

Die Förderung deduktiver Schlussfolgerungen bei Grundschulkindern in naturwissenschaftlichen Kontexten

Fostering deductive reasoning in contexts of primary science education

Die Induzierung eines kognitiven Konflikts durch diskrepante Ereignisse ist im Rahmen von Conceptual-Change-Ansätzen eine verbreitete Lehrmethode. Sie setzt voraus, dass Lernende über die erforderlichen deduktiven Denkfähigkeiten verfügen, um in Hinblick auf zu prüfende Vorstellungen bestätigende, widerlegende und irrelevante Ereignisse voneinander zu unterscheiden. Dies stellt eine besondere Herausforderung für Grundschul Kinder dar. Entsprechend zeigten 43 Erst- und Viertklässler bei einer truth-testing task zur Beurteilung von Bedingungen für das Sprungverhalten verschiedener Gegenstände kein konditionales Verständnis einfacher naturwissenschaftlicher Annahmen, sondern überwiegend ein inkonsistentes Antwortverhalten (Experiment 1). Beim Vergleich von zwei Formen instruktionaler Unterstützung zur korrekten Beurteilung von Ereignissen mit 42 Viertklässlern erwies sich ein adaptives Nachfragen bei inkorrekten Beurteilungen als statistisch signifikant wirksamer als ein vorhergehendes Klassifizieren der Ereignisse, welches die kognitive Verarbeitung der Eigenschaften der Ereignisse unterstützen sollte (Experiment 2).

Schlüsselwörter: deduktives Denken, kognitiver Konflikt, Conceptual Change, Naturwissenschaftsdidaktik

Within the framework of conceptual change, evoking cognitive conflict by means of discrepant events is a popular teaching method in science education. It presumes, however, that learners dispose of the necessary deductive reasoning abilities to differentiate events as confirmatory, disproving or irrelevant with respect to extant conceptions. This is a challenge for primary school children. Accordingly, 43 first and fourth graders did not show a conditional understanding of simple scientific assumptions in a truth-testing task, which thematically featured conditions making different objects bounce, but rather showed an inconsistent response behaviour (Experiment 1). Comparing two forms of instructional support for correct judgement of events with 42 fourth graders, an adaptive repeated questioning in case of incorrect judgements proved significantly more effective than a preceding classification of events, which was intended to support cognitive processing of event features (Experiment 2).

Keywords: deductive reasoning, cognitive conflict, conceptual change, science education

1. Einführung

Der gezielte Einsatz von Demonstrations- und Schülerversuchen mit überraschenden Versuchsausgängen – in der englischsprachigen Literatur als *discrepant events* oder als *anomalous data* bezeichnet – ist fester Bestandteil des Methodenrepertoires der Naturwissenschaftsdidaktik. Ihre lerntheoretische Legitimation erfahren solche Methoden aus konstruktivistischen Sichtweisen innerhalb der Naturwissenschaftsdidaktik (Sandoval, 1995; White & Gunstone, 2008), welche die Veränderung bestehender Vorstellungen von Lernenden in den Mittelpunkt stellen und der Erzeugung kognitiver Konflikte eine zentrale Rolle zuschreiben (Posner, Strike, Hewson & Gertzog, 1982; Duit & Treagust, 2003). Die empirische Befundlage zur Wirksamkeit der kognitiven Konfliktstrategie ist allerdings uneinheitlich. Entsprechend sollten mögliche Hemmnisse für konzeptuelle Veränderungen und mögliche instruktionale Unterstützungen zur Überwindung dieser Hemmnisse ein Gebiet weiterer Forschung darstellen (Duit & Treagust, 2003; Limón, 2001).

Obwohl Lernende, zumindest in Hinblick auf eine Reihe von naturwissenschaftlichen Inhalten, über deduktive Denkfähigkeiten verfügen müssen, um von beobachteten Ereignissen auf die Haltbarkeit ihrer Vorstellungen zu schließen, wurde den logischen Denkfähigkeiten als möglicher Einflussgröße für den Ablauf von instruktional induzierten konzeptuellen Umstrukturierungen bisher wenig Beachtung geschenkt (vgl. als Ausnahme Kang, Scharmann & Noh, 2004). Für das naturwissenschaftliche Lernen im Primarbereich ist die Frage nach dem Zusammenhang zwischen logischen Fähigkeiten und konzeptuellen Veränderungen besonders drängend, da kognitionspsychologische Studien nahe legen, dass das Denken von Grundschulkindern spezifische Einschränkungen im Bereich des deduktiven Schlussfolgerns, insbesondere im Umgang mit irrelevanten Informationen und uneindeutigen Schlussfolgerungen, aufweist (Barrouillet, Gauffroy & Lecas, 2008; English, 1997; Markovits, Schleifer & Fortier, 1989). Vor diesem Hintergrund sollte erkundet werden, ob und – wenn ja – welche instruktionalen Maßnahmen deduktive Schlussfolgerungen unterstützen und somit indirekt zur Förderung von konzeptuellen Veränderungen bei Grundschulkindern beitragen können.

2. Die kognitive Konfliktstrategie

Aufbauend auf die Forschung zu Präkonzepten von Lernenden (vgl. Duit, 2007) postulierten bereits Posner et al. (1982), aber auch neuere Ansätze (Duit & Treagust, 2003; Rea-Ramirez & Clement, 1998), als notwendige Bedingung für konzeptuelle Veränderungen die lernerseitige Unzufriedenheit mit eigenen Konzepten, also das Bestehen eines kognitiven Konflikts zwischen vorhandenen Vorstellungen und Phänomenen, welche nicht mit diesen Vorstellungen erklärt werden können. Zwar konnte in zahlreichen Studien eine Wirksamkeit der kognitiven Konfliktstrategie nachgewiesen werden (Gorsky & Finegold, 1994; Guzzetti, Snyder, Glass & Gamas, 1993; Kang et al., 2004), aber ebenso deutlich zeigte sich, dass konzeptuelle Veränderungen keineswegs leicht durch kognitive Konflikte herbeizuführen sind (Chinn & Brewer, 1993). Die Rolle des kognitiven Konflikts bei konzeptuellen Veränderungen wurde daher zunehmend hinterfragt, relativiert und kritisch gesehen (diSessa, 2006; Stavy, 1991; Duit, Treagust & Widodo, 2008).

Für den Erfolg der Konfliktstrategie ist unabdingbar, dass Lernende auch tatsächlich einen bedeutungsvollen kognitiven Konflikt erleben (Chan, Burtis & Bereiter, 1997; Dreyfus, Jungwirth & Elioitch, 1990; Hashweh, 1986). Diese Forderung ist keineswegs trivial. Argumente und Ereignisse, welche Lehrpersonen als geeignet zur Auslösung kognitiver Konflikte erscheinen mögen, müssen bei Lernenden nicht zwingend auf Bedingungen stoßen, welche kognitive Konflikte entstehen lassen. Wir gehen davon aus, dass die Fähigkeit zum deduktiven Schließen bei der Interpretation diskrepanter Ereignisse ein zentraler Faktor für die Auslösung eines kognitiven Konflikts und die nachfolgenden konzeptuellen Veränderungen ist. Allerdings wurden Zusammenhänge zwischen deduktiven Fähigkeiten, kognitiven Konflikten und konzeptuellen Veränderungen bisher nicht systematisch untersucht (siehe Kang et al., 2004; Lawson & Thompson, 1988, für korrelative Zusammenhänge).

3. Die Entwicklung des deduktiven Schließens über hypothetische Annahmen

Seit der berühmten *Wason Selection Task* (Wason, 1966) bildeten Aufgaben, in denen von geltenden Zuständen auf die Gültigkeit von zu überprüfenden Aussagen geschlossen werden muss, ein zentrales Feld der psychologischen Forschung zum deduktiven Denken. Hierbei wurden insbesondere sog. konditionale Aussagen (z.B. „Wenn es regnet, wird die Straße nass.“) in den Blick genommen, um zu eruieren, inwieweit Personen das Mittel der Falsifikation nutzen, um ihre Erkenntnisse einer kritischen Prüfung zu unterziehen. Zahlreiche Untersuchungen zeigten, dass bei Erwachsenen eine Falsifikationsstrategie, wenn überhaupt, nur bei thematischen Einkleidungen angewendet wurde (Wason & Shapiro, 1971).

Vor dem Hintergrund der *Mental Models Theory* (Johnson-Laird & Byrne, 2002) – welche die Fähigkeit zu fehlerfreien deduktiven Schlüssen mit der Kapazität des Arbeitsgedächtnisses erklärt – legten Barrouillet et al. (2008) ein Entwicklungsmodell des deduktiven Schlussfolgerns vor. In einer sog. *truth-testing task* boten sie Dritt-, Sechst- und Neuntklässlern sowie Erwachsenen zunächst Sätze der Form „Wenn der Kreis rot ist, dann ist der Stern schwarz.“ ($p \rightarrow q$) als zu überprüfende abstrakte Aussagen dar. Auf einer Leinwand wurden dann Kreise und Sterne präsentiert, deren Farbgebung so manipuliert war, dass sie Realisierungen der vier möglichen Zustände $p \rightarrow q$, $p \rightarrow \neg q$, $\neg p \rightarrow q$ und $\neg p \rightarrow \neg q$ darstellten. Für die oben genannte Aussage wurde der Zustand $\neg p \rightarrow \neg q$ beispielsweise durch die Präsentation eines blauen Kreises und eines grünen Sterns verwirklicht. Die Versuchsteilnehmer mussten dann angeben, ob der beobachtete Zustand die abstrakte Aussage bestätigte, widerlegte oder keine eindeutige Schlussfolgerung zuließ, also irrelevant war.

Tab. 1: Entwicklungsniveaus des Verständnisses abstrakter Aussagen und erwartete Einschätzungen der zugehörigen Zustände (nach Barrouillet et al., 2008)

Zustand	Verständnisniveau		
	Konjunktiv	Bikonditional	Konditional
$p \ q$	+	+	+
$p \ \neg q$	-	-	-
$\neg p \ q$	-	-	o
$\neg p \ \neg q$	-	o	o

Anm.: + = bestätigend; o = irrelevant; - = widerlegend.

Im Kern wurden drei Entwicklungsniveaus des Verständnisses deduktiver Schlussfolgerungen identifiziert, welche durch je spezifische Antwortmuster gekennzeichnet waren (vgl. Tab. 1). Bei Drittklässlern konnte ein *konjunktives* Verständnis zu prüfender abstrakter Aussagen nachgewiesen werden, welches die Verbindung zwischen Antezedenz und Konsequenz als einfaches „und“ begreift. Im Gegensatz hierzu zeigten Sechstklässler vornehmlich ein *bikonditionales* Verständnis, welches die in der abstrakten Aussage genannte Antezedenz als einzigen und ausschließlichen Grund für die Konsequenz versteht. Bei Neuntklässlern trat erstmals ein *konditionales* Verständnis der untersuchten abstrakten Aussagen auf, welches sich bei Erwachsenen als die bestimmende Form des Verständnisses durchsetzte.

In Anknüpfung an die Studie von Barrouillet et al. (2008) untersuchten wir, welches Verständnisniveau Kinder im Grundschulalter bei deduktiven Schlussfolgerungen mit kontextualisierten naturwissenschaftlichen Inhalten zeigen (Experiment 1) und ob es möglich ist, durch instruktionale Unterstützung das deduktive Schließen von Grundschulkindern zu verbessern (Experiment 2).

4. Experiment 1

Zwar gibt es viele Studien zum deduktiven Denken in thematischen Inhaltsgebieten, eine Betrachtung von explizit für die naturwissenschaftliche Bildung relevanten Inhalten fand bisher jedoch bei Grundschulkindern nicht statt. Mit dem ersten Experiment sollte daher das Niveau des deduktiven Schließens von Grundschulkindern bei naturwissenschaftlichen Inhalten erfasst werden. Zu diesem Zweck wurden sowohl Erst- als auch Viertklässler befragt. Aufgrund der kontextualisierten Inhalte erwarteten wir, dass die von uns untersuchten Viertklässler ein höheres Niveau deduktiven Schließens aufweisen würden als die von Barrouillet et al. (2008) mit abstrakten Inhalten untersuchten Drittklässler.

Als Inhaltsgebiet wurde das Thema der Elastizität und Plastizität von Materialien gewählt, welches sich exemplarisch unter der Leitfrage „Wie kommt es, dass ein Ball springt?“ zusammenfassen lässt (siehe Thiel, 1973). Zum einen entwickeln bereits junge Kinder im genannten Inhaltsgebiet aufgrund eigener Erfahrungen eine Reihe von Präkonzepten, welche nicht mit einer wissenschaftlichen Sichtweise übereinstimmen, während gleichzeitig die Gesamtzahl relevanter Präkonzepte überschaubar bleibt. Zum anderen können zu den Präkonzepten Materialien bereitgestellt werden, welche als bestätigend, widerlegend oder irrelevant interpretiert werden müssen (vgl. Anhang).

4.1 Stichprobe

Nach direkter Ansprache der verantwortlichen Lehrpersonen nahmen 20 Erstklässler (14 Mädchen und 6 Jungen) und 23 Viertklässler (10 Mädchen und 13 Jungen) aus zwei intakten Klassen zweier verschiedener Grundschulen Nordrhein-Westfalens freiwillig an der Untersuchung teil. Die Viertklässler wurden zur Durchführung der Untersuchung in die Universität eingeladen. Die Erstklässler waren durchschnittlich 7.48 Jahre alt, die Viertklässler durchschnittlich 10.24 Jahre.

4.2 Material und Versuchsablauf

Die Fähigkeit zu korrekten Schlussfolgerungen wurde mittels einer *truth-testing task* erfasst. Hierbei mussten die Kinder die Bedeutung verschiedener Ereignisse für die Gültigkeit vorgegebener Annahmen einschätzen. Jedem Kind wurden dieselben zwei Annahmen vorgelegt. Die beiden Annahmen waren, in Anlehnung an den Wortlaut typischer Kinderäußerungen, als einfache Aussagesätze formuliert (z.B. „Weiche Bälle springen.“). In der Terminologie der deduktiven Logik benannten die Annahmen einerseits als Antezedenz verschiedene Balleigenschaften und als Konsequenz stets die Sprungfähigkeit. Aus der systematischen Kombination von Antezedenz und Konsequenz sowie den zugehörigen Negationen ergaben sich dann zu jeder Annahme vier mögliche Ereignisse (vgl. Anhang). Die Kinder mussten für jedes dieser Ereignisse beurteilen, ob es die vorgegebene Annahme bestätigte, widerlegte oder aber irrelevant für die Gültigkeit der Annahme war.

Drei Versuchsleiterinnen und ein Versuchsleiter führten mithilfe eines standardisierten Leitfadens Einzelinterviews durch, wobei eine vergleichbare Durchführung in einer gemeinsamen Versuchsleiterschulung durch Probeinterviews eingeübt worden war. An kritischen Stellen der Interviews waren Formulierungen wörtlich dem Leitfaden zu entnehmen. Die Interviews, welche zwischen 25 und 45 Minuten dauerten, wurden audio-digital aufgezeichnet. Sie begannen mit einer kurzen Demonstration, in welcher ein Ball und ein Knetklumpen simultan auf den Boden fallen gelassen wurden. Anhand dieser Demonstration wurde gemeinsam mit dem Kind die übergeordnete Fragestellung für die folgenden Aufgaben erschlossen („Wie kommt es, dass ein Ball bzw. ein Gegenstand springt?“). Sodann wurde den Kindern der Ablauf der nachfolgenden *truth-testing task* anhand einer Beispiellannahme („Große Bälle springen.“) mit zwei zugehörigen Ereignissen erläutert (vgl. Anhang). Die Bearbeitung der kritischen Aufgaben begann mit der Einführung der ersten zu testenden Annahme („Weiche Bälle springen.“). Die Annahme war auf eine Karte gedruckt und wurde als Idee eines anderen Kindes vorgestellt, welche über-

prüft werden müsse. Darauf folgte die Präsentation des ersten Ereignisses. Hierzu wurde das Sprungverhalten des betroffenen Gegenstandes demonstriert und durch den Versuchsleiter verbalisiert („Die Murmel ist hart und springt.“). Die Kinder mussten dann das Ereignis durch die Auswahl einer der hierfür vorgelegten Antwortkarten als bestätigend („Dieser Ball zeigt: Die Vermutung stimmt.“), widerlegend („Dieser Ball zeigt: Die Vermutung stimmt nicht.“) oder irrelevant („Dieser Ball hat damit nichts zu tun.“) in Hinblick auf die Gültigkeit der vorgegebenen Annahme einschätzen. Dieses Vorgehen wurde für die übrigen drei zur Annahme gehörenden Ereignisse wiederholt, um dann mit der Einführung der zweiten Annahme wieder von vorne zu beginnen. Alle Kinder bearbeiteten die zwei Annahmen und zugehörigen acht Ereignisse (vgl. Anhang) in derselben pseudo-randomisierten Reihenfolge.

4.3 Ergebnisse und Diskussion

Die Einschätzungen der Ereignisse (Tab. 2) wurden mittels einer MANOVA mit dem Innersubjektfaktor Ereignis ($p q, p \neg q, \neg p q, \neg p \neg q$) und dem quasi-experimentellen Zwischensubjektfaktor Alter (Erst- vs. Viertklässler) analysiert. Die drei abhängigen Variablen dieser Auswertung waren die jeweils über die zwei zu prüfenden Annahmen hinweg zusammengefassten Häufigkeiten an bestätigenden, irrelevanten und widerlegenden Einschätzungen (mit möglichen Werten von 0, 1 oder 2). In dieser multivariaten Analyse zeigte sich neben dem Haupteffekt des Faktors Ereignis, Wilks' $\Lambda = .041, F(7, 35) = 116.35, p < .001$, auch die Interaktion der Faktoren Ereignis und Alter, Wilks' $\Lambda = .677, F(7, 35) = 2.39, p = 0.042$, als signifikant.

Tab. 2: Mittelwerte (und Standardabweichungen) der Nennungshäufigkeiten der Antwortkategorien in Experiment 1 getrennt nach Alter

Ereignis	Klasse 1			Klasse 4		
	Bestätigend	Irrelevant	Widerlegend	Bestätigend	Irrelevant	Widerlegend
$p q$	1.90 (0.31)	0.05 (0.22)	0.05 (0.22)	1.91 (0.42)	0.09 (0.42)	0.00 (0.00)
$p \neg q$	0.10 (0.31)	0.45 (0.69)	1.40 (0.75)	0.00 (0.00)	0.13 (0.34)	1.80 (0.34)
$\neg p q$	0.55 (0.89)	0.45 (0.69)	1.00 (0.86)	0.04 (0.21)	0.57 (0.79)	1.39 (0.78)
$\neg p \neg q$	0.70 (0.87)	0.65 (0.75)	0.65 (0.81)	1.09 (0.90)	0.70 (0.88)	0.22 (0.52)

Nachfolgend wurden univariate Analysen der drei abhängigen Variablen vorgenommen. In der Analyse der bestätigenden Einschätzungen erwiesen sich der Haupteffekt des Faktors Ereignis, $F(3, 123) = 86.21, p < .001$, und der Interaktionseffekt der Faktoren Ereignis und Alter, $F(3, 123) = 4.27, p = .007$, als signifikant. Für die irrelevanten Einschätzungen konnte lediglich ein signifikanter Haupteffekt des Faktors Ereignis, $F(3, 123) = 9.03, p < .001$, nachgewiesen werden. Für die widerlegenden Einschätzungen schließlich wurden ein signifikanter Haupteffekt des Faktors Ereignis, $F(3, 123) = 67.20, p < .001$, und ein signifikanter Interaktionseffekt der Faktoren Ereignis und Alter, $F(3, 123) = 5.59, p = .001$, gefunden.

Die signifikanten Haupteffekte des Faktors Ereignis belegen, dass die einzelnen Ereignisse unterschiedlich eingeschätzt wurden. Ereignisse der Form $p q$

wurden fast ausschließlich als bestätigend gewertet, während Ereignisse der Formen $p \neg q$ und $\neg p q$ überwiegend als widerlegend betrachtet wurden. Im Falle von Ereignissen der Form $\neg p \neg q$ ist eine generelle Präferenz für eine der Antwortkategorien über beide Altersgruppen hinweg nicht klar erkennbar. Bei der Vergabe von irrelevanten Einschätzungen gab es keine Unterschiede zwischen den Altersgruppen, wie die Abwesenheit eines entsprechenden Interaktionseffektes belegt. Die signifikanten Interaktionseffekte für die beiden anderen Antwortkategorien spiegeln wider, dass die Neigung zu widerlegenden Einschätzungen für Ereignisse der Formen $p \neg q$ und $\neg p q$ bei den Viertklässlern besonders ausgeprägt war und dass die Viertklässler im Vergleich zu den Erstklässlern Ereignisse der Form $\neg p \neg q$ eher als bestätigend interpretierten. Beide Altersgruppen verfehlten das Niveau eines konditionalen Verständnisses der präsentierten Annahmen.

Neben den drei im Entwicklungsmodell von Barrouillet et. al. (2008) beschriebenen Verständnisniveaus mit zugehörigen Antwortmustern (vgl. Tab. 1), konnte eine weitere Form des Verständnisses gefunden werden. Das zugehörige Antwortmuster entsprach demjenigen des bikonditionalen Verständnisses, außer dass das Ereignis $\neg p \neg q$ als bestätigend eingestuft wurde. Offensichtlich wurden hier die Annahmen bikonditional, aber als Regel verstanden. Die Annahmen wurden also als geltende Regeln gesehen, und es wurde beurteilt, ob die Ereignisse diesen Regeln gehorchten, anstatt die Gültigkeit der Annahmen mithilfe der Ereignisse zu überprüfen. In diesem Sinne handelte es sich um ein *regelmäßiges* Verständnis. Nur 4 der Erstklässler (1mal konjunktiv, 2mal regelmäßig und 1mal konditional) und 11 der Viertklässler (8mal regelmäßig und 3mal konditional) zeigten über die Ereignisse beider Annahmen hinweg dasselbe Antwortmuster. Das deduktive Schließen der hier befragten Viertklässler setzte sich also vom Denken der von Barrouillet et al. (2008) untersuchten Drittklässler, welche abstrakte Annahmen in erster Linie konjunktiv interpretierten, durch das Auftreten des regelmäßigen Verständnisses ab. Insgesamt zeichnete sich das Antwortverhalten allerdings durch Instabilität aus.

5. Experiment 2

Im zweiten Experiment sollte erkundet werden, ob es möglich ist, das im ersten Experiment bestätigte niedrige Niveau des deduktiven Schließens von Grundschulkindern durch Instruktionsmaßnahmen zu verbessern. Hierzu wurden Viertklässler als Zielgruppe gewählt, da hier eine Förderung des Verständnisses auf ein konditionales Niveau möglich schien. Um in einer *truth-testing task* zu konditional korrekten Schlüssen zu gelangen, müssen die Ausprägungen des präsentierten Ereignisses ($p/\neg p$ bzw. $q/\neg q$) zunächst wahrgenommen und dann mit mentalen Modellen abgeglichen werden, welche aufgrund der zu prüfenden Annahme konstruiert wurden. Um den Kindern diesen Verarbeitungsprozess zu erleichtern, sollte die simultane Beachtung von Merkmalsausprägungen ($p/\neg p$) und Sprungverhalten ($q/\neg q$) durch instruktionale Maßnahmen unterstützt werden. In der ersten Bedingung (Sortieren) wurde vor dem Hintergrund einer Entlastung des Arbeitsgedächtnisses eine visuelle Darstellung in Form eines Setzkastens mit passend beschrifteten Zeilen und Spalten (z.B. weich/hart und springt/springt nicht) genutzt, welche während der gesamten Aufgabenbearbeitung präsent war. Die Kinder mussten zunächst die vier zu

einer Annahme gehörenden Ereignisse in den Setzkasten einordnen. Durch Kontrollfragen nach Abschluss der Sortierung wurde sichergestellt, dass die Merkmale der einzelnen Ereignisse fehlerfrei verstanden worden waren (vgl. Ziegler, 1990). Danach wurden die Ereignisse in Hinblick auf die jeweilige Annahme beurteilt. In der zweiten Bedingung (Nachfragen) wurden unter Berücksichtigung von Ergebnissen der Instruktionsforschung zu prozessorientierten kognitiven Strukturierungen (z.B. Einsiedler & Hardy, 2010; Krammer, 2009) Fragen verwendet, welche die Kinder dazu anregen sollten, die Merkmalsausprägungen ($p/\neg p$) und das Sprungverhalten ($q/\neg q$) in den betrachteten Annahmen und Ereignissen angemessen zu verarbeiten. Durch gezielte Nachfragen wurden die Kinder aufgefordert zu benennen, welches besondere Merkmal ($p/\neg p$) und welches Sprungverhalten ($q/\neg q$) die zu prüfende Annahme beinhaltet, und welches besondere Merkmal ($p/\neg p$) und welches Sprungverhalten ($q/\neg q$) am vorliegenden Ereignis zu beobachten waren (z.B. „Springt dieser Ball aus Knete?“). Diese Bedingung wurde als adaptive Hilfe konzipiert, welche im Falle von Fehleinschätzungen der Kinder zum Tragen kam. Konkret wurden die Nachfragen dann mit der Aufforderung, sich noch einmal gemeinsam dem betroffenen Ereignis zuzuwenden, eingeleitet („Lass uns über diesen Ball noch einmal nachdenken.“). Dabei wurde darauf geachtet, den Kindern keine Rückmeldung bezüglich der Richtigkeit ihrer vorigen Antwort zu geben. Anschließend war das fragliche Ereignis erneut als bestätigend, widerlegend oder irrelevant einzuschätzen und diese Einschätzung mit einer freien Antwort zu begründen. Während die Bedingung Sortieren also instruktional als relativ leicht zu realisierende, anschauliche Hilfe bei der kognitiven Verarbeitung von Ereignissen angesehen werden kann, nutzte die Bedingung Nachfragen individuell angepasste Prompts, welche vor dem Hintergrund von Erkenntnissen über die Bedeutung von kognitiv anregenden Fragen vielversprechend erscheinen.

5.1 Stichprobe

Es wurden 42 Kinder (24 Mädchen und 18 Jungen) zweier Klassen der vierten Jahrgangsstufe einer vorstädtischen Grundschule des Ruhrgebiets befragt. Auf Anfrage bei den verantwortlichen Lehrpersonen nahmen die Klassen freiwillig an der Untersuchung teil. Das Durchschnittsalter der Kinder betrug 10.45 Jahre. Nach einer Parallelisierung nach Geschlecht und Lehrereinschätzung der kognitiven Fähigkeit wurden die Kinder zufällig einer der beiden Versuchsbedingungen zugewiesen.

5.2 Material und Versuchsablauf

Die Kinder wurden in Einzelinterviews, von denen Audioaufnahmen gemacht wurden, in einem separaten Raum der Schule befragt. Drei Versuchsleiterinnen, welche zuvor eine gemeinsame Schulung erhalten hatten (vgl. Experiment 1), führten parallel Interviews von 20 bis 40 Minuten Dauer durch. Die *truth-testing task* des ersten Experiments wurde jeweils um die spezifische instruktionale Maßnahme ergänzt (s. oben). Darüber hinaus wurden in beiden Bedingungen die zu prüfenden Annahmen unter Verwendung der Worte *wenn* und *dann* (z.B. „Wenn ein Ball weich ist, dann springt er.“) umformuliert, um das simultane Vorhandensein von Antezedenz und Konsequenz sprachlich salienter zu machen. In der Bedingung des Sortierens wurde die Beispielaufgabe mit

vier Ereignissen durchgespielt, während in der Bedingung Nachfragen nur zwei Ereignisse als Beispiele bearbeitet wurden. Allen Kindern wurden im Anschluss dieselben vier Annahmen mit zugehörigen Ereignissen zur Bearbeitung vorgelegt (vgl. Anhang). Ansonsten verliefen die Interviews wie im ersten Experiment.

5.3 Ergebnisse und Diskussion

Die Einschätzungen der Ereignisse (Tab. 3) wurden einer MANOVA mit dem Innersubjektfaktor Ereignis ($p q, p \neg q, \neg p q, \neg p \neg q$) und dem Zwischensubjektfaktor Instruktion (Sortieren vs. Nachfragen) unterzogen. Als die drei abhängigen Variablen dieser Auswertung fungierten die jeweils über die vier zu prüfenden Annahmen hinweg zusammengefassten Häufigkeiten an bestätigenden, irrelevanten und widerlegenden Einschätzungen (mit möglichen Werten von 0 bis 4). In dieser multivariaten Analyse erwiesen sich sowohl die Haupteffekte der Faktoren Ereignis, Wilks' $\Lambda = .009, F(6, 35) = 642.76, p < .001$, und Instruktion, Wilks' $\Lambda = .670, F(2, 39) = 9.61, p < .001$, als auch die Interaktion dieser beiden Faktoren, Wilks' $\Lambda = .509, F(6, 35) = 5.64, p < .001$, als signifikant.

Tab. 3: Mittelwerte (und Standardabweichungen) der Nennungshäufigkeiten der Antwortkategorien in Experiment 2 getrennt nach Interventionsbedingung

Ereignis	Sortieren			Nachfragen		
	Bestätigend	Irrelevant	Widerlegend	Bestätigend	Irrelevant	Widerlegend
$p q$	3.81 (0.51)	0.00 (0.00)	0.19 (0.51)	4.00 (0.00)	0.00 (0.00)	0.00 (0.00)
$p \neg q$	0.10 (0.30)	0.76 (1.22)	3.14 (1.24)	0.00 (0.00)	0.19 (0.68)	3.81 (0.68)
$\neg p q$	0.05 (0.22)	1.05 (1.40)	2.90 (1.48)	0.00 (0.00)	2.86 (1.56)	1.14 (1.56)
$\neg p \neg q$	0.90 (1.55)	1.43 (1.72)	1.67 (1.62)	0.48 (1.08)	3.24 (1.38)	0.29 (0.90)

Zur weiteren Exploration dieser Effekte wurden univariate Analysen getrennt für die drei abhängigen Variablen berechnet. Für die bestätigenden Einschätzungen wurde ein signifikanter Haupteffekt des Faktors Ereignis, $F(3, 120) = 282.74, p < .001$, gefunden, ebenso für die irrelevanten Einschätzungen, $F(3, 120) = 41.07, p < .001$. Darüber hinaus ergab die Analyse der irrelevanten Einschätzungen einen signifikanten Haupteffekt des Faktors Instruktion, $F(1, 40) = 14.44, p < .001$, sowie einen signifikanten Interaktionseffekt der Faktoren Ereignis und Instruktion, $F(3, 129) = 12.26, p < .001$. In der Analyse der widerlegenden Einschätzungen erwiesen sich der Haupteffekt des Faktors Ereignis, $F(3, 120) = 63.28, p < .001$, der Haupteffekt des Faktors Instruktion, $F(1, 40) = 19.70, p < .001$, sowie der Interaktionseffekt der Faktoren Ereignis und Instruktion, $F(3, 120) = 9.26, p < .001$, als signifikant.

Die signifikanten Haupteffekte des Faktors Ereignis zeigen, dass sich die typischen Einschätzungen der einzelnen Ereignisse systematisch voneinander unterschieden. Das Ereignis $p q$ wurde fast ausschließlich als bestätigend beurteilt, während das Ereignis $p \neg q$ vornehmlich als widerlegend eingeschätzt wurde. Die übrigen beiden Ereignisse wurden in erster Linie als widerlegend oder irrelevant bewertet. Die signifikanten Haupteffekte des Faktors Instruktion, welche

durch eine signifikante Interaktion mit dem Faktor Ereignis qualifiziert werden, repräsentieren den eigentlich interessierenden Effekt der Instruktionsvariation. In der Bedingung Nachfragen wurden die Ereignisse $\neg p q$ und $\neg p \neg q$ hauptsächlich als irrelevant beurteilt, wohingegen diese beiden Zustände in der Bedingung Sortieren eher als widerlegend angesehen wurden. In der Bedingung Nachfragen wurden zusammengenommen 77 fehlerhafte erste Beurteilungen abgegeben (bei einer Gesamtzahl von 336 ersten Einschätzungen), wobei die Anzahl je Kind zwischen 0 und 8 variierte ($M = 3.67$, $SD = 2.83$). In 35 Fällen wurde das inkorrekte Urteil beibehalten, in 2 weiteren Fällen fand ein Wechsel zur anderen inkorrekten Beurteilung statt, während in 40 Fällen eine Änderung zur korrekten Antwort zu verzeichnen war. In 5 Fällen waren Ereignisse der Form $p \neg q$ betroffen (2mal kein Wechsel auf korrekte Antwort vs. 3mal Wechsel auf korrekte Antwort), wohingegen es sich in 42 Fällen um Ereignisse der Form $\neg p q$ (24mal vs. 18mal) und in 30 Fällen um Ereignisse der Form $\neg p \neg q$ (11mal vs. 19mal) handelte. Ergänzend hierzu wurden die von den Kindern für ihre endgültigen Antwortwahlen abgegebenen Begründungen dahingehend kategorisiert, ob sie korrekte Rechtfertigungen enthielten. Für Ereignisse der Form $p \neg q$ wurde eine Begründung als korrekt angesehen, wenn sie sowohl auf die Kongruenz von Antezedenz in Annahme und Ereignis als auch auf den Unterschied zwischen beidem in der Konsequenz hinwies und so die Annahme falsifizierte. Für Ereignisse der Formen $\neg p q$ und $\neg p \neg q$ wurden Rechtfertigungen als korrekt angesehen, wenn sie die Diskrepanz in der Antezedenz zwischen Annahme und Ereignis darlegten. Die Kodierung wurde blind gegenüber den betroffenen Ereignissen und zugehörigen Antwortwahlen allein auf Basis der transkribierten Begründungen vorgenommen. Alle Begründungen wurden auf gleiche Weise von einem Zweitkodierer kategorisiert. Die Beurteilerübereinstimmung betrug $\kappa = .79$. In den 37 Fällen, in denen kein Wechsel auf die korrekte Antwort stattfand, wurden – im Widerspruch zur fehlerhaften endgültigen Einschätzung des Ereignisses – 4mal korrekte Rechtfertigungen abgegeben (ca. 11% der Fälle). In den 40 Fällen mit einem Wechsel auf die korrekte Einschätzung wurden 31mal korrekte Begründungen gegeben (ca. 78% der Fälle).

Ähnlich wie im ersten Experiment wurden die vier zu einer Annahme gehörenden Einschätzungen auf Individualebene zu Antwortmustern zusammengefasst (vgl. Barrouillet et al., 2008). Lag bei einem Kind für mindestens drei der vier Annahmen dasselbe Antwortmuster vor, so wurde es dem entsprechenden Verständnisniveau zugewiesen. Nach dieser Definition antworteten 15 der Kinder der Bedingung Nachfragen konsistent (1mal regelhaft, 2mal bikonditional und 12 konditional). Demgegenüber konnten nur 8 der Kinder der Bedingung Sortieren als konsistent klassifiziert werden (2mal konjunktiv, 3mal regelhaft, 2mal bikonditional und 1mal konditional).

Insgesamt zeigten die Kinder bei prozessorientierter instruktionaler Unterstützung durch Nachfragen ein Niveau des deduktiven Schließens, welches das von Neuntklässlern bei abstrakten Inhalten spontan gezeigte Niveau sogar überstieg (Barrouillet et al., 2008). Prinzipiell ist eine wirksame Unterstützung des deduktiven Schlussfolgerns im Grundschulalter also möglich. Neben den eingesetzten adaptiven Fragen könnten allerdings weitere Faktoren zur Wirksamkeit der Bedingung Nachfragen beigetragen haben. Möglicherweise wussten die Kinder aus der Aufforderung zum nochmaligen Nachdenken, dass ein Wechsel der gewählten Antwortkategorie von ihnen erwartet würde. Die hohe Über-

einstimmung von korrekten Antwortwahlen mit korrekten Rechtfertigungen spricht allerdings dafür, dass viele der betroffenen Kinder nicht gedankenlos wechselten, sondern (zumindest zeitweise) zu einem neuen Verständnis der vorgelegten Ereignisse gelangten. Darüber hinaus erhielten die Kinder durch den Anstoß zum erneuten Nachdenken jedoch auch eine zweite Gelegenheit, sich mit der Einschätzung von Ereignissen zu beschäftigen, so dass sich die beiden Versuchsbedingungen hinsichtlich des Ausmaßes der Beschäftigung mit dem Inhalt voneinander unterschieden. Allerdings ist auch zu bedenken, dass in der Bedingung Sortieren jedes Kind alle kritischen Ereignisse genau zweimal und die vier Ereignisse der Beispielaufgabe verarbeitete, in der Bedingung Nachfragen hingegen nur im Falle einer fehlerhaften Einschätzung eine zweimalige Auseinandersetzung mit der jeweiligen Annahme und dem betroffenen Ereignis stattfand. Außerdem wurden nur zwei Ereignisse der Beispielaufgabe bearbeitet. In der Bedingung Sortieren wurde also zwar häufiger mit Ereignissen umgegangen, in der Bedingung Nachfragen erfolgte die wiederholte Bearbeitung jedoch adaptiv und betraf auch Annahmen.

6. Generelle Diskussion

In beiden Untersuchungen wurden Grundschulkindern einfache Annahmen bezüglich eines physikalischen Phänomens präsentiert. Die Kinder beobachteten ausgewählte Ereignisse und beurteilten diese in Hinblick auf die Gültigkeit der vorgelegten Annahmen. Bei der spontanen Bewertung dieser Ereignisse verfehlten die Grundschul Kinder, wie zu erwarten, meist das Niveau eines konditionalen Verständnisses (Experiment 1). Stattdessen wurden die Annahmen häufig bikonditional in einem regelhaften Sinne interpretiert, d.h. die Kinder betrachteten die Annahmen als geltende Regeln und beurteilten, ob die Ereignisse diesen Regeln gehorchten, anstatt die Gültigkeit der Annahmen kritisch anhand der Ereignisse zu prüfen. Der hypothetische Charakter der vorgelegten Annahmen wurde nicht erkannt. Offensichtlich weist die verwendete *truth-testing task* auf einen Berührungspunkt zwischen deduktivem Denken und Wissenschaftsverständnis, wobei gegenwärtig diskutiert wird, ob die Beachtung des hypothetischen Charakters mancher Annahmen ein inhärenter Bestandteil deduktiver Denkprozesse ist (Barrouillet et al., 2008) oder aber eine besondere meta-logische Fähigkeit darstellt (Moshman, 1990) und welche Bedeutung dem deduktiven Schließen im Rahmen des wissenschaftlichen Denkens zukommt (vgl. Zimmerman, 2007).

Abweichend von den Ergebnissen des ersten Experimentes zeigten Grundschul Kinder ein vorwiegend konditionales Verständnis der vorgelegten Annahmen, wenn sie durch adaptives Nachfragen bei der Verarbeitung der Merkmale von Annahmen und Ereignissen unterstützt wurden (Experiment 2). Ähnlich förderliche Effekte aufmerksamkeitslenkender Fragen sind auch aus der entwicklungspsychologischen Forschung zur konzeptuellen Kategorienbildung bekannt (Saalbach & Schalk, in Druck). Vor diesem Hintergrund eröffnet sich für zukünftige Untersuchungen die Frage, auf welche Faktoren genau die Wirksamkeit unserer auf Nachfragen gestützten Instruktionsbedingung zurückzuführen ist. Ebenso sollte geklärt werden, ob sich die Ergebnisse auf Annahmen mit anderen relevanten Merkmalen oder auf Annahmen zu anderen naturwissenschaftlichen Phänomenen generalisieren lassen.

Die beobachteten korrekten Schlussfolgerungen sind allerdings nicht gleichzusetzen mit dem Erreichen inhaltlicher Lernziele. Ist deduktives Denken im Rahmen diskrepanter Ereignisse nun tatsächlich instrumentell für die Anregung konzeptueller Veränderungen auf der Inhaltsebene? Die Anforderungen der *truth-testing task* bilden einen entscheidenden Anteil jener kognitiven Prozesse ab, welche von Lernenden geleistet werden müssen, um in einem naturwissenschaftlichen Unterricht eigene Vorstellungen überprüfen und Interpretationen anderer Lernender nachvollziehen zu können. Eben solche Denkprozesse beinhaltet ein an sozial-konstruktivistische Ideen angelehntes Unterrichtssetting (Reusser, 2006). Die Umsetzung von Instruktionsmaßnahmen zum deduktiven Schließen in die Unterrichtssituation stellt also ein weiteres Feld für nachfolgende Forschung dar. Im Sinne einer psychologisch fundierten Didaktik, nach welcher die Lehrperson auf Seiten der Lernenden die Anwendung passender Denkoperationen auf den jeweiligen Lerngegenstand anregt, könnten die in dieser Arbeit eingesetzten Fragen ein Baustein für eine adaptive Unterrichtsgestaltung im Zusammenhang mit diskrepanten Ereignissen sein (Beck et al., 2008).

Literatur

- Barrouillet, P., Gauffroy, C. & Lecas, J. F. (2008). Mental models and the suppositional account of conditionals. *Psychological Review*, 115, 760–772.
- Beck, E., Baer, M., Guldemann, T., Bischoff, S., Brühwiler, C., Müller, P., et al. (2008). *Adaptive Lehrkompetenz*. Münster: Waxmann.
- Chan, C., Burtis, J. & Bereiter, C. (1997). Knowledge building as a mediator of conflict in conceptual change. *Cognition and Instruction*, 15, 1–40.
- Chinn, C. A. & Brewer, W. F. (1993). The role of anomalous data in knowledge acquisition: a theoretical framework and implications for science education. *Review of Educational Research*, 63, 1–49.
- diSessa, A. (2006). A history of conceptual change research. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 265–281). Cambridge, UK: Cambridge University Press.
- Dreyfus, A., Jungwirth, E. & Eliovitch, R. (1990). Applying the “cognitive conflict” strategy for conceptual change – some implications, difficulties and problems. *Science Education*, 74, 555–569.
- Duit, R. (2007). *STCSE-Bibliography: Students’ and teachers’ conceptions and science education*. Kiel: IPN.
- Duit, R. & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, 25, 671–688.
- Duit, R., Treagust, D. & Widodo, A. (2008). Teaching science for conceptual change: Theory and practice. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 629–646). New York: Routledge.
- Einsiedler, W. & Hardy, I. (Hrsg.). (2010). Kognitive Strukturierung im Unterricht [Themenheft]. *Unterrichtswissenschaft*, 38(3).
- English, L. (1997). Interventions in children’s deductive reasoning with indeterminate problems. *Contemporary Educational Psychology*, 22, 338–362.
- Gorsky, P. & Finegold, M. (1994). The role of anomaly and of cognitive dissonance in restructuring students’ concepts of force. *Instructional Science*, 22, 75–90.
- Guzzetti, B. J., Snyder, T. E., Glass, G. V. & Gamas, W. S. (1993). Promoting conceptual change in science: A comparative meta-analysis of instructional interventions from reading education and science education. *Reading Research Quarterly*, 28, 116–159.

- Hashweh, M. Z. (1986). Toward an explanation of conceptual change. *European Journal of Science Education*, 8, 229–249.
- Johnson-Laird, P. N. & Byrne, R. M. J. (2002). Conditionals: A theory of meaning, pragmatics, and inference. *Psychological Review*, 109, 646–678.
- Kang, S., Scharmann, L. C. & Noh, T. (2004). Reexamining the role of cognitive conflict in science concept learning. *Research in Science Education*, 34, 71–96.
- Krammer, K. (2009). *Individuelle Lernunterstützung in Schülerarbeitsphasen*. Münster: Waxmann.
- Lawson, A. E. & Thompson, L. D. (1988). Formal reasoning ability and misconceptions concerning genetics and natural selection. *Journal of Research in Science Teaching*, 25, 733–746.
- Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and Instruction*, 11, 357–380.
- Markovits, H., Schleifer, M. & Fortier, L. (1989). The development of elementary deductive reasoning in young children. *Developmental Psychology*, 25, 787–793.
- Moshman, D. (1990). The development of metalogical understanding. In W. F. Overton (Ed.), *Reasoning, necessity, and logic: Developmental perspectives* (pp. 110–135). London: Routledge.
- Posner, G. J., Strike, K. A., Hewson, P. W. & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, 66, 211–227.
- Rea-Ramirez, M. A. & Clement, J. (1998). *In search of dissonance: The evolution of dissonance in conceptual change theory*. Paper presented at the Annual Meeting of the National Association for Research in Science Teaching.
- Reusser, K. (2006). Konstruktivismus – vom epistemologischen Leitbegriff zur Erneuerung der didaktischen Kultur. In M. Baer, M. Fuchs, P. Füglistner, K. Reusser & H. Wyss (Hrsg.), *Didaktik auf psychologischer Grundlage: Von Hans Aebli's kognitionspsychologischer Didaktik zur modernen Lehr- und Lernforschung* (S. 151–168). Bern: h.e.p.
- Saalbach, H. & Schalk, L. (in Druck). Fragen stellen hilft: Die Aktivierung von Vorwissen fördert die Nutzung kategorialer Beziehungen in Wortlernaufgaben bei jungen Kindern. In F. Vogt, M. Leuchter, A. Tettenborn, E. Wannack, U. Hottinger & M. Jäger (Hrsg.), *Entwicklung und Lernen junger Kinder*. Münster: Waxmann.
- Sandoval, J. (1995). Teaching in subject matter areas: Science. *Annual Review of Psychology*, 46, 355–374.
- Stavy, R. (1991). Using analogy to overcome misconceptions about conservation of matter. *Journal of Research in Science Teaching*, 28, 305–313.
- Thiel, S. (1973). Grundschulkinde zwischen Umgangserfahrung und Naturwissenschaft. In M. Wagenschein, A. Banholzer & S. Thiel (Hrsg.), *Kinder auf dem Weg zur Physik* (S. 90–180). Stuttgart: Klett.
- Wason, P. C. (1966). Reasoning. In B. Foss (Ed.), *New horizons in psychology*. Harmondsworth, UK: Penguin.
- Wason, P. C. & Shapiro, D. (1971). Natural and contrived experience in a reasoning problem. *Quarterly Journal of Experimental Psychology*, 23, 63–71.
- White, R. T. & Gunstone, R. F. (2008). The conceptual change approach and the teaching of science. In S. Vosniadou (Ed.), *International handbook of research on conceptual change* (pp. 619–628). New York: Routledge.
- Ziegler, A. (1990). Deduktives Schließen mit mentalen Modellen. *Sprache und Kognition*, 9, 82–91.
- Zimmerman, C. (2007). The development of scientific thinking skills in elementary and middle school. *Developmental Review*, 27, 172–223.

Anhang

Annahmen und zugehörige Ereignisse

Bälle mit Luft springen.^a/Wenn ein Ball mit Luft gefüllt ist, dann springt er.^b [$p \ q$]

Der Hunde-Ball ist mit Luft und springt. [$p \ q$]

Der hohle Fensterkitt-Ball ist mit Luft und springt nicht. [$p \ \neg q$]

Die Eisenkugel ist ohne Luft und springt. [$\neg p \ q$]

Der rote Knet-Ball ist ohne Luft und springt nicht. [$\neg p \ \neg q$]

Weiche Bälle springen.^a/Wenn ein Ball weich ist, dann springt er.^b [$p \ q$]

Softball [$p \ q$], Jonglier-Ball [$p \ \neg q$], Murmel [$\neg p \ q$], trockener Knet-Klumpen [$\neg p \ \neg q$]

Wenn ein Ball rund ist, dann springt er.^b [$p \ q$]

Tischtennisball^b [$p \ q$], Knet-Ball^b [$p \ \neg q$], Flummi-Ei^b [$\neg p \ q$], gefüllter Luftballon^b [$\neg p \ \neg q$]

Wenn ein Ball aus Gummi ist, dann springt er.^b [$p \ q$]

Flummi^b [$p \ q$], Knäuel aus Gummibändern^b [$p \ \neg q$], Golfball^b [$\neg p \ q$], Fensterkitt^b [$\neg p \ \neg q$]

Beispielaufgabe

Große Bälle springen.^a/Wenn ein Ball groß ist, dann springt er.^b [$p \ q$]

Hunde-Ball^c [$p \ q$] gefüllter Luftballon [$p \ \neg q$], Holzkugel [$\neg p \ q$], Knet-Kügelchen^c [$\neg p \ \neg q$]

^a Experiment 1; ^b Experiment 2; ^c Experiment 2 nur Bedingung Sortieren.

Anschriften der Autoren:

Dipl.-Psych. Steffen Tröbst und Prof. Dr. Kornelia Möller, Westfälische Wilhelms-Universität Münster, Seminar für Didaktik des Sachunterrichts, Leonardo-Campus 11, 48149 Münster

Prof. Dr. Ilonca Hardy, Goethe-Universität Frankfurt, Institut für Pädagogik der Elementar- und Primarstufe, Fachbereich Erziehungswissenschaften, Senckenberganlage 13-17, 60054 Frankfurt am Main

Beitrag eingegangen: 9.9.10; revidiert: 28.10.10; angenommen: 3.11.10