wo würdest du es dann hinlegen?“**

So lange fragen, bis alle Dinge auf der richtigen Platte liegen

Protokoll 1b)

Protokoll 1c) Bälle wegnehmen

# Sortieren der Gegenstände nach Sprungverhalten

**„Nun wollen wir mal sehen, welche dieser Dinge springen und welche nicht springen können.“**

**„Dafür habe ich hier noch etwas mitgebracht. Diese Karte bedeutet, dass das Ding springen kann. (Karte für springt zeigen) und diese bedeutet, dass das Ding nicht springen kann, also nur fällt und auf liegen bleibt. (Karte für springt nicht zeigen).“**  
(Symbolkarten für springt und springt nicht auf jede Platte legen.)

**„Nun lass die Dinge mal fallen, um zu sehen, ob sie springen können oder nicht.“**  
**Und sortiere die vier Dinge danach, ob sie springen oder nicht.**

**Die Dinge, die hart sind und springen legst du hier hin** (Symbolkarte „springt“ im Feld „hart“), **die Dinge, die hart sind und nicht springen legst du hier hin** (Symbolkarte „springt nicht“ im Feld „hart“),

**die Dinge, die weich sind und springen kommen hier hin** (Symbolkarte „springt“ im Feld „weich“)

**und die Dinge, die weich sind und nicht springen legst du hier hin** (Symbolkarte „springt nicht“ im Feld „weich“).

Kind sortieren lassen. [richtige Lösung: Murmel hart, springt; Knetkugel hart, springt nicht; Softball: weich, springt; Jonglierball: weich, springt nicht]

Protokoll 2a)

Bei Fehler:

**I: „Wir haben ja eben schon gesehen, dass dieses Ding hart / weich ist. Und nun sag mir doch noch mal wo denn die harten / weichen Dinge hinkommen? (falls schon gesagt: Deswegen haben wir es ja auf das harte / weiche Feld gelegt.) und nun probiere doch noch einmal aus, ob das Ding springt oder nicht.“**

Kind ausprobieren lassen (springt)

**Wo würdest du es jetzt hinlegen?“**

So lange fragen, bis alle Dinge auf der richtigen Platte liegen

Bälle liegen lassen



# Zusammenfassung

**„Jetzt wissen wir, welche Dinge hart und welche weich sind.“**

Bälle wegnehmen.

**„Und jetzt sag mir doch nochmal warum du die Bälle auf die jeweiligen Felder gelegt hast. Also warum liegt zum Beispiel die Knetkugel dort?“**

[auf Murmel bzw. Feld zeigen]

Kind antworten lassen

Protokoll 3a)

**„Und warum liegt der Softball dort?“**

Kind antworten lassen

Protokoll 3b)

**„Was ist mit der Murmel?“**

Kind antworten lassen

Protokoll 3c)

**„Und warum liegt der Jonglierball auf diesem Feld?“**

Kind antworten lassen

Protokoll 3d)

**„Jetzt wissen wir, welche harten Dinge springen und welche nicht. Wir wissen welche weichen Dinge springen und welche nicht.“**

Bälle wegnehmen.

# Interviewleitfaden Intervention II mit Modeling

Vermutung „Dinge, die mit  
Luft gefüllt sind, springen.“

(Wasserball mit wenig Luft [p  
-q]; Spielball [p q]; Würfel  
[-p q]; aufgeschnittener  
Luftballon [-p -q];

Gegenstand Explizieren des  
Gedankenganges, 2.+3.  
Gegenstand zeigen der  
Gedanken bis zur  
Entscheidung, 4. Gegenstand  
adaptives Nachfragen

1. Sortieren der Gegenstände  
→ nach Eigenschaften  
Nach Sprungverhalten

2. Präsentieren der Vermutung  
→ Spielfigur einführen

3. Überprüfen der Vermutung  
Modeling /Fading Adaptives/  
Nachfragen  
Bei erstem Gegenstand  
Wasserball kein Protokoll bei 5  
6 und 7

# Sortieren der Gegenstände nach Eigenschaften

**„Schau´ mal, hier sind vier Dinge: ein aufgeschnittener Luftballon, ein Spielball, ein aufblasbarer Wasserball und einen Würfel.“**

(Bälle zeigen)

**„Hier habe ich zwei kleine Luftkissen mitgebracht. Eins ist nicht aufgeblasen und flach, also ohne Luft (auf nicht aufgeblasenes Luftkissen zeigen)**

**und eins aufgeblasen, also mit Luft (auf aufgeblasenes Luftkissen zeigen).**

**„Nun sortiere doch mal die vier Dinge danach, ob sie mit Luft oder ohne Luft sind.**

**Die Dinge ohne Luft legst du hier hin (auf nicht aufgeblasenes Luftkissen zeigen)**

**und die Dinge mit Luft legst du hier hin (auf aufgeblasenes Luftkissen zeigen).**

Kind sortieren lassen. [richtige Lösung: Spielball, aufblasbarer Wasserball: mit Luft; aufgeschnittener Luftballon, Würfel: ohne Luft]

Protokoll 1a)

Bei Fehler:

**„Schau doch noch einmal genau hin. Ist das Ding mit Luft oder ohne Luft?“**

Kind anfassen lassen.

**„Und wo würdest du es dann hinlegen?“**

So lange fragen, bis alle Dinge auf dem richtigen Kissen liegen

Protokoll 1b)

Protokoll 1c)

**„Jetzt wissen wir, welche Dinge mit Luft und welche ohne Luft sind.“**

Bälle wegnehmen

(

# Sortieren der Gegenstände nach Sprungverhalten

**„Nun wollen wir mal sehen, welche dieser Dinge springen und welche nicht springen können.“**

**„Dafür habe ich hier noch etwas mitgebracht.“**

(Symbolkarten für springt und springt nicht auf jedes Luftkissen legen.)

**„Nun lass die Dinge mal fallen um zu sehen, ob sie springen können oder nicht.“**

Kind ausprobieren lassen.

**„Und nun sortiere die vier Dinge danach, ob sie springen oder nicht.“**

**Die Dinge, die mit Luft sind und springen legst du hier hin.** (Symbolkarte „springt“ auf Kissen mit Luft)

**die Dinge, die mit Luft sind und nicht springen legst du hier hin.** (Symbolkarte „springt nicht“ auf Kissen mit Luft)

**die Dinge, die ohne Luft sind und springen kommen hier hin.** (Symbolkarte „springt“ auf nicht- Kissen mit Luft)

**und die Dinge, die ohne Luft sind und nicht springen legst du hier hin.** (Symbolkarte „springt nicht“ auf Kissen mit Luft)

Kind sortieren lassen. [richtige Lösung: Tennisball: mit Luft, springt; Wasserball: mit Luft, springt nicht; aufgeschnittener Luftballon: ohne Luft, springt nicht; Würfel: ohne Luft, springt]

Protokoll 2a)

Bei Fehler:

**„Schau doch noch einmal genau hin. Ist das Ding mit Luft oder ohne Luft?“**

Kind antworten lassen

**und nun probiere doch noch einmal aus, ob das Ding springt oder nicht.**

Kind ausprobieren lassen

**Wo würdest du es jetzt hinlegen?“** So lange fragen, bis alle Dinge auf der richtigen Platte liegen

Protokoll 2b)

Protokoll 2c)

Bälle liegen lassen

# Zusammenfassung

**„Und jetzt sag mir doch nochmal warum du die Bälle auf die jeweiligen Felder gelegt hast. Also warum liegt zum Beispiel der Wasserball dort?“**

[auf Wasserball bzw. Feld zeigen]

Kind antworten lassen

Protokoll 3a)

**„Und warum liegt der Spielball dort?“**

Kind antworten lassen

Protokoll 3b)

**„Was ist mit dem aufgeschnittene Luftballon?“**

Kind antworten lassen

Protokoll 3c)

**„Und warum liegt der Würfel auf diesem Feld?“**

Kind antworten lassen

Protokoll 3d)

**„Jetzt wissen wir, welche Dinge ohne Luft springen und welche nicht. Wir wissen welche Dinge mit Luft springen und welche nicht.“**

## Erklärung: Was ist eine Vermutung?

**Wir wollen heute verschiedene Vermutungen überprüfen.**

**Eine Vermutung ist etwas, das ich zwar glaube, aber noch überprüfen muss.**

(Kästchen hervorholen und zeigen)

**Ich vermute, dass in dem Kästchen ein Boxer versteckt ist. Was glaubst du?** (Kind antworten lassen)

**Wie können wir herausbekommen, ob deine oder meine Vermutung stimmt?** (Kind antworten lassen)

Wir können doch einfach nachschauen, um festzustellen, ob deine oder meine Vermutung stimmt. (Kind nachschauen lassen).

Ja, richtig, da ist eine Tänzerin drin.

Welche Vermutung stimmt, deine oder meine Vermutung? Oder hat keiner von uns beiden Recht? (Kind antworten lassen).

Eine Vermutung kann stimmen, muss aber sie kann auch falsch sein. Das müssen wir immer überprüfen.

# Präsentation der Vermutung

**„Nun haben wir ja schon einige Kinder gefragt, was sie denken, warum ein Ball springt. Mia meint, dass Dinge die mit Luft sind, springen.“**

**„Das hier ist Mia. (auf weibliche Playmobilfigur zeigen)**

**Wo würdest du Mia nun hinstellen? Mia meint ja, dass Dinge die mit Luft sind springen. Die Dinge, von denen sie spricht, müssen also mit Luft sein und springen können.** (Richtige Antwort: „Auf die Symbolkarte „springt“ auf dem Kissen mit Luft“.)

Protokoll 4a)

Bei Fehler:

**„Spricht Mia von Dingen mit Luft oder ohne Luft? Auf welches Kissen würdest du Mia also stellen?“**

Kind antworten lassen (richtige Antwort: auf das Kissen mit Luft)

**„Und, meint Mia, dass diese Dinge springen können oder, dass sie nicht springen können?“**

Kind antworten lassen (richtige Antwort: springen)

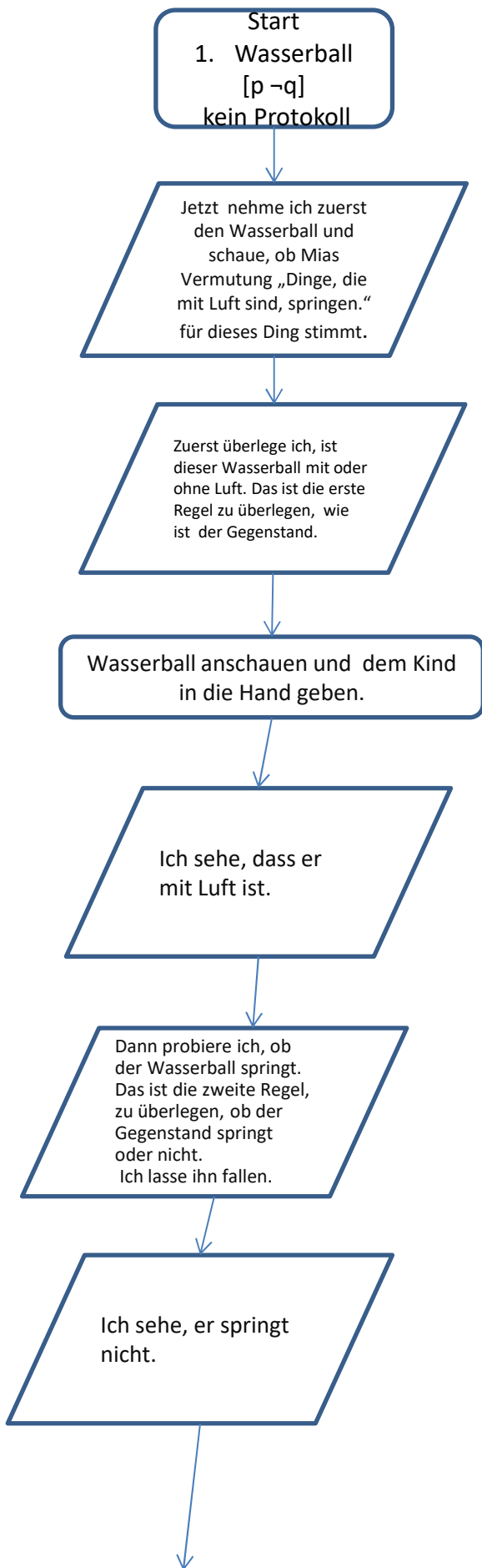
**„Auf welches Feld musst du Mia also stellen?“**

Kind antworten lassen (richtige Antwort: auf das Kissen mit Luft, Symbolkarte „springt“)

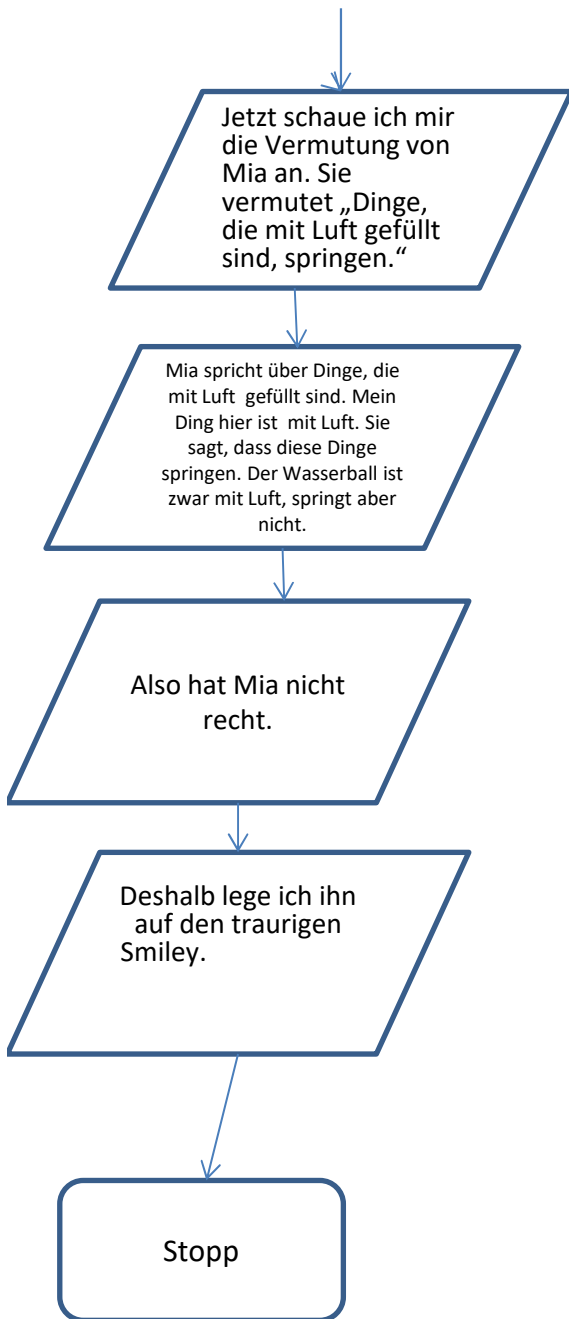
Solange fragen, bis die Spielfigur auf dem richtigen Feld steht.

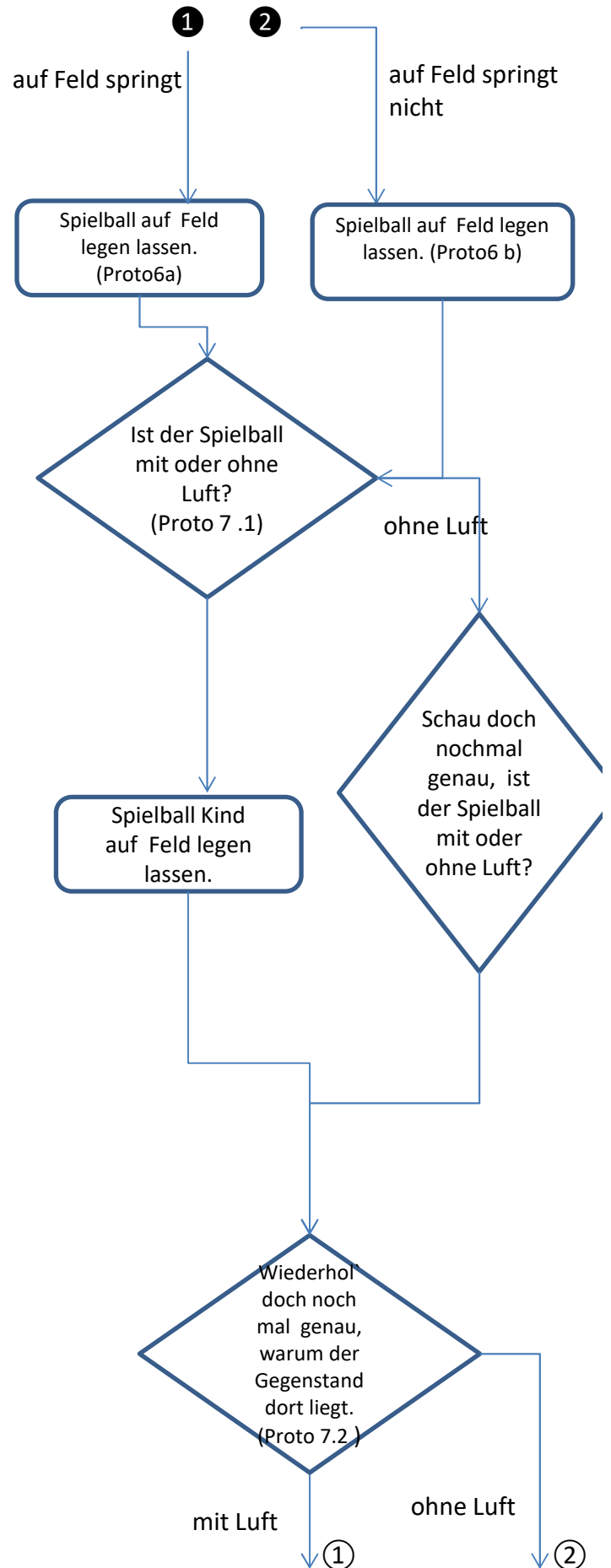
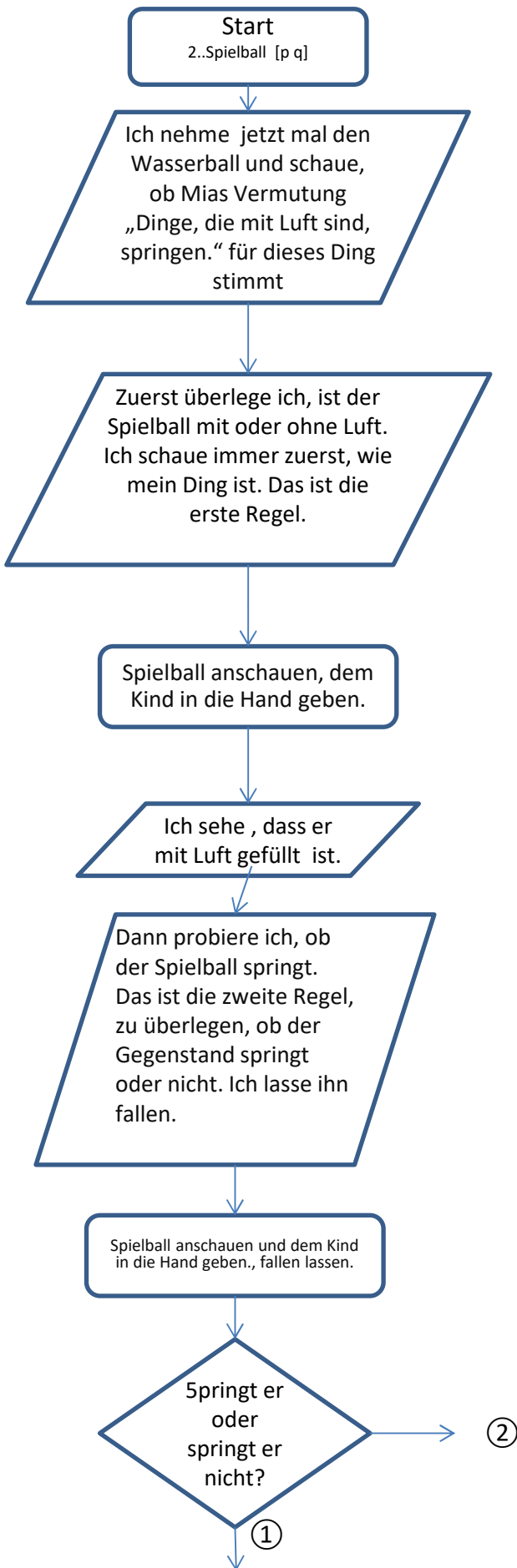
Protokoll 4b)

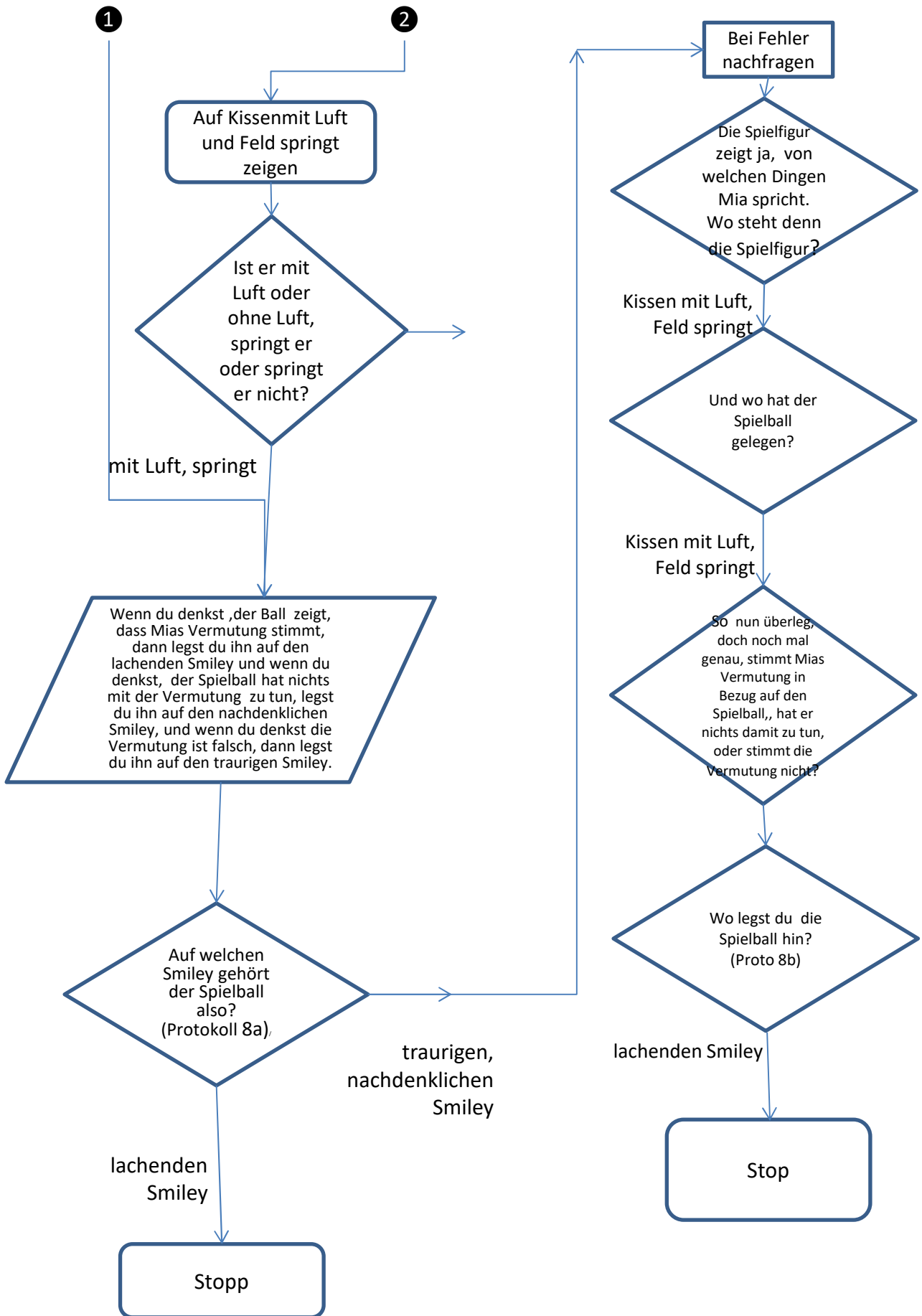
Protokoll 4c)

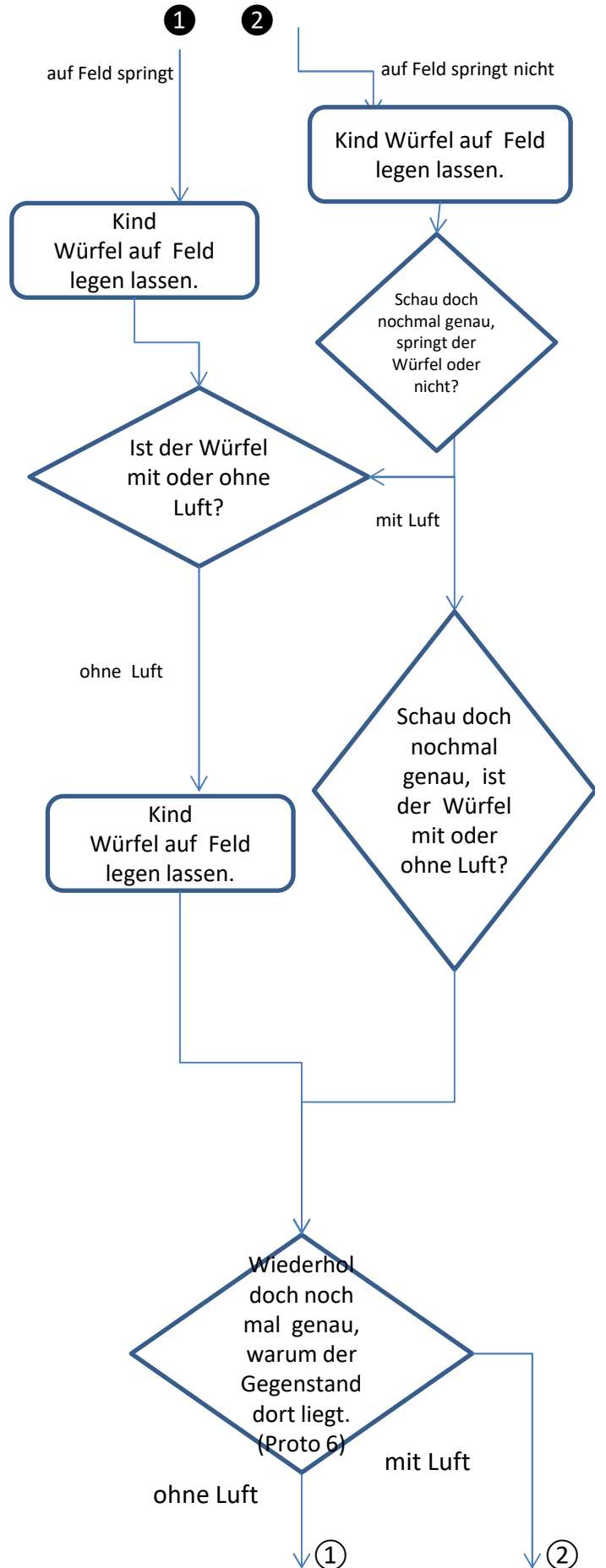
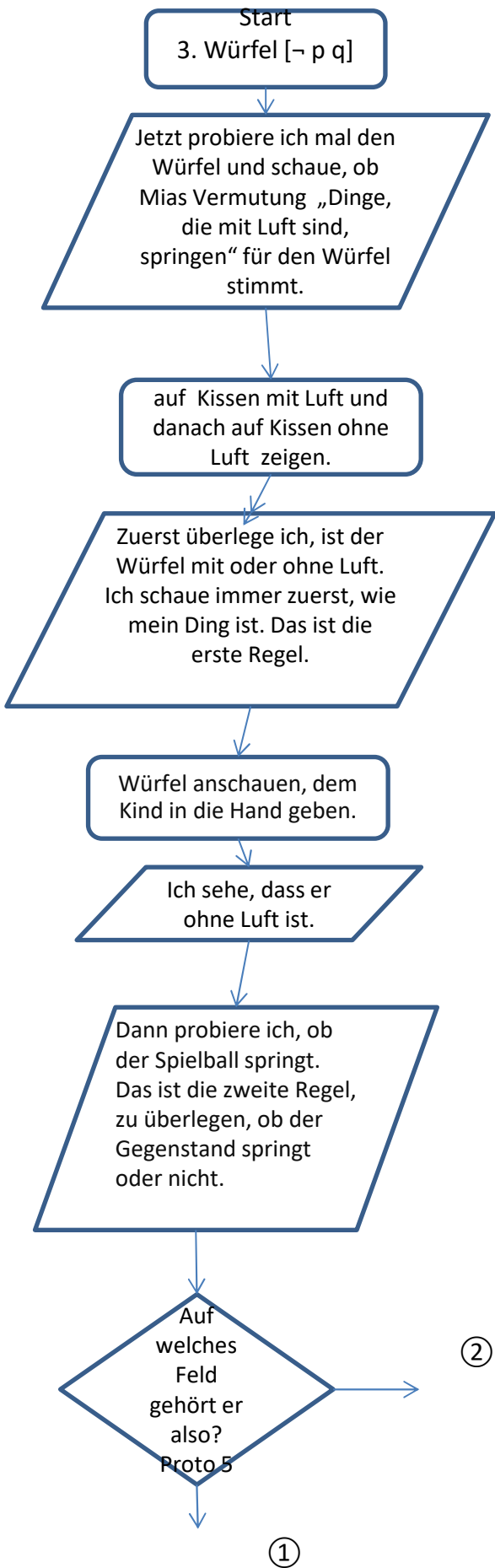


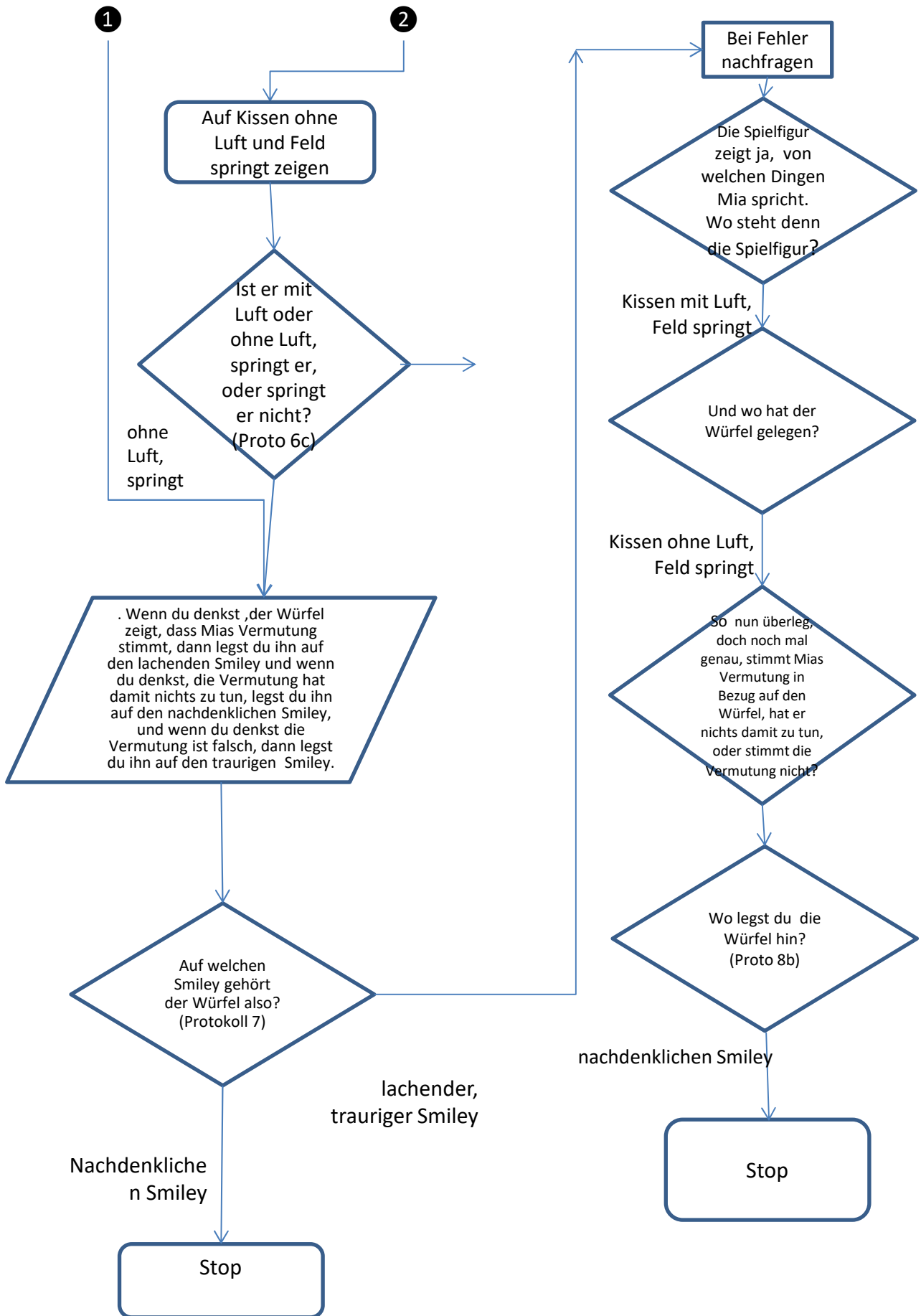


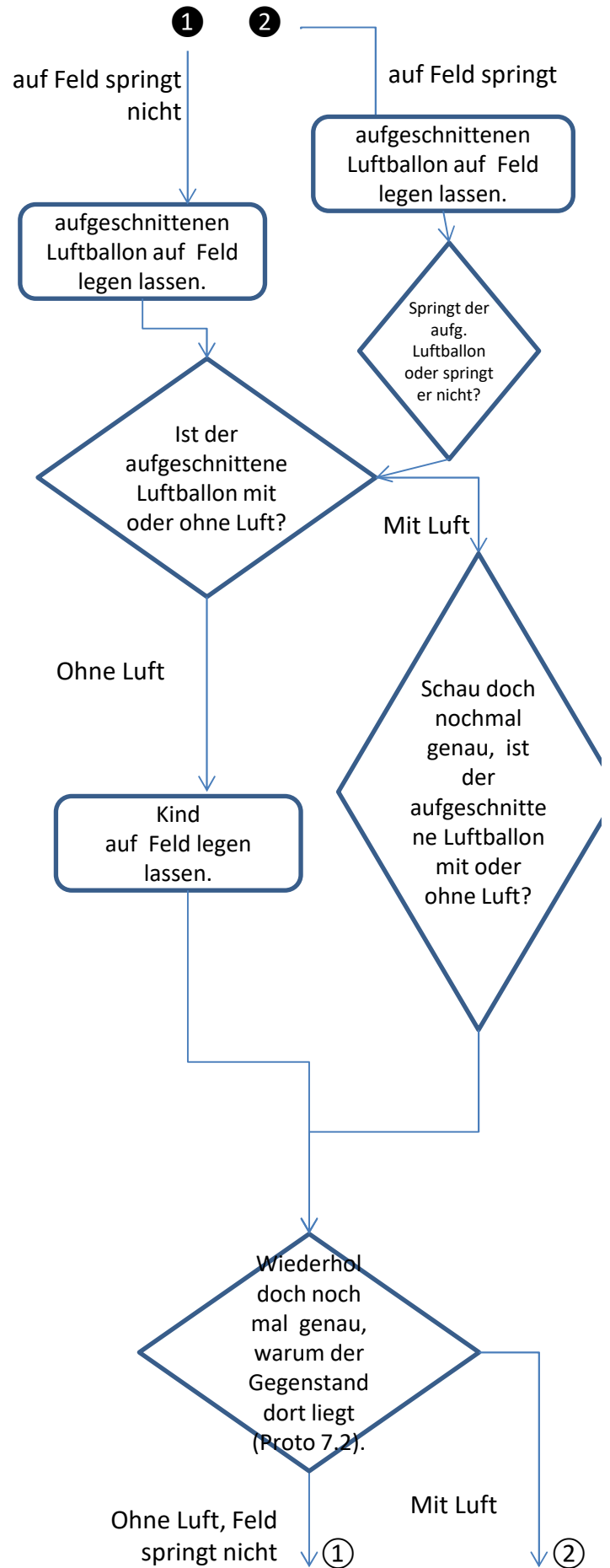
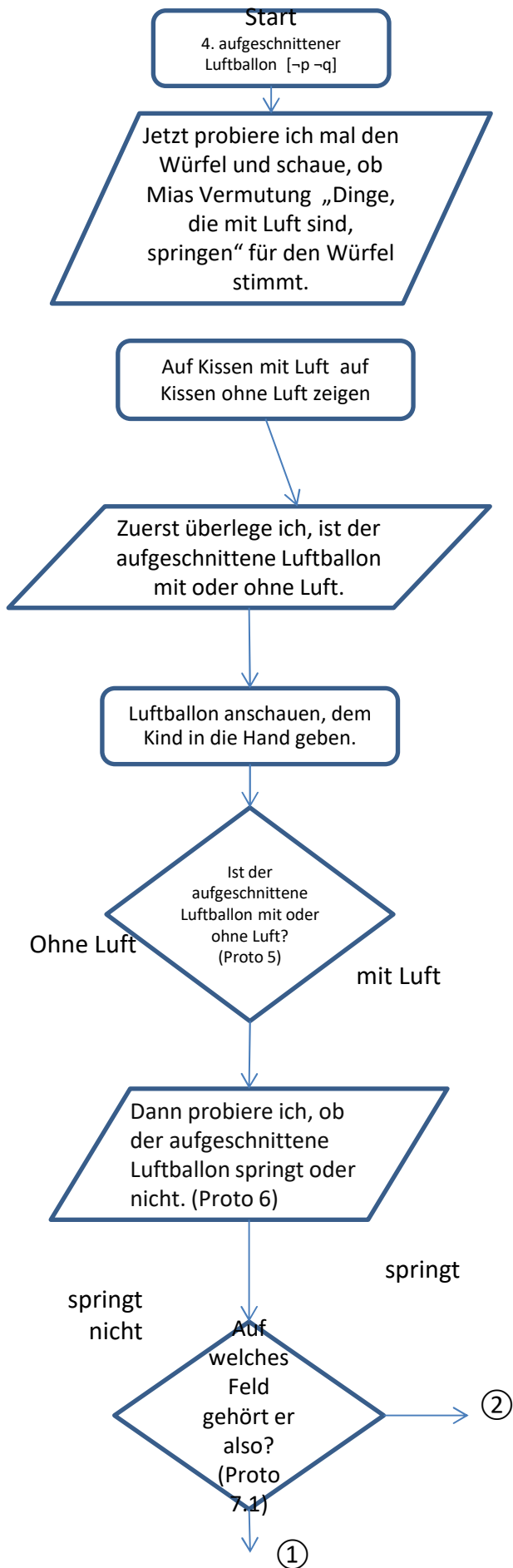


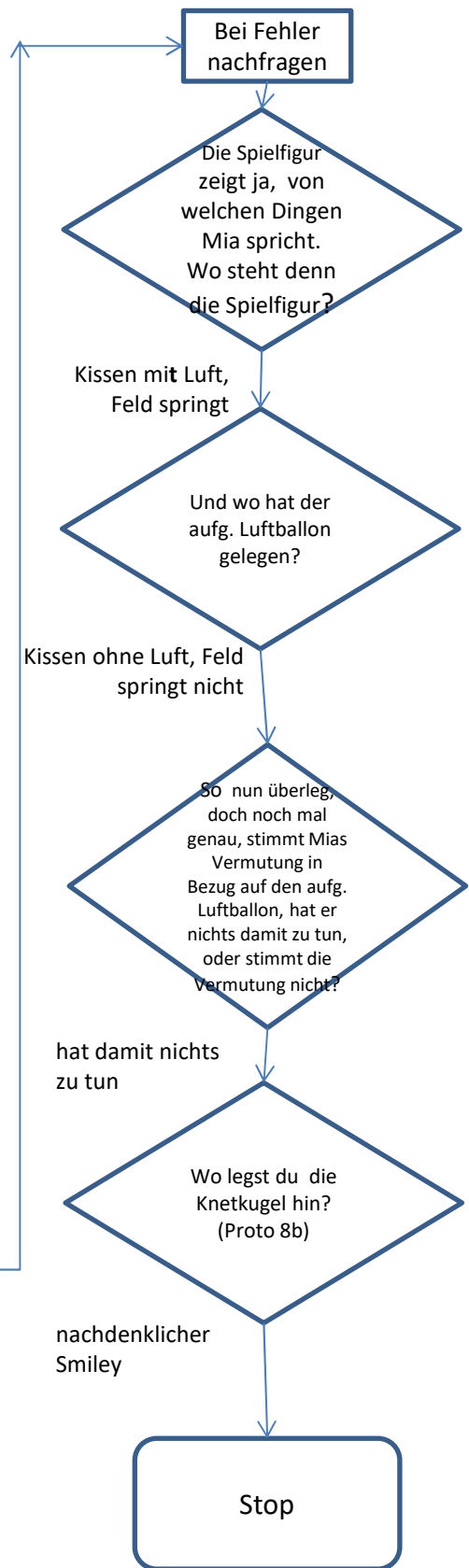
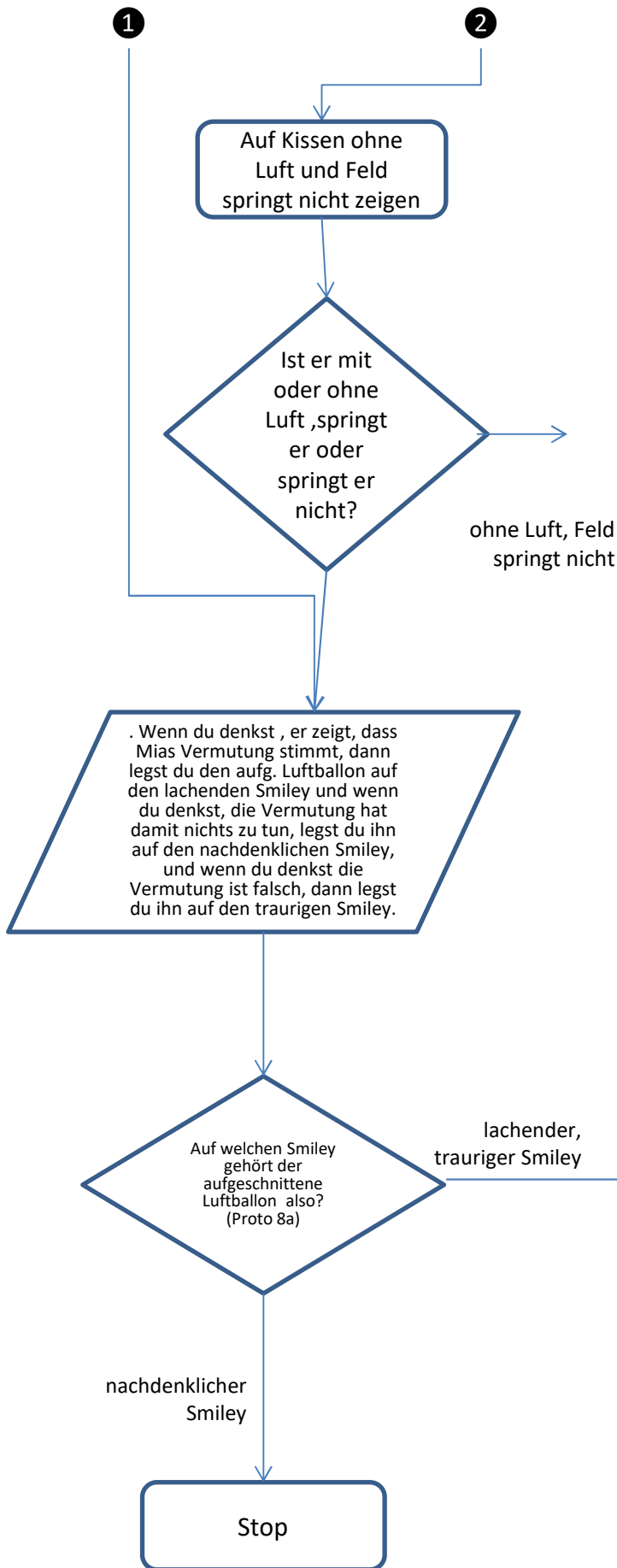












## Legende:

