

# Wirkungen einer Öffnung des Unterrichts

## Befunde und Probleme der empirischen Forschung

- revisited 2008<sup>323</sup>

### 1 Der Gegenstand: verschiedene Aspekte der Öffnung

„Öffnung des Unterrichts“ ist eine Reformperspektive, die drei verschiedene Probleme in den Blick nimmt:

1. Wie ist die *Rolle von Kindern als SchülerInnen* definiert? Dies gilt insbesondere im Verhältnis zur Lehrperson und im Blick auf die Wahl- und Mitwirkungsmöglichkeiten, die ihnen *im Unterricht* gewährt werden („offener Unterricht“ im engeren Sinne)<sup>324</sup>.

2. Wie wird das *Verhältnis der Schule zur Lebenswelt* der SchülerInnen bestimmt, aus der diese tagtäglich kommen und in die sie mittags wieder zurückkehren? Welche Funktion hat die Schule in der *Gemeinde* („community“), die sie trägt? („offene Schule“ als Erfahrungsraum)<sup>325</sup>.

3. Wie sind die Verantwortlichkeiten verteilt zwischen den unmittelbar am Unterricht Beteiligten (vor allem LehrerIn und SchülerInnen) einerseits und denjenigen, die Unterricht *von außen* - z. B. über Lehrpläne oder Lehrwerke - planen? („offenes Curriculum“)<sup>326</sup>.

Der folgende Beitrag konzentriert sich auf die erste Perspektive, die Schülerzentrierung des Unterrichts, auch weil diese die empirische Forschung im Wesentlichen bestimmt hat. Dabei sollte aber nicht übersehen werden, dass die *Öffnung des Unterrichts* in hohem Ma-

<sup>323</sup> Für die Überarbeitung verdanke ich viele Hinweise den beiden Dissertation meiner Kollegen Falko Peschel (2003) und Horst Heinzel (2006), denen ich für diese vorzüglichen Arbeiten auch hier noch einmal danken möchte.

<sup>324</sup> Vgl. Rumpf (1971); Garlichs u. a. (1974); Ramseger (1977); für die USA vor allem Silberman (1973).

<sup>325</sup> Vgl. Hentig 1973; Zimmer 1973).

<sup>326</sup> Vgl. Brügelmann (1972); Sachs/ Scheilke (1973); Deutscher Bildungsrat (1974).

ße abhängig von dem institutionellen Selbstverständnis der *Schule* und dass ihr Gelingen angewiesen ist auf eine *fachdidaktisch* stimmige Übersetzung der allgemeinen Prinzipien.

„Offener Unterricht“ ist also ein mehrdimensionaler Begriff. Schon wegen dieser Mehrdeutigkeit ist es kein Wunder, dass die Diskussion kontrovers geführt wird, dass die Befunde der Empirie widersprüchlich erscheinen. Als entscheidend setze ich zur Abgrenzung des offenen vom lehrerzentrierten Unterricht im Folgenden das Kriterium der Selbstständigkeit beim und der Mitverantwortung für das eigene Lernen. Diese kann sich in drei Dimensionen zeigen<sup>327</sup>

- *methodisch-organisatorisch* in Freiräumen bei der Bearbeitung inhaltlich (d. h. in Lösung und Lösungsweg) geschlossener Aufgaben;
- *didaktisch-inhaltlich* in der Zulässigkeit/ Herausforderung von Varianten bei der Lösung vorgegebener Aufgaben auf der Grundlage individueller Erfahrungen und Konzepte;
- *pädagogisch-institutionell* in der (Mit-)Entscheidung von SchülerInnen über Ziele, Inhalte und Rahmenbedingungen der (individuellen oder gemeinsamen) Arbeit.

Schon diese Dimensionierung macht deutlich, dass sich Öffnung des Unterrichts nicht auf die methodische Ebene beschränkt und dass sie vor allem nicht mit Binnendifferenzierung gleichzusetzen ist, sondern tiefer greift und grundsätzlicher begründet wird.

Bei der Einordnung von Untersuchungen zur Öffnung des Unterrichts sind zunächst verschiedene Ebenen der Auseinandersetzung mit diesem Ansatz zu unterscheiden:

- *Definitionen* und Einordnungen in *systematischer* Perspektive<sup>328</sup>
- *theoretische Begründungen*, z. T. in *historischer* Perspektive<sup>329</sup>
- *Vorschläge* für die Umsetzung in *praktischer* Perspektive<sup>330</sup>
- *Einschätzungen* von Breite und Intensität der Realisierung sowie von ihren Wirkungen aus *empirischer* Perspektive<sup>331</sup>.

Gegenstand dieses Beitrags ist die letzte Ebene. Im Wesentlichen müssen wir dabei auf die Forschung im angelsächsischen Raum zurückgreifen. Deutschsprachige Untersuchungen

<sup>327</sup> Ausführlicher dazu mein einleitender Beitrag zu diesem Band; vgl. im Einzelnen zur Definitionsproblematik unten Kap. 3.1.

<sup>328</sup> Z. B. Ramseger (1977); Wallrabenstein (1991, 170f.).

<sup>329</sup> Z. B. Benner (1989); Jürgens (1994, 24ff.).

<sup>330</sup> Z. B. Kasper u. a. (1989); Wallrabenstein (1991).

<sup>331</sup> Z. B. Einsiedler (1981); Jürgens (1995b).

gibt es nicht viele. Andererseits ist es nur am Rande möglich, auch Studien zu berücksichtigen, die in einer ähnlichen Perspektive, aber mit besonderen Akzenten durchgeführt wurden, z. B. zum entdeckenden Lernen, zur Individualisierung des Unterrichts, zur Schülerorientierung, zu Grundqualifikationen vs. Fachleistungen<sup>332</sup>.

Zwar gibt es Überschneidungen, weil dieselben Fragen zu verschiedenen Zeiten oder in verschiedenen Disziplinen mit unterschiedlichen Etiketten benannt wurden; es gibt aber auch bedeutsame Unterschiede im Fokus der Untersuchungen. In den 1960er und teilweise auch in den 1970er Jahren war für die Forschung zu Unterrichtsformen beispielsweise das Konzept des „entdeckenden Lernens“ bestimmend<sup>333</sup>. Die Fragestellung allerdings war primär lerntheoretisch und fachdidaktisch („Wie eignen sich Kinder bestimmte Inhalte an?“), weniger schultheoretisch und pädagogisch („Was trägt die Schule zur Entwicklung der Kinder als individueller Persönlichkeit und als Bürger in einer demokratischen Gesellschaft bei?“). Damit war sie auch nicht auf die Sozialformen und Mitwirkungsrechte in den Klassen und in der Schule gerichtet<sup>334</sup>. Beide sind aber als komplementär zu sehen.

## 2 Ein Rückblick auf die frühere Forschung

Programmatisch hat die Öffnung des Unterrichts seit den 1970er Jahren eine zunehmende Zustimmung gefunden – vor allem im Grundschulbereich<sup>335</sup>. Kritik und Gegenpositionen wurden in der Regel allenfalls vorsichtig formuliert<sup>336</sup>. Stärker als die Ziele und Prinzipien werden die Umsetzbarkeit und die Wirkungen offenen Unterrichts in Zweifel gezogen – vor allem im Blick auf die Vergleichbarkeit fachlicher Leistungen mit den Ergebnissen in traditionellem Unterricht.

In den USA haben empirische Studien zu offenem Unterricht eine lange Tradition. Horwitz (1979, 73) berichtet über eine Vielzahl US-amerikanischer Studien zur „progressive education“ aus den 1930er und 1940er Jahren<sup>337</sup>. Solche reformpädagogisch inspirierten

<sup>332</sup> Vgl. die Einordnung einschlägiger Befunde in den Metaanalysen von Walberg (1986); Fraser u. a. (1989).

<sup>333</sup> Vgl. Bruner (1961); Einsiedler (1976); Neber (1981).

<sup>334</sup> Vgl. dazu die Ansätze der Schulkritik, u. a. bei Freire (1977) und Illich (1972; 1973), und in der sog. Antipädagogik, vor allem bei v. Braunmühl (1977) und v. Schoenebeck (1985).

<sup>335</sup> Vgl. die vielfältigen Beiträge in den Publikationen des Grundschulverbandes und in den einschlägigen Fachzeitschriften.

<sup>336</sup> Vgl. Schulz (1989); Winkel (1993); Grüntgens (1997); s. als Ausnahme: Günther (1988; 1996) und populistische Vorwürfe gegen eine angebliche „Spaßpädagogik“ (Kraus 2000) oder eine „Kuschelpädagogik“ (so Bundespräsident Herzog in seiner sog. „Ruck“-Rede im Berliner Hotel Adlon, dokumentiert u. a. in: Frankfurter Rundschau vom 6. November 1997).

<sup>337</sup> Vgl. vor allem die berühmte 8-Year-Study von Chamberlin u. a. (1942).

Ansätze waren schon damals in der öffentlichen Meinung umstritten. Sie wurden deshalb zum willkommenen Experimentierfeld für die gerade aufblühende Testpsychologie.

Die qualitativen Überblicke über diese Studien von Wrightstone (1938), Baker u. a. (1941) und Leonard/ Eurich (1942) kommen zu einem übereinstimmenden Urteil: Gegenüber traditionellem Unterricht gebe es keine Nachteile in den Fachleistungen; auf der anderen Seite seien deutliche Vorteile in verschiedenen Dimensionen der Persönlichkeitsentwicklung (Initiative, soziales Verhalten, Problemlösen u. ä.) festzustellen. Ähnlich fielen auch die Ergebnisse einer englischen Längsschnittstudie von Gardner (1942; 1950; 1966) aus.

Eine zweite Welle empirischer Studien folgte in den 1960er und 1970er Jahren. Horwitz (1979) fasste ihre Ergebnisse relativ grob zusammen, indem er für verschiedene Lernzielbereiche ausählte, wie viele Studien Vorteile für offene bzw. für traditionelle Ansätze nachwiesen. Zwar stellte er im Mittel einen leichten Rückstand offener Ansätze in den fachlichen Leistungen fest; diese waren aber eher geringfügig, und ihnen standen wiederum deutliche Vorteile im Persönlichkeitsbereich gegenüber<sup>338</sup>. Studien zum „demokratischen“ vs. „autoritären“ Führungsstil erbrachten im Leistungsbereich ebenfalls eine Patt-Situation: bessere fachlichen Lernerfolge „demokratisch“ geführter Klassen in 11 Studien, schlechtere in 8 Studien und keinen Unterschied in 13 Studien (Anderson 1959).

Die Schwäche von Horwitz' „box-score“-Verfahren: Große und kleine Studien, geringe und markante Unterschiede schlugen gleichwertig zu Buche<sup>339</sup>. Immerhin wurde so sichtbar, dass eine große Zahl von Studien gar keine Unterschiede gefunden hatte.

Differenzierter sind Metanalysen, die die Werte der einzelnen Untersuchungen quantitativ verrechnen. Die resultierenden „Effekt-Stärken“<sup>340</sup> sind zugleich ein Maß für die praktische Bedeutsamkeit von Unterschieden. Dabei gelten nach Cohen (1977) Effektstärken von .20 als schwach, von .50 als mäßig und von .80 als stark.

Auf diese Weise haben verschiedene Autoren die Ergebnisse aus empirischen Originalstudien der 1960er und 1970er Jahre zusammengefasst<sup>341</sup>. Auch ihr Fazit: In den Fachleis-

<sup>338</sup> Analog Walberg (1982).

<sup>339</sup> S. aber zur – im Ergebnis – starken Übereinstimmung von Box Score und Effektstärke-Maßen: Walberg (1985, 218, 220).

<sup>340</sup> Die Effektstärke wird statistisch errechnet als Differenz der Mittelwerte zwischen zwei Gruppen, geteilt durch die Standardabweichung [also durch das Maß der Streuung] der Werte in der Kontrollgruppe; vgl. Walberg (1985, 216); ausführlicher: Lind (2005).

<sup>341</sup> Vgl. Peterson (1979; 1980), Hetzel u. a. (1980) und Hedges u. a. (1981; zusammengefasst in Giaconia/ Hedges 1982).

tungen sind Klassen mit offenem Unterricht im Mittel geringfügig unterlegen, in den Einstellungen (zum Lernen, zu den Lehrpersonen) und in den Grundqualifikationen dagegen erreichen sie bessere Ergebnisse.

Die folgende Übersicht<sup>342</sup> macht die Tendenzen in der rechten Spalte auf einen Blick sichtbar (vgl. zu den genauen Werten die beiden mittleren Spalten):

| Bereich                  | Effektstärke in den Analysen |                            | Richtung und Ausmaß <sup>343</sup><br>der Unterschiede über<br>alle Studien hinweg |
|--------------------------|------------------------------|----------------------------|--|
|                          | Peterson<br>45 Studien       | Hedges u.a.<br>153 Studien |  |
| Allgemeine Schulleistung | -.12                         |                            | ■  |
| Mathematik               | -.14                         | -.04                       | ■  |
| Lesen                    | -.13                         | -.08                       | ■  |
| Sprache                  |                              | -.07                       | ■  |
| Verschiedenes            |                              | -.15                       | ■  |
| Denkfähigkeit allgemein  |                              | +.18                       | ■  |
| Kreativität              | +.18                         | +.29                       | ■  |
| Unabhängigkeit           | +.30                         | +.28                       | ■  |
| Neugier                  | +.14                         |                            | ■  |
| Angst                    | -.07                         |                            | ■  |
| Selbstkonzept            | +.16                         | +.07                       | ■  |
| "locus of control"       | +.03                         |                            | ■  |
| Kooperation              |                              | +.23                       | ■  |
| Einstellung zur Schule   | +.12                         | +.17                       | ■  |
| Einstellung zur Lehrerin | +.42                         |                            | ■  |

Die Auswertung zeigt also durchgängige, aber nur geringfügige Vorteile traditionellen Unterrichts in den Fachleistungen, durchgängige und deutlichere (aber mit .10 - .40 immer noch begrenzte) Vorteile offener Ansätze im Persönlichkeits- und Einstellungsbereich:

Auf den ersten Blick kommt Haenisch (1991) in seiner Zusammenfassung ausgewählter deutschsprachiger Studien zum Prinzip der "Schülerorientierung" zu einer positiveren Einschätzung:

1. Schülerorientierung reduziert Unterrichtsstörungen und Disziplinprobleme
2. Schülerorientierung fördert Schulleistung und Motivation

<sup>342</sup> Differenzierter: Einsiedler (1990b); Schöll (1996).

<sup>343</sup> Blöcke, die vom Nullpunkt nach rechts zeigen, veranschaulichen einen Vorteil offener Ansätze, Blöcke nach links verweisen auf Vorteile traditionellen Unterrichts (jeder Block steht für eine Effektstärke von etwa .05)

### 3. Schülerorientierung erhöht die Lernfreude und mildert die Schulmüdigkeit.

Haenisch differenziert seinen Kernbegriff der Schülerorientierung jedoch in bedeutsamer Weise, wenn er den in diesen Perspektiven erfolgreichen Unterricht auch durch drei „aber“ charakterisiert<sup>344</sup>:

- „Bezogen auf die Themen und Inhalte ... wird auch deutlich aufgabenorientiert gearbeitet und der Unterricht ist häufig systematisch auf Logik und Ordnung aufgebaut.“
- Bezogen auf die Methoden, Arbeits- und Lernformen ... haben [LehrerInnen] - wenn immer es nötig ist - Mut, den Unterricht zu steuern, ordnende Regie zu führen und Hilfe zu geben; der Unterricht ist gut strukturiert und flexibel zugleich.
- Bezogen auf die Rolle als LehrerIn... LehrerInnen treten selbstbewusst auf, sind konsequent und sorgen für klare Verhältnisse, sind (auch) Autoritätsperson und können sich durchsetzen.“

Diese Einschränkungen entsprechen auf den ersten Blick gängigen Befunden der Unterrichtsforschung. Vor allem wird im Anschluss an Helmke/ Weinert (1997a, 248) aus der Münchener SCHOLASTIK-Studie zur Entwicklung im Grundschulalter immer wieder die Bedeutung „effizienter Klassenführung“ und „Strukturierung“ des Unterrichts betont<sup>345</sup>.

Allerdings ist die Datengrundlage weniger deutlich als die Folgerungen der AutorInnen: Die Korrelationen zwischen verschiedenen Merkmalen des Unterrichts und dem Lernfortschritt der entsprechenden Klassen liegen für die Rechtschreibung zwischen -.04 und .27\*<sup>346</sup>. Dies sind nicht eindeutige, sehr lose und statistisch nur teilweise signifikante Zusammenhänge. Nicht überraschend zeigt die Einzelfallanalyse der sechs erfolgreichsten LehrerInnen so bizarr unterschiedliche Profile, dass es schwer fällt, daraus Empfehlungen für einen erfolgreichen Unterrichtsstil abzuleiten (a. a. O., 250). Hinzu kommt, dass „Struktur“ nur als Merkmale der Lehrersteuerung erfasst wurde und die teilnehmenden

<sup>344</sup> A. a. O., 13-14. Die Einschätzungen ergeben sich allerdings primär aus Studien im Sekundarbereich.

<sup>345</sup> Vgl. etwa Lipowsky (2007a), der wie fast alle AutorInnen nicht diskutiert, aus wessen Sicht ein Unterricht „strukturiert“ erscheint bzw. sein sollte: aus fachdidaktischer oder sozialpsychologischer Sicht? aus Lehrer oder Schülersicht? Diese Kritik trifft auch die Kriterien für „guten Unterricht“ in den viel zitierten Handbüchern von Helmke (2003) und Meyer (2004) und Studien wie die viel zitierten Befunde aus den Münchener Studien zum Sachunterricht, z. B. Hardy u. a. (2006).

<sup>346</sup> Im Mittel bei .10 bis .20, nicht signifikant; für Mathematik etwas höher mit im Durchschnitt .30\*.

Klassen die Bandbreite von lehrer- bis schülerzentriertem Unterricht gar nicht abgebildet haben<sup>347</sup>.

Die Befunde der Forschung sind also nicht so einfach zu interpretieren, wie manche Zusammenfassungen suggerieren. Wesentlich sind folgende Punkte:

1. Über viele Studien hinweg zeigt sich, dass es *durchaus Unterschiede* zwischen sog. offenem und sog. lehrerzentriertem (bzw. durch Material stark vorstrukturiertem) Unterricht in beobachtbaren Verhaltensmerkmalen von SchülerInnen gibt – sowohl im Leistungs- als auch im Persönlichkeitsbereich. Die Dimension "Offenheit" ist also nicht irrelevant für die Beurteilung von didaktisch-methodischen Alternativen.

2. Die Unterschiede sind in der Regel *bereichsspezifisch*.

- Im fachlichen Bereich sind die Leistungen offener Ansätze im Mittel etwas niedriger, zugleich streuen die Leistungen oft breiter.
- Dagegen schneiden offene Ansätze im Bereich der Grundqualifikationen, der Einstellungen und der Persönlichkeitsentwicklung besser ab.

3. Die Unterschiede sind – gemittelt über verschiedene Studien hinweg – *gering*.

Aus diesen Tatbeständen werden entgegengesetzte Folgerungen gezogen:

Die bildungspolitisch konservative Position lautet: Die zu geringen Vorteile lohnen den ökonomischen Aufwand und die sozialen und emotionalen Kosten einer Schulreform zur Durchsetzung von Offenheit nicht. Erst wenn empirisch bewiesen werden kann, dass offener Unterricht generell besser ist, darf er generell eingeführt werden.

Die liberale Position widerspricht: Zwar sei es richtig, dass sich offener Unterricht nicht „verordnen“ lässt<sup>348</sup>. Die Beweislast für eine Öffnung dem offenen Unterricht aufzuerlegen

---

<sup>347</sup> Vgl. Brügelmann (1998c) und zur Kritik an dieser viel (und oft leichtfertig zitierten) Studie (s. zuletzt Einsiedler u. a. 2008, 346) auch Rauin (2004).

<sup>348</sup> Und schon gar nicht flächendeckend, vgl. Hofkins (1991); Marston (1991).

sei aber falsch, sie liege vielmehr bei denen, die Freiräume einschränken wollen, und zwar aus zwei Gründen.

Verfassungsrechtlich gilt: Die Freiheit der Person darf durch den Staat, also auch durch seine Schule, nur mit besonderen Gründen eingeschränkt werden<sup>349</sup>. Da offener Unterricht aber keine bedeutsamen Nachteile im Leistungsbereich zeigt, gibt es keine Rechtfertigung, die Selbständigkeit von Kindern über die Ordnung hinaus einzuschränken, die für ein Zusammenleben und eine Zusammenarbeit in der Klasse erforderlich sind<sup>350</sup>.

Pädagogisch ist zu beachten: Wenn Selbständigkeit, Mitverantwortung und Eigenaktivität als *Ziele* der Schule wichtig sind, dann ist ein Unterricht vorzuziehen, der mit diesen Prinzipien übereinstimmt, solange keine Verluste/ Nachteile in anderen bedeutsamen Zielbereichen nachgewiesen sind.

Dies aber ist die Situation, wie die Metaanalysen zeigen.

Mit einer Beschränkung auf die oben genannten Befunde (1) bis (3) wird allerdings ein Problem verdrängt:

4. Nicht nur sind die Unterschiede gering; oft weisen die Unterschiede sogar in einem Teil der Studien in die eine, in anderen Studien in die andere Richtung – und das mit *erheblicher Streuung*, zum Teil auch zwischen Klassen innerhalb der einzelnen Studien.

Diese *Widersprüche zwischen einzelnen Studien* werfen verschiedene Fragen auf, die eine genauere Prüfung erfordern.

### 3 Einzelbefunde und forschungsmethodische Probleme der Studien

---

<sup>349</sup> S. dazu auch die Diskussion um die UN-Kinderrechtskonvention und ihre Bedeutung für die Schule, aktuell in Backhaus/ Knorre u. a. (2008).

<sup>350</sup> S. dazu auch die Argumente der Antipädagogik für ihre Forderung nach mehr Respekt gegenüber den Rechten der Kinder, z.B. bei Schoenebeck (1985).

Als Beleg für unzureichende Schulleistungen in offenerem Unterricht wird immer wieder die Studie von Bennett u. a. (1976) zitiert, in der Lerneffekte bei „informellem“ vs. „formellem“ Unterrichtsstil in englischen Grundschulen verglichen wurden.

Doch deren Befunde können unter verschiedenen Gesichtspunkten in Frage gestellt werden. Die Kritik ist über diese Studie hinaus bedeutsam, markiert sie doch typische Schwierigkeiten, einen so komplexen pädagogischen Ansatz im Feld zu evaluieren: unterschiedliche *Kontexte* (einige Schulen arbeiteten im Rahmen eines testbezogenen Auslesesystems, andere nicht); unterschiedliche *Umsetzungen* (einige der als "informal classrooms" klassifizierten Lerngruppen haben faktisch eine "laissez-faire" Form praktiziert); unterschiedliche *Ausgangslagen* (im Mittel waren die "informal classrooms" wesentlich leistungsschwächer, die Differenz zu den lehrerzentrierten Klassen entsprach immerhin fast dem Jahreszuwachs der "informal classrooms").

Kasper (1992, 108-109) zweifelt darüber hinaus die Aussagekraft dieser Studie für deutsche Konzepte offenen Unterrichts wegen ihrer mangelnden Übertragbarkeit an. Sie stützt ihre Kritik auf die ORACLE-Studien von Galton u. a. (1980), Galton/ Simon (1980) und Simon/ Willcocks (1981). Diese zeigen,

- dass Individualisierung in englischen Grundschulen zwar weit verbreitet war,
- dass sie aber qualitativ oft nur eine Verlagerung der Steuerung vom Lehrer auf stark vorstrukturierte Aufgaben/ Materialien bedeutete, was zur Folge hatte:
  - ein Nebeneinander statt eines Miteinander bei der Arbeit (bei engem Zusammenhang zwischen geringer Interaktion und Abbrechen bzw. Unstetigkeit der Arbeit);
  - repetitive Arbeitsformen statt entdeckenden Lernens (Lernmaterialien differenzieren nur quantitativ, gehen an spezifischen Interessen, Zugangsweisen, Schwierigkeiten einzelner Kinder vorbei);
  - implizite Voraussetzung kognitiver und sozialer Lernstrategien, statt sie im Unterricht systematisch zu entwickeln.

Forschungsmethodisch wird darüber hinaus kritisiert, dass die Gruppen ("formal" vs. "informal") erst im Nachhinein aus der verfügbaren Stichprobe gebildet (und nicht nach vorab definierten Kriterien für Offenheit bestimmt) wurden. Außerdem gibt es technische Kontroversen darüber, wie die Unterschiede im Lernfortschritt beider Gruppen statistisch bestimmt wurden.

Mit Bezug auf spätere Re-Analysen der Bennett-Studie und auf zwei weitere Untersuchungen stellen Gray/ Satterly (1981, 191) für England fest, dass sich in den Befunden zwar Tendenzen zugunsten besserer Leistungen formellen Unterrichts in den Bereichen Lesen, Sprache und Mathematik erkennen lassen, dass diese aber erstens statistisch nicht signifikant und zweitens in ihrer praktischen Bedeutung zu vernachlässigen seien.

Die Bennett-Gruppe hat ihre Folgerungen wegen der Schwächen in der statistischen Analyse später selbst relativiert<sup>351</sup>. Aber dies ist weder in der Fach- noch in der öffentlichen Diskussion wahrgenommen worden, obwohl die breiter fundierten Metaanalysen aus den USA ein für den offenen Unterricht deutlich positiveres Bild ergaben<sup>352</sup>. Zu Einzelaspekten, insbesondere von Vorteilen einer klaren "Strukturierung" des Unterrichts und von Nachteilen einer "zu weitgehenden" Eigenverantwortung von SchülerInnen, wurden allerdings später Befunde vorgelegt, die durchaus einige Folgerungen aus der Bennett-Studie (und der fast zeitgleichen Studie von Rutter u. a. 1979) stützen<sup>353</sup>.

Allgemein können Gründe für unterschiedliche Befunde in folgenden Differenzen zwischen den Untersuchungen gefunden werden:

- Unterschiede im Verständnis der ForscherInnen von Öffnung des Unterrichts und der Auslegung des *Konzepts* in der untersuchten Praxis, d. h. auch unterrichtsmethodischer Aspekte wie "Strukturierung", "Klarheit", "Führung" (s. unten 3.1-3.3);
- unterschiedliche *Erfolgskriterien*, Lernbereiche/ Fächer, Stichproben und Kontexte der Untersuchungen (s. 3.4-3.7);
- Unterschiede im *Untersuchungsdesign* und
- Unterschiede in den eingesetzten *Erhebungsverfahren* (s. 3.8-3.9).

Hinzu kommt die grundsätzliche Problematik, dass der Unterricht in vielen Untersuchungen wie eine "black box" behandelt wird, so dass Beziehungen zwischen Bedingungen des Unterrichts und seinen Effekten zwar statistisch, aber nicht inhaltlich begründet werden können<sup>354</sup>.

Die Bedeutung dieser Faktoren soll im Folgenden genauer untersucht werden.

---

<sup>351</sup> Vgl. Aitken u. a. (1981a+b).

<sup>352</sup> Vgl. zur Kritik an der unkritischen Rezeption der Bennet-Studie: Walberg (1985, 226); Fraser u. a. (1987, 161-163).

<sup>353</sup> Vgl. Mortimore u. a. (1989).

<sup>354</sup> Vgl. Giest (1997).

### 3.1 Unterschiedliche Definitionen von offenem Unterricht werden unterstellt

In seiner Auswertung von Studien zum Einfluss eines partnerschaftlichen Erziehungsstils in Familien stellt Schimpke (2007; 2008) fest, dass die 202 verfügbaren empirischen Untersuchungen kaum vergleichbar sind, weil sie die Alternativen zu traditionellen, eher autoritären Erziehungsformen in verschiedener Weise definieren. Ebenso gehen verschiedene Untersuchungen im Schulbereich nicht ohne Weiteres von demselben Verständnis offenen Unterrichts aus. Dieses wird zudem oft nicht zureichend geklärt, so dass die Bildung von Untersuchungsgruppen teilweise a-theoretisch erfolgt<sup>355</sup>.

In der oben zitierten Studie von Bennett u. a. (1976/79) beispielsweise werden die Aussagen der LehrerInnen zu sechs verschiedenen Aspekten des Unterrichts (deren systematische Verknüpfung eine Vielfalt unterschiedlicher Unterrichts-"Profile" ergeben würde) in einer ersten Verdichtung induktiv zu zwölf Lehrstilen typisiert, die dann noch einmal zu drei recht groben Typen "informell", "formell" und "gemischt" zusammengefasst werden.

Günther (1988) bezieht in seine Stichprobe zur "Freiarbeit" Klassen aus Regelschulen mit nicht weiter spezifiziertem Konzept aus Freinet-, Montessori- und Petersen-Schulen ein, deren Verständnis von "Offenheit" (wenn sie sich diesen Begriff denn überhaupt zu eigen machen) aus verschiedenen didaktischen Traditionen stammt und in unterschiedlichen methodischen Praktiken niederschlägt<sup>356</sup>.

Diese Kritik unscharfer Begrifflichkeit richtet sich nicht allein gegen die Untersuchungen<sup>357</sup>, sondern spiegelt auch eine unbefriedigende "Selbstdarstellung" offenen Unter-

<sup>355</sup> Vgl. zu diesen Schwierigkeiten auch Kap. 2 meines Beitrags „Die Öffnung des Unterrichts muss radikaler gedacht...“ in diesem Band und zu einer sorgfältigen Diskussion der begrifflichen Unklarheiten auch Heinzl (2006, Kap. 5 und 6).

<sup>356</sup> Vgl. etwa die Sicht von Individualisierung und die Funktion von Materialien bei Montessori vs. Freinet.

<sup>357</sup> Genauso kritisch müsste nach der Definition des "Gegenbilds" gefragt werden. Denn was unter Etiketten wie "Regelunterricht", "geschlossener" oder "traditioneller Unterricht" als Kontrollgruppe herhalten muss, ist häufig sehr heterogen (belegt wurde dies z. B. für die westdeutsche Stichprobe im "Schreibvergleich BRDDR", die weit in die beiden Vergleichsgruppen "systematisch gestufter Rechtschreiblehrgang" (DDR) und "Lesen durch Schreiben" (Schweiz) hineinreichte; vgl. Brügelmann u. a. 1994).

richts durch seine VertreterInnen wider. Der Vorwurf wird aber zunehmend ernst genommen. So grenzt Einsiedler (1988) offenen Unterricht bewusst ab von innerer Differenzierung, die sich auf eine "methodisch-organisatorische Öffnung von oben" beschränkt.

Ramseger (1987a, 6) betont die Verbindung von drei Perspektiven als charakteristisch für offenen Unterricht:

- bestimmte Vorstellungen über das Lernen von Menschen,
- bestimmte Einstellungen zum Kind und zur Kindheit,
- bestimmte Ansprüche an die Funktion der öffentlichen Schule in einer Demokratie.

Peschel (1996; 2002) konkretisiert diese über eine bloß methodische Differenzierung hinausreichende Öffnung, indem er verschiedene Formen offenen Unterrichts in einer Matrix nach der Ausprägung von Selbst- bzw. Mitbestimmung der SchülerInnen in verschiedenen Dimensionen didaktisch-methodischer Entscheidungen ordnet. Als wesentlichen Schritt in dieser Perspektive betonen Schmidt/ Wopp (1996) die Aufgabe des Unterrichts, die Methoden- und Planungskompetenz der SchülerInnen zu fördern, aber auch zu fordern.

Von anderen AutorInnen werden die Grundlagen offenen Unterrichts durch Bezug auf bestimmte psychologische Theorien genauer bestimmt, womit auch die Unterschiede in seinem Verständnis deutlicher werden, vor allem im Blick auf die Auslegung der Subjektrolle des Kindes. Unterschiedliche Bezugstheorien sind vor allem:

- der Konstruktivismus in der Nachfolge Piagets<sup>358</sup>;
- die Theorie des persönlichen (Sach-)Interesses nach Schiefele und Krapp<sup>359</sup>;
- das Konzept metakognitiver Selbststeuerung des Lernens<sup>360</sup>;
- die humanistische Psychologie im Anschluss an C. Rogers<sup>361</sup>;

---

Eigentlich müssten als Kontrollgruppen zum offenen Unterricht Klassen ausgewählt werden, die ausdrücklich das Konzept "direct teaching" repräsentieren (vgl. zur Definition: Dichanz/ Zahorik 1986, Weinert 1996c, 30; zu einer positiven Einschätzung der Befunde: Weinert 1996, 6-8), wobei es theoretisch anregend sein könnte, besonders erfolgreiche Klassen aus beiden Gruppen zu identifizieren und deren Unterricht differenziert zu untersuchen (Giest 1997 mit Bezug auf Einsiedler 1997; Helmke/ Weinert 1997b; Weinert 1997b). In einem solchen Vergleich wäre auch deutlich zu machen, dass bei einer Öffnung von Unterricht nicht um den Verzicht auf Strukturierung, sondern um eine andere Art der Strukturierung mit spezifischen Qualitäten geht (vgl. meinen Versuch einer Systematisierung in diesem Band).

<sup>358</sup> Vgl. Jürgens (1994); Brügelmann ( „Die Öffnung des Unterrichts muss radikaler gedacht...“ i.d.B).

<sup>359</sup> Vgl. Jürgens (1994); Richter (1996); Hartinger (2005b).

<sup>360</sup> Vgl. Beck u. a. (1991; 1995); Nauck (1993b+c).

<sup>361</sup> Vgl. Groddeck (1983); Jürgens (1994).

- das Modell eines kooperativen Lernens mit- und voneinander<sup>362</sup>;
- die Tätigkeitstheorie im Anschluss an Leontjev und Galperin<sup>363</sup>.

Diese Versuche einer theoretischen Fundierung – über bloße Definitionen des eigenen Selbstverständnisses hinaus – lassen sich nach ihren Ausgangspunkten wie folgt systematisieren:

- aus einem veränderten Bildungs- und Wissen(schaft)sbegriff als sich ergänzende Prinzipien entfaltet: institutionelle, thematische und methodische Öffnung<sup>364</sup>;
- pädagogisch-psychologisch ebenfalls als komplementäre Dimensionen gedacht: Offenheit in der Organisationsform, im inhaltlichen Bereich, im kognitiven Bereich, im sozio-emotionalen Bereich, gegenüber der Welt außerhalb der Schule<sup>365</sup>;
- sozialpsychologisch durch Verknüpfung von Beziehungs- und Inhaltsebene<sup>366</sup>;
- nach dem Kriterium zunehmender Selbständigkeit der SchülerInnen als Erweiterung der betroffenen Dimensionen von Unterricht gestuft: methodische, didaktische, pädagogische Öffnung<sup>367</sup>.

Scheidet man Dopplungen aus, so kommt man analytisch auf fünf Grunddimensionen und ihnen korrespondierende Begründungen der Öffnung von Unterricht:

- *persönliche* Öffnung der Lehrerin zu den SchülerInnen, aber auch als bewusste Eröffnung von Beziehungen zwischen diesen, weil Schule nicht mehr vorrangig Ort des Sachlernens, sondern auch des gestalteten sozialen Lernens ist;

<sup>362</sup> Vgl. Brügelmann (1994); Shwalb/ Shwalb (1995).

<sup>363</sup> Vgl. Lompscher (1985; 1997); s. a. den Unterrichtsversuch von Trautmann (1990) und den kritischen Vergleich der kulturhistorischen mit der konstruktivistischen Sicht bei Bauersfeld (1993/95).

<sup>364</sup> Vgl. Ramseger (1977); Benner (1989).

<sup>365</sup> Vgl. Wagner (1978)

<sup>366</sup> Vgl. Bönsch (1979, 37,39). Unter kommunikativem Aspekt sind Kriterien: Gegenseitigkeit, Einfühlungsvermögen, Einwirkungsfähigkeit, Anerkennungsbereitschaft; unter inhaltlichem Aspekt: Durchschaubarkeit, Nachvollziehbarkeit, Veränderbarkeit, Relevanz.

<sup>367</sup> Vgl. dazu Kap. 3 in meinem Beitrag „Die Öffnung des Unterrichts muss radikaler gedacht ...“ in diesem Band; Peschel (2002, Kap. 2.3; 2003, Kap. 2.3).

- *institutionelle* Öffnung der Schule zur Lebenswelt der SchülerInnen hin, weil Lernen nicht mehr im alltäglichen Lebensvollzug aufgehen kann, aber von diesem her und wieder auf diesen hin motiviert werden muss;
- *organisatorische* Öffnung der Arbeitsformen, weil Kinder unterschiedliche Voraussetzungen mitbringen und Lernstile bevorzugen;
- *inhaltliche* Öffnung der Lernwege und Denkmöglichkeiten, weil Lernen nicht passive Nachahmung, sozusagen nur Übernahme von fertigem Wissen sein kann, sondern immer eigenständige (Re-)Konstruktion ist und deshalb prozessuale Erfahrungen und einen Bezug auf die persönliche Bedeutung für den Lernenden (im kognitiven wie emotionalen Sinn) erfordert;
- *politisch-pädagogische* Öffnung der Entscheidungsverfahren im Unterricht, weil Selbstbestimmung nur durch Selbstbestimmung gelernt werden kann und Schule auch ein Raum demokratischen Lebens ist.

Vor allem von LehrerInnen wird in den letzten Jahren zusätzlich als spezifische Dimension eine

- *methodische* Öffnung durch ein Lernen "mit allen Sinnen", also auch über Wahrnehmungs- und Ausdrucksformen neben der (Schrift-)Sprache betont.

Clusteranalysen unserer Befragung von LehrerInnen in vier Schulaufsichtsbezirken zeigen, dass sich deren Vorstellungen zwar nicht sauber einzelnen dieser Dimensionen zuordnen lassen, dass es aber sehr wohl Gruppen mit unterschiedlich akzentuierten Profilen gibt<sup>368</sup>. Ein Vergleich von Aussagen, die 400 LehrerInnen selbst formuliert haben, mit ihren Auswahlantworten auf vorgegebene Alternativen in demselben Fragebogen (Nohn 1997) macht andererseits deutlich, wie schwierig eine solche Einordnung wegen der offensichtlich unscharfen Bedeutung vieler Begriffe in diesem Bereich ist.

**Zwischenfazit:** Zukünftige Untersuchungen oder die Bewertung vorliegender Befunde sollten sich beziehen auf ein explizit formuliertes (Selbst-)Verständnis offenen Unterrichts,

<sup>368</sup> Vgl. dazu Kap. 3.2 in meinem Beitrag „Wie verbreitet...“ in diesem Band.

dieses verorten in Beziehung auf konkurrierende Ansätzen (und ggf. in Abgrenzung von ihnen) und sich um eine genauere Operationalisierung bemühen (s. dazu auch Kap. 3.2 und 3.3).

### 3.2 Unterschiedliche Aspekte offenen Unterrichts werden untersucht

Dies ist kritisch zu bewerten, wenn es ohne Begründung geschieht wie bei Günther (1996, Kap. 4), der z. B. in der Interpretation seiner eigenen Studie von 1988 zwischen den keineswegs deckungsfähigen Begriffen "Freiarbeit" und "Offenem Unterricht" hin- und herwechselt (s. auch oben 3.1).

Andererseits ist die Konzentration auf einzelne Aspekte offenen Unterrichts eine Folge der negativen Erfahrungen aus Versuchen, Unterrichtskonzepte als Ganze zu vergleichen. Nach einer Phase solcher Globalvergleiche<sup>369</sup> hat sich die Forschung um eine Dimensionierung des Ansatzes bemüht und darum, den Einfluss einzelner Bedingungen bzw. Merkmale der Öffnung auf die Entwicklung der Kinder zu erfassen.

Giaconia/ Hedges (1982, 593-4) unterscheiden in ihrer Metaanalyse beispielsweise folgende Bedingungen der Offenheit<sup>370</sup>:

- Selbstbestimmung des Kindes ("role of child in learning")
- Förderdiagnostik ("diagnostic evaluation")
- Vielfalt der Lehr-/ Lernmittel ("materials to manipulate[!?]")
- Individualisierung des Unterrichts ("individualized instruction")
- altersgemischte Lerngruppen ("multiage grouping of students")
- Lehrerteam ("team teaching" ... incl. "use of parents"]
- räumliche Offenheit ("open space").

Sie fanden, dass in "offenen" Klassen folgende Bedingungen positiv (+) oder negativ (-) besonders einflussreich waren...

- im Bereich *Einstellungen* usw.:
  - (+) hohe Selbststeuerung durch die Kinder
  - (+) diagnostische Lernkontrolle mit dem Ziel individueller Passung

<sup>369</sup> Vgl. Peterson (1979; 1980).

<sup>370</sup> Dt. Zusammenfassung hier nach: Scheerer-Neumann (1989, 55).

- (+) vielfältiges Lernmaterial
- (+) individualisierte Unterweisung;
- im Bereich der *Schulleistungen*:
  - (+) gute Materialausstattung für selbständiges Lernen
  - (-) offener Raum.

Diese Befunde lassen allerdings immer noch vielfältigen Auslegungen Raum. So kann "selbständiges Lernen" bedeuten, dass sich das Kind - ohne unmittelbaren Kontakt zur Lehrerin - "selbständig" mit Aufgaben und Materialien auseinandersetzt, die ihrerseits aber von der Lehrerin in der Bearbeitungsform vorstrukturiert und auf eine eindeutige Lösung hin angelegt sind; "selbständiges Lernen" kann aber auch meinen, dass die Kinder eigene Zugänge zu Problemen und Verfahren zu ihrer Lösung erarbeiten, die dann auch zu unterschiedlichen Ergebnissen führen können.

Überdies werden auch von Giaconia/ Hedges spezifische fachdidaktische Ansätze noch gar nicht erfasst (vgl. zu diesem Problem unten Kap. 3.4)

Beobachtete Effekte sind also auf konkretere Merkmale des Unterrichts zu beziehen, als sie in den Metaanalysen fassbar werden. Dazu gibt es eine Reihe von Vorschlägen, die für die Situation in us-amerikanischen Schulen auch schon vor 25 Jahren in Form von Beobachtungsverfahren operationalisiert wurden<sup>371</sup>. Diesen Rastern wird von KritikerInnen<sup>372</sup> allerdings zum Teil bloße Additivität ohne Dimensionierung und Gewichtung vorgeworfen, so dass in den Punktwert sehr unterschiedliche Aspekte gleichwertig eingehen.

Insofern hat Marshall (1981) zu Recht darauf hingewiesen, dass selbst gleiche Werte auf der Offenheits-Skala von Walberg/ Thomas (1972) nicht dasselbe bedeuten müssen, da das Instrument acht verschiedene Faktoren unterstellt, deren Profil auch bei gleichem Summenwert sehr unterschiedlich ausfallen kann. Also darf es nicht verwundern, wenn einmal positive Effekte nur im Leistungsbereich, ein ander Mal nur im Einstellungsbereich festgestellt werden - je nachdem, welcher Faktor (z. B. Vielfältigkeit des Materials vs. ho-

<sup>371</sup> Vgl. Traub u. a. (1972); Walberg/ Thomas (1972); vgl. für die deutsche Schulsituation: Brügelmann u. a. (1997).

<sup>372</sup> Vgl. Marshall (1974); Bennett (1976)

hes Maß an Selbständigkeit) im Einzelfall das Profil qualitativ bestimmt, auch wenn quantitativ unterschiedliche Teilwerte sich zu einem identischen Gesamtwert addieren.

**Zwischenfazit:** Auswertungen vorliegender Studien sollten sich also beziehen auf eine Differenzierung von Merkmalen im Rahmen einer übergreifenden Konzeption. Zukünftige Untersuchungen müssten klären, ob es zwei, drei, fünf, sechs oder gar acht bis zehn unabhängige Dimensionen<sup>373</sup> gibt bzw. welche Beziehungen zwischen ihnen bestehen.

### 3.3 Die realisierte Offenheit wird unterschiedlich klassifiziert

Selbst wenn geklärt ist, welcher Ansatz untersucht werden soll (s. 3.1), heißt das noch nicht, dass dieser Ansatz auch tatsächlich untersucht wird.

Zwei Probleme tauchen auf. Erstens: Werden Versuchsgruppen nach Selbst- oder nach Fremdeinschätzungen gebildet? Zweitens: Werden sie nach vorweg definierten Kriterien oder ex-post durch gleichmäßige Aufteilung der anfallenden Stichprobe (z. B. Halbierung oder Vergleich der Extremdrittel) gebildet?

Empirische Studien sind aufwändig. Deshalb beschränken sie sich oft auf ein Pre-/Post-Test-Design. Dazu werden in der Regel zwei Gruppen gebildet, die unterschiedlichen "Interventionen" ausgesetzt werden. Diese sind - im Experiment - vorweg als Programm definiert, werden bei Feldversuchen meist im Nachhinein durch Befragung der LehrerInnen erfasst. Ob das Programm wie geplant umgesetzt wurde, wie weit sich also die Ansprüche bzw. Wahrnehmungen der LehrerInnen mit dem decken, was in ihrem Unterricht tatsächlich abläuft, muss offen bleiben.

Schon Alltagsbeobachtungen zeigen, dass LehrerInnen mit demselben Lehrgang ganz verschiedenen Unterricht machen. Auch die Lehrplanforschung hat mehrfach gestufte "Veränderungen des Curriculum auf dem Weg vom Autor zum Kind" (Brügelmann 1978) ausfindig gemacht. Dies gilt selbst für sog. "teacher-proof programs" auf Computern<sup>374</sup>.

<sup>373</sup> S. dazu. Weiss u. a. (1972).

<sup>374</sup> S. u.a. Office of Technology Assessment (1988, 77).

Aber nur wenige Studien dokumentieren auch den tatsächlichen Unterricht<sup>375</sup>. Dies wäre besonders wichtig bei einem so auslegungsfähigen Konzept wie dem offenen Unterricht. Ein Beispiel: SchülerInnen können an der Planung von Unterricht beteiligt werden, indem sie gemeinsam, also als Gruppe, einen Wochenplan für alle vereinbaren oder indem sie individuell, sozusagen als Vertrag, mit der Lehrerin absprechen, welche Aufgaben sie sich vornehmen wollen. Sowohl motivational als auch im Sinne der kognitiven Passung von Aufgabe und Entwicklungsstand sind jeweils unterschiedliche Effekte zu erwarten, die nur dann korrekt interpretierbar sind, wenn man die jeweilige Praxis der "Mitwirkung" kennt.

Bussis/ Chittenden (1970, 62) haben solche Unterschiede bereits vor mehr als 25 Jahren als problematisch für vergleichende Evaluationen offenen Unterrichts hervorgehoben, leider ohne erkennbare Wirkung. So belegt Marshall (1981, 187) zehn Jahre später eindrucksvoll, dass in zusammenfassende Analysen sehr unterschiedliche "Realitäten" offenen Unterrichts eingegangen sind: Klassen, die bei Klass/ Hodge (1978) als "offen" eingestuft wurden, erreichen im Klassifikationsschema des Öffnungsgrades von Walberg/ Thomas (1972) Mittelwerte wie die Teil-Stichprobe "traditionell" unterrichteter Klassen bei Evans (1971) - oder sogar niedrigere als bei Groobman u. a. (1976).

**Zwischenfazit:** Punktuelle Vergleiche von Leistungen und Einstellungen sind zu ergänzen durch Prozess- und Kontextstudien, die den Verlauf und die Bedingungen des Unterrichts erfassen. Selbstaussagen über Unterrichtskonzeptionen sind zu kontrollieren durch Fremdbeobachtungen<sup>376</sup>.

### 3.4 Unterschiedliche Lernbereiche/ Fächertraditionen werden einbezogen

Dass die Befunde je nach untersuchtem Fach verschieden ausfallen, kann zwei Gründe haben.

Die Fächer haben von ihrem Zielprofil (stärker sach- vs. person-orientiert; reproduktiv vs. kreativ; Teilfertigkeiten vs. Globalleistungen) und von ihrem methodischen Vorgehen her

<sup>375</sup> Vgl. etwa Bennett u. a. (1976); Helmke u. a. (1988).

<sup>376</sup> Zu den Differenzen zwischen Selbsteinschätzungen und Fremdbeobachtungen vgl. Kap. 3.5 in meinem Beitrag „Wie verbreitet...“ in diesem Band und Hanke (2005, 164).

(stärker lehrgangsgebundene Instruktion vs. entdeckendes Lernen) Traditionen, die den Prinzipien offenen Unterrichts unterschiedlich nah sind.

Kompliziert wird die Situation, wenn die Dimension offen vs. geschlossen überlagert wird durch didaktische oder unterrichtsmethodische Dimensionen. Das trifft beispielsweise im **Lese- und Schreibunterricht** zu. So fand Madamba (1981) in ihrer Metaanalyse von 72 experimentellen Studien keinen Unterschied in den Effekten von offenen vs. traditionellen Ansätzen (u. a. im Leseverstehen, im Wortschatz und im Selbstkonzept). Andererseits belegen Pflaum u. a. (1980) in ihrer Metaanalyse von Lesemethoden Vorteile synthetischer Ansätze.

Nun ist der Spracherfahrungsansatz in angelsächsischen Ländern in der Regel verbunden mit einer ganzheitlichen Leselehre, die Aufgaben zur Lautanalyse und zur Einsicht in die Buchstabenstruktur der Alphabetschrift vernachlässigt<sup>377</sup>. Insofern verwundert es nicht, dass KritikerInnen<sup>378</sup> eine Überlegenheit von Lehrgängen feststellen, die eben zugleich stärker auf "phonics", also synthetisch orientiert sind<sup>379</sup>. Dennoch hat der mit hohen öffentlichen Subventionen entwickelte und zunächst hoch gelobte Phonics-Lehrgang „Reading First“ die Erwartungen nicht erfüllen können<sup>380</sup> und auch eine neuere Metaanalyse von Jeynes (2008) zur Förderung von Minderheiten stellt die früheren Einschätzungen in Frage.

Im deutschsprachigen Raum steht zudem im Zentrum des Spracherfahrungsansatzes das freie Schreiben, das über das "Konstruieren" der Wörter die Aufmerksamkeit der Kinder von Anfang an auf die Beziehung zwischen Lauten und Buchstaben(gruppen) richtet. Offene Konzeptionen zum Schriftspracherwerb im deutschsprachigen Raum sind also sowohl weniger stark an einen Lehrgangsaufbau gebunden als auch geprägt durch die Bedeutung des freien Schreibens schon im ersten Schuljahr. Entsprechend vorsichtig sind amerikanische Befunde zu übertragen, auch wenn sie in der Tendenz eher die offeneren Ansätze begünstigen<sup>381</sup>.

<sup>377</sup> Vgl. Goodman (1967; 1986); Smith (1971; 1978).

<sup>378</sup> Z. B. Chall (1967; 1979; 1989); Liberman/ Liberman (1992); Stanovich/ Stanovich (1995).

<sup>379</sup> Ein zentraler fachmethodischer Faktor, wie die neuere Forschung immer wieder belegt, vgl. u. a. Schneider (1995); Marx (1997).

<sup>380</sup> Vgl. Toppo (2008).

<sup>381</sup> Nach Metaanalyse von Stahl/ Miller (1989) von 180 Studien: 43 = 22% zugunsten des Spracherfahrungsansatzes; 115 = 66% unentschieden; 22 = 12% zugunsten Lehrgang mit Dekodierschwerpunkt; s. die Zusammen-

Auch die deutsche Situation lässt sich nicht einfach auf einen Nenner bringen, weil die Studien sehr unterschiedlich angelegt sind<sup>382</sup>.

Herff (1993) hat für den Regierungsbezirk Köln von LehrerInnen einschätzen lassen, wie viele Kinder ihrer Klassen am Ende des ersten Schuljahres "bisher nicht gelesene, laut-treue Wörter selbständig erlesen" konnten. Entgegen der verbreiteten These vom allgemeinen Leistungsverfall kam sie gegenüber 78% im Jahre 1969/70 für 1988/89 auf 90% (beschränkt auf deutschsprachige SchülerInnen: 93%; a. a. O., 24). In den 0,5% (= 7) Klassen, die nach dem Ansatz "Lesen durch Schreiben" unterrichtet wurden, lagen die deutschsprachigen SchülerInnen auf demselben Niveau (a. a. O., 52). In 2% der Klassen wurde mit einer Eigenfibel (ohne käuflichen Lehrgang) gearbeitet, in weiteren 8% ebenfalls mit einer Eigenfibel und nur teilweise mit einer Fibel. In diesen 128 Klassen (knapp 10% der Stichprobe) schnitten die deutschsprachigen SchülerInnen geringfügig besser ab als in Fibelklassen (94,5% "LeserInnen" vs. 92,7%), während der Vorteil bei Aussiedlerkindern (89,6% vs. 81,6%) und bei ausländischen SchülerInnen (79,3% vs. 75,1%) deutlicher war. Forschungsmethodisch problematisch bei dieser Studie ist das weiche Kriterium und die Erhebung über die Lehrperson selbst. Zudem ist wenig über evtl. Unterschiede in den Teilstichproben und in den jeweiligen Unterrichtsbedingungen bekannt.

Im Blick auf die Rechtschreibung unterstellt Hüttis-Graff (1997) mit ihrer Auswertung von 21 Klassen zumindest für "Lesen durch Schreiben" eine andere Tendenz als Herff<sup>383</sup>. Was sich hinter den Etiketten ("Lesen durch Schreiben", "Eigenfibel") konkret verbirgt, wird allerdings nicht erläutert, die Auswahl der (zudem sehr kleinen) Stichprobe nicht begründet, auch hier die Vergleichbarkeit der Lernvoraussetzungen und Rahmenbedingungen nicht belegt.

Besser abgesichert sind demgegenüber die Ergebnisse einer repräsentativen Studie von Peter May (1995), ebenfalls in Hamburg. Er kommt zwar zunächst zu einem ähnlichen Er-

fassung bei Walter (1996); noch stärker zugunsten des Spracherfahrungsansatzes die dort ebenfalls berichteten Befunde aus 21 Studien, die Gee (1995) ausgewertet, in ihrer forschungsmethodischen Qualität mehrheitlich aber kritisiert hat.

<sup>382</sup> Vgl. zu neueren deutschen Studien, die insgesamt in ein Patt münden, aber mit interessanten Differenzen im Einzelnen: Brügelmann/ Brinkmann (2006) und das Fazit unten.

<sup>383</sup> Vgl. die Tabelle S. 36 und den Kommentar S. 37, insbesondere Anm. 6, in Herff (1997).

gebnis wie Hüttis-Graff: "Im Mittel erzielen Klassen, die mit einer Fibel unterrichtet werden, bessere Rechtschreibleistungen und einen geringeren Anteil von Kindern mit schwachen Rechtschreibleistungen als Klassen, die ohne Fibel unterrichtet werden." (348) Ergänzende Daten zeigen aber, dass die fibellosen Klassen häufiger in einem sozio-ökonomisch belasteten Umfeld arbeiten: "Bei den Klassen, deren außerschulisches Umfeld nach Einschätzung der Schulaufsicht stark bzw. sehr stark belastet ist, zeigen sich keine signifikanten Leistungsunterschiede zwischen Klassen, die mit bzw. ohne Fibel unterrichtet werden." (348)

In dieselbe Richtung deuten die Befunde aus dem "Schreibvergleich BRDDR"<sup>384</sup>: Am Ende der ersten Klassen erzielten Schweizer Klassen, die nach "Lesen durch Schreiben" unterrichtet wurden, sowohl im freien Text als auch im Diktat deutlich bessere Ergebnisse als die westdeutsche Stichprobe (in der sich nur ein Fünftel der LehrerInnen "nicht" oder "kaum" an die Vorgaben des Fibellehrgangs gehalten haben); im Vergleich zu der besonders stark lehrgangsgebundenen Stichprobe "Ost" schnitten die "Lesen durch Schreiben"-Klassen im freien Text besser (62.3% vs. 58.7% der Wörter richtig), im Diktat etwas schwächer ab (77.7% zu 80.7% der Wörter richtig)<sup>385</sup>.

Im Schreibvergleich "BRDDR" waren Ende vierter Klasse die Unterschiede im freien Text zwischen allen drei Gruppen - Stichprobe "West" 90.8% richtig; "Lesen durch Schreiben" Schweiz 91.1% richtig; Stichprobe "Ost" 92.6% richtig - nur noch gering.

Das Problem dieser Erhebung ist ihre kurzfristige Organisation wegen der besonderen historischen Bedingungen. Daraus folgt, dass die Wahl der Stichprobe nicht konsequent nach einem Versuchsplan bestimmt werden konnte, dass die Schreibproben von den LehrerInnen selbst durchgeführt werden mussten und dass kein echter Längsschnitt möglich war. Insbesondere konnte der Unterricht von Klasse 2 bis 4, also nach dem Erstunterricht, nicht eindeutig verschiedenen Konzepten zugeordnet werden.

Allerdings bestätigen andere Untersuchungen zum Vergleich "Ost" und "West" den zentralen Befund, dass im Lesen und Schreiben keine bedeutsamen Leistungsunterschiede gefun-

<sup>384</sup> Vgl. Brügelmann u. a. (1994b, 137).

<sup>385</sup> Bestätigt wird der Vorteil eines selbständigen "Konstruierens" von Wörtern im Anfangsunterricht durch zwei us-amerikanische Studien zum "invented spelling" auf Lese- und Schreibleistungen in der 2. Klasse (Clarke (1988); Foorman u. a. (1991)).

den werden konnten. Das heißt, dass der inhaltlich und methodisch nicht vergleichbar systematisierte und zudem weniger stark auf explizite Instruktion und intensive Übung angelegte Lese- und Schreibunterricht in den westlichen Ländern im Mittel zu keinen Leistungsverlusten geführt hat<sup>386</sup>. Im unteren Leistungsbereich sieht das teilweise etwas anders aus<sup>387</sup>.

Damit lässt sich zum „freien Schreiben“ in den ersten Schuljahren resümieren<sup>388</sup>:

„Betrachtet man die vielfältigen Befunde im Überblick, ergibt sich kein einheitliches Bild. Zwar kann die Befürchtung, freies Schreiben schade der Lese- und (Recht-) Schreibentwicklung, als widerlegt gelten<sup>389</sup>: Nicht nur viele Erfahrungsberichte, sondern auch sorgfältig dokumentierte Studien zeigen, wie stark alle Kinder, auch die leistungsschwachen<sup>390</sup>, vom freien Schreiben profitieren können<sup>391</sup>. In größeren Stichproben fächern sich die Effekte allerdings breit auf - übrigens ein typisches Phänomen für pädagogische Konzepte, wenn sie in die Breite gehen. Insofern lässt sich eine generelle Überlegenheit von Ansätzen, die freiem Schreiben im Anfangsunterricht eine zentrale Funktion einräumen, gegenüber stärker lehrgangs- und normorientierten Ansätzen nicht belegen.

<sup>386</sup> Vgl. Flor u. a. (1992); Lehmann (1993); Schnabel u. a. (1997).

<sup>387</sup> Vgl. May 1991; s. andererseits das besonders günstige Abschneiden der schwachen SchülerInnen in "Lesen durch Schreiben"-Klassen am Ende des 1. Schuljahres in Brügelmann u. a. (1994, 131), und die positiven Befunde aus der Förderung sog. "Risikokinder" bei Winsor/ Pearson (1992). Auch Folgerungen aus neueren Studien, z. B. zum Berliner Projekt „BeLesen“, die zunächst für eine Benachteiligung leistungsschwacher Kinder durch offenere Ansätze in Anspruch genommen wurden, ließen sich bei genaueren Analysen nicht belegen (vgl. die kritische Würdigung von Schröder-Lenzen/ Mücke (2005) in Brügelmann/ Brinkmann (2006, 15-17 und Anhang I); s. zu den bis Klasse 4 vollständig egalisierten Methodeneffekten auch Merckens (2008). S. aktuell aus den USA die bereits erwähnte Metaanalyse von Jeynes (2008) und den Misserfolg des hoch subventionierten Phonics-Lehrgangs „Reading First“ (vgl. Toppo 2008).

<sup>388</sup> Vgl. unseren kritischen Forschungsüberblick in Brügelmann/ Brinkmann (2006), der detailliert aufzeigt, wie schwer die Befunde zum Lese- und Schreibunterricht auf einen Nenner zu bringen sind - selbst bezogen auf ein-und-dieselbe Studie, wenn man etwa verschiedene Zeitpunkte der Wirkungskontrolle vergleicht.

<sup>389</sup> Varianten des Spracherfahrungsansatzes schneiden insgesamt etwas besser ab als die Reinform „Lesen durch Schreiben“, aber Lesen durch Schreiben ist gerade im Anfangsunterricht oft erfolgreicher als Lehrgänge.

<sup>390</sup> Dass dieser Effekt durch eine akustische Rückmeldung - wie in neueren Programmen zu „sprechenden Anlauttabellen“ - gerade für leistungsschwache RechtschreiberInnen noch gesteigert werden kann, hat Dahl (1990) schon vor vielen Jahren experimentell nachgewiesen. Genutzt wird diese Möglichkeit in neueren Formen der „Sprechenden Anlauttabelle“, z. B. in der ABC-Lernlandschaft von Brinkmann u. a. (2008).

<sup>391</sup> Dieser Befund steht in Übereinstimmung mit Feinanalysen der Rechtschreibentwicklung, die zeigen, dass Rechtschreiblernen nicht der Spiegel von Rechtschreibunterricht ist, d. h. dass Rechtschreibung nicht Wort für Wort übernommen, sondern gerade in der Anfangsphase immer wieder neu konstruiert wird. So zeigen die Gruppenvergleiche von May (1995), dass die Gruppen leistungsschwacher und -starker RechtschreiberInnen über vergleichbare Fehlschreibungen zur Normschreibung gelangen - auch unabhängig vom Lehrgang (vgl. May 1991); die individuellen Lernwege innerhalb dieser Gruppen wiederum verlaufen sehr unterschiedlich - auch bei gleichem Unterricht (vgl. Brinkmann 1997, 342ff.; 2003).

Bedenkt man die Vorbemerkung, verwundert das nicht: Unter den allgemeinen Etiketten werden sehr unterschiedliche Konzepte verglichen. Wie diese dann noch von einzelnen LehrerInnen konkret umgesetzt worden sind, ist gar nicht oder nur sehr grob erfasst worden. Die große Streuung *innerhalb* einzelner Ansätze (auch *innerhalb* der einzelnen Untersuchungen) ist der einzige durchgängig sichere Befund."

Für den **Mathematikunterricht** gibt es erst wenige einschlägige Studien<sup>392</sup>, am ehesten noch zum Vergleich inhaltlich und damit auch mathematisch komplexer vs. kleinschrittig aufbereiteter Aufgaben.

Historisch interessant ist ein Feldversuch aus den 1930er Jahren von Benezet (1988). Er verglich 100 Klassen, in denen nach Lehrgang in die Grundrechenarten eingeführt wurde, mit einer gleich großen Stichprobe, die bis zur sechsten Klasse Mathematik an Sachproblemen einführte. Der positive Befund: Nach nur einem Jahr Unterweisung in formellen Rechenverfahren konnte die Experimentalgruppe in dieser Hinsicht zu der Kontrollgruppe aufschließen, war ihr überdies in anderen Leistungen, die durch den häufigen mündlichen und schriftlichen Ausdruck gefordert wurden (Problemdiskussion; Ausdruck; Rechtschreiben), überlegen.

Analog sind die Befunde, die Lingelbach (1995, 82) aus einer Studie von Peterson u.a. (1989) berichtet, nämlich "... dass in einer Gruppe von Expertenlehrkräften, die ihren Unterricht vornehmlich über Textaufgaben gestalteten und wenig Zeit auf das reine Üben von Grundrechenfertigkeiten verwendeten, keine Einbußen in den Grundrechenfähigkeiten der Schüler zu erkennen waren".

Aktuell werden die Befunde aus der TIMS-Studie in der Perspektive diskutiert, ob Mathematikleistungen einen frühen Umgang mit komplexen Aufgaben erfordern, an denen Kinder eigene Strategien zur Lösung von Problemen entwickeln können, die erst allmählich über die Diskussion des Divergenten "vom Singulären zum Regulären führen"<sup>393</sup>. Stern (1997, 164) stellte dazu in der Münchener SCHOLASTIK-Studie fest, dass sowohl leistungsstarke als auch leistungsschwache Klasse dann besonders große Fortschritte machen

<sup>392</sup> Zum Teil werden verwandte Probleme untersucht in den Metaanalysen zu "traditionellem" vs. "modernem" Mathematikunterricht bei Athapily u. a. (1983) und zur Individualisierung bei Horak (1981), allgemein: Walberg u. a. (1984); s. a. den Überblick von Selter (1998).

<sup>393</sup> So Gallin/ Ruf (1990); vgl. auch den Beitrag von Radatz (1998).

ten, wenn ihre LehrerInnen "eine konstruktivistische Auffassung vom Kompetenzerwerb" vertraten:

"Lehrer, die eine konstruktivistische Grundhaltung zum Lernvorgang einnehmen, halten es für günstiger, Schülern Freiheiten zu lassen in der Art und Weise, wie sie Aufgaben lösen. ... Eine rezeptive Grundhaltung geht einher mit der Vorstellung, dass man nur Aufgaben vorgeben darf, für deren Lösung bereits genaue Anweisungen gegeben wurden." (163)

Im **Sachunterricht** sind häufiger Studien zu spezifischeren Fragestellungen durchgeführt worden: Dabei geht es einerseits um Selbsttätigkeit, zum anderen um entdeckendes Lernen als methodische Form der Öffnung von Aufgaben (s. die Zusammenfassung bei Einsiedler 1996). Nach der Metaanalyse von Wise/ Okey sind Programme "entdeckenden Lernens" um .32 Effektstärken herkömmlichem Unterricht überlegen, aber auch "direkte Unterweisung" schneidet noch um .23 Effektstärken besser ab als üblicher Unterricht. Willett u.a. (1983) fanden in ihrer Metaanalyse für selbstgesteuertes Lernen einen Vorteil von nur .08, während stärker lehrerInnen-gesteuerte Programme deutlichere Vorteile erzielten (programmierte Unterweisung .17, "mastery learning" sogar .64). Wieder anders fallen die Befunde des Münsteraner BIQUA-Projekts zu „konstruktivistischen“ Lernumgebungen aus<sup>394</sup>: Diese erwiesen sich herkömmlichem Unterricht überlegen, wobei aber ein stärker sequenziertes Lernangebot mit strukturierenden Hilfen durch die Lehrperson einem freien Materialangebot gegenüber zumindest in den langfristigen Lerneffekten noch einmal überlegen war<sup>395</sup>. Hilfreich wäre eine Differenzierung der Befunde nach Schulstufen und Teilfächern des Sachunterrichts, um diese Widersprüche auflösen zu können.

Für handlungs-zentrierte Programme (ESS, SAPA, SCIS) liegen Befunde aus 57 kontrollierten Untersuchungen vor. Bredderman (1983) berichtet aus seiner Meta-Analyse Vorteile gegenüber traditionellen Programmen, die sich statistisch in Effektstärken von .52 für naturwissenschaftliche Verfahren, .16 für inhaltliches Wissen und .28 für den affektiven Bereich niederschlagen<sup>396</sup>.

<sup>394</sup> Vgl. zu Anlage und Ergebnissen des Projekts Möller u. a. (2002); Hardy u. a. (2006).

<sup>395</sup> Peschel (2003, Kap. 8.1) kritisiert an diese Versuchen, dass es sich um kurzfristige und (thematisch) immer noch stark fokussierte Interventionen handelt, die nicht als Beispiel für einen offenen Unterricht herhalten könnten, in dem die Kinder Zeit haben eigene Interessen zu entwickeln und diesen zu selbst gewählter Zeit nachzugehen.

<sup>396</sup> Vgl. ergänzend zum *inquiry learning* im naturwissenschaftlichen Unterricht die Metanalysen von El-Nemr (1980); Lott (1983); Sweitzer/ Anderson (1983).

Auch wenn diese Untersuchungen nur bestimmte fachliche Ausschnitte der Öffnung des Unterrichts erfassen, wird doch erkennbar, dass es nicht reicht, Unterricht nach allgemein-didaktischen bzw. -pädagogischen Kriterien zu unterscheiden.

Kaum in die Diskussion einbezogen werden Befunde zum **sozialen und politischen Lernen**, die sich aus Unterrichtsarrangements ergeben, in denen SchülerInnen mehr Mitsprache haben. In unserer Auswertung verschiedener Studien<sup>397</sup> zeigte sich ein durchaus bedeutender Einfluss auf die Entwicklung demokratischer Einstellungen und Verhaltensweisen.

Positive Effekte sind in ganz verschiedenen Dimensionen nachgewiesen, wobei die meisten Projekte allerdings im Sekundar- oder Hochschulbereich durchgeführt worden sind<sup>398</sup>:

- Abbau von Vorurteilen,
- positivere Wahrnehmung zwischen sozialen oder ethnischen Gruppen,
- gesteigertes Verantwortungsbewusstsein,
- höheres Selbstwertgefühl und höhere Selbstwirksamkeit,
- verbesserte soziale Kompetenzen,
- höhere Sensibilität für Probleme in der Gemeinde und
- eine ausgeprägtere politische Identität.

Fatke/Schneider (2005) untersuchten in ihrer Evaluationsstudie zur Initiative „mitWirkung“ die Beteiligung von Kinder und Jugendlichen in Kommunen. Sie stellten ebenfalls eine umso höhere Beteiligung an Projekten am Wohnort fest, je höher die Partizipationsintensität in der Schule war.

Solche Erfahrungen können nicht nur in außerschulischen Projekten (als „community service“) gewonnen werden, wie Ohlmeier (2006) in einer qualitativen Fallstudie ausgewählter Klassenkonferenzen in einer Grundschule gezeigt hat. Er stellt ein zunehmendes Spektrum an sozialen und politischen Erfahrungsmöglichkeiten fest, je mehr sich demokratische Verhandlungsverfahren etabliert haben. Insbesondere erwerben die GrundschülerInnen soziale Kompetenzen, die für eine zivilgesellschaftlich ausgerichtete, politische Demokratie von Bedeutung sind<sup>399</sup>.

<sup>397</sup> Vgl. die Auswertung bei Brügelmann (2008c).

<sup>398</sup> Vgl. die Zusammenfassung von Evaluationsstudien in den USA bei Sliwka (2004, 5, 10-11).

<sup>399</sup> Vgl. Ohlmeier (2007, 70).

Auch für den Sekundarbereich folgern Torney-Purta/Barber (2005) aus den IEA-Studien der 1970er Jahre<sup>400</sup>, dass nicht allein die Vermittlung von Wissen über demokratische Prozesse, sondern auch deren Praktizierung im Schulalltag wichtig ist, damit SchülerInnen Zuversicht entwickeln, dass ihre Handlungen etwas bewirken können: Das Klassenklima erklärt Unterschiede sowohl im erfragten Wissen als auch im Blick auf die Partizipation an politischen Prozessen außerhalb der Schule.

**Zwischenfazit:** Je nach Lernbereich lassen sich positive Wirkungen einer Öffnung des Unterrichts in verschiedenen Dimensionen feststellen. Unterschiedliche Befunde aus Studien in den einzelnen Lernbereichen verweisen aber auf die Bedeutung fachdidaktischer Annahmen der jeweiligen Programme<sup>401</sup>. Diese relativieren die Generalisierbarkeit von Aussagen über das pädagogische Konzept. Nach Möglichkeit sollten Untersuchungen Aktivitäten und Leistungskontrollen in verschiedenen Lernbereichen einbeziehen, um fachspezifische Differenzen zu erfassen.

### 3.5 Die Wirkungen werden anhand unterschiedlicher Kriterien beurteilt

Bereits in den Pauschalbefunden wurden Unterschiede zwischen dem Leistungs- und dem Persönlichkeitsbereich deutlich (s. oben Kap. 2.). Hinzu kommen unterschiedliche Anforderungen in den einzelnen Bereichen (z. B. Worterkennen vs. Textverständnis beim Lesen). Wird "Leseleistung" als "rasches Benennen von Wörtern" operationalisiert, sind im Blick auf die Bedeutung offener Unterrichtsformen andere Ergebnisse denkbar als bei der inhaltlichen Beantwortung von Fragen zu einem Text oder bei der Angabe privater Lesezeit pro Woche. Je nach den verwendeten Test-, Beobachtungs- oder Einschätzungsverfahren differenzieren sich die Ergebnisse auch für dieselben Kriterien weiter aus.

Vorweg bedeutsam ist die Unterscheidung zwischen Prozess- Merkmalen (z. B. selbständiges Arbeiten im Unterricht) und Produkt-Merkmalen (z. B. Rechenleistungen am Ende einer Einheit).

Als *Prozess-Merkmale* bezeichnen wir nicht allgemein Aspekte des Unterrichts, sondern speziell Verhaltensweisen von Kindern im Unterricht. Sie werden in einigen Untersuchun-

<sup>400</sup> Vgl. die Zusammenfassung in Torney u. a. (1975).

<sup>401</sup> Zu einer ähnlichen Einschätzung kommt auch Lipowsky (2007).

gen gesondert als dessen Primär-Effekte erfasst<sup>402</sup>, können aber auch als unmittelbare Reaktion auf die Unterrichtsform und damit als vermittelnde Variable zwischen Lernangebot und Lerneffekten (im engeren Sinn) interpretiert werden.

Günther (1988; 1996) kritisiert generell, dass in den von ihm untersuchten Klassen die Ansprüche offenen Unterrichts nicht eingelöst worden seien: 80% der SchülerInnen beschäftigten sich mit vorprogrammiertem Unterrichtsmaterial; 30% erschienen als "nicht ausgeglichen und fröhlich"; 40% verrichteten ihre Arbeit "verdrossen und freudlos"; 60% hatten Probleme, "wie sie sich angemessen durchsetzen sollten", etwa 50%, "ihre Lerninitiative gegen ihren Hang zur Passivität zu entwickeln".

Diese Einschätzungen sind insofern nicht sinnvoll zu bewerten, als zum einen die Stichprobe unklar definiert ist (s. oben 1.) und als es zweitens keine Vergleichsdaten gibt: weder zum Verhalten dieser Kinder in anderen Unterrichtsformen, noch aus Klassen, die nach einem anderen Konzept arbeiten<sup>403</sup>.

Inhaltlich in eine ähnliche Richtung weisen aber Befunde von Bennett u. a. (1976) in England, dass SchülerInnen im formellen Unterricht häufiger an Aufgaben arbeiten; dies gelte verstärkt für besonders leistungsstarke bzw. leistungsschwache Kinder. Andererseits stellte Bennett im informellen Unterricht mehr (durchaus aufgabenbezogene) Gespräche fest<sup>404</sup>.

Positiv ist demgegenüber das Fazit von Wagner/ Schöll (1992, 48) aus ihrer systematischen Beobachtung in offenen Lernphasen einer 4. Grundschulklasse:

"Während eines Beobachtungszeitraums von 30 Minuten arbeitet ein Schüler im Durchschnitt 22,8 Minuten selbständig, 1,8 Minuten unselbständig und war 5,4 Minuten mit nicht-arbeitsbezogenen Dingen beschäftigt. Leistungsstarke ... 24 Minuten selbständig, leistungsschwache ... immerhin noch 21 Minuten."

Während auch diese Studie unter einem fehlenden Vergleichsmaßstab (Kontrollgruppe) leidet, untersuchten Laus/ Schöll (1995) die Aufmerksamkeit von Kindern einer Klasse im

<sup>402</sup> Also als "Produkte" im weiteren Sinn, s. dazu unten.

<sup>403</sup> S. zu diesem Problem unten (3.8.).

<sup>404</sup> S. a. Bennett (1979, 126f.).

Kontrast von zwei verschiedenen Unterrichtsformen: Alle Leistungsgruppen zeigten in offenen Situationen eine höhere Aufmerksamkeit und längere Konzentration, dies galt besonders für leistungsschwache SchülerInnen.

Dieser Befund bestätigt das Fazit aus der 1 1/2-jährigen Beobachtung einer 3./4. Klasse durch Wagner (1978, 58), dass die SchülerInnen in offenen Situationen (zunächst 1-2, später 6-10 von 17 Stunden pro Woche) ein "deutlich größeres Engagement und eine höhere Arbeitsintensität zeigen als im normalen Frontalunterricht." Ergänzend konnte Hartinger (1997; 1998) im Vergleich von drei Klassen zeigen, dass sich in offenem Unterricht inhaltsbezogene Interessen auch auf Dauer stärker entwickeln. Auch die vielfältigen Interessen und das breite Engagement der Kinder im Längsschnitt von Peschel (2003) belegen das Potenzial offenen Unterrichts für die Aktivierung von SchülerInnen.

Im Blick auf "Produkt"-Merkmale des Unterrichts wird gegen die vorliegenden Vergleichsuntersuchungen schon seit langem eingewandt,

- dass die gängigen Instrumente eher reproduktive als produktive Kompetenzen, eher Teilleistungen als Grundqualifikationen, eher fachliche Leistungen als Persönlichkeitsmerkmale erfassen und damit wesentliche Ziele offenen Unterrichts unterbewerten (Meier 1973);
- dass schon die Testsituation selbst im Widerspruch zu den Zielen offenen Unterrichts stehe und eher den Erfahrungen von Kindern aus traditionellen Klassen entspreche, offenen Unterricht damit zusätzlich benachteilige (DeRivera 1973).

Die (fehlende) Passung von Messverfahren ist einerseits ein "Totschlag"-Argument, mit dem sich unliebsame Befunde immer abwehren lassen. Andererseits ist nicht zu übersehen, dass diejenigen Leistungen leichter zu messen sind, die im traditionellen Unterricht einen vergleichsweise höheren Stellenwert haben. Von daher stellt sich der Forschung die Aufgabe, Verfahren zu entwickeln, die einerseits *inhaltlich* Leistungen erfassen, die im Rahmen offenen Unterrichts im Vordergrund stehen, andererseits dies in *Formen* zu tun, die mit den Arbeitsformen offenen Unterrichts verträglich sind<sup>405</sup>.

<sup>405</sup> So schon Bussis/ Chittenden (1970); dass dies ein generelles Problem bei der Evaluation alternativer didaktischer Ansätze ist, zeigt erneut die Diskussion um "situiertes Lernen", vgl. etwa Renkl (1996); s. zusam-

Nachdenklich stimmt auch die sehr lockere Beziehung zwischen konventionell gemessenen Unterrichtseffekten und verschiedenen Indikatoren für Erfolg im späteren Leben (z.B. Einkommen, Selbsteinschätzung der Zufriedenheit, Selbst- und Fremdeinschätzung beruflicher Effektivität)<sup>406</sup>.

Im Blick auf die Erfolgskriterien kommen weitere Probleme hinzu. So sind bei der Bewertung von Trends, wie sie in Kap. 2 für Fachleistungen vs. Persönlichkeitsentwicklung berichtet wurde, auch bei eindeutiger Differenz der Mittelwerte immer noch drei Fragen umstritten, von denen die ersten beiden mögliche Zielkonflikte der Grundschule beschreiben<sup>407</sup>:

- Gibt es Unterrichtsformen, die es erlauben, spezifische fachliche Kompetenzen und allgemeine Aspekte der Persönlichkeit wie Selbstständigkeit, Selbstkonzept, Kooperationsfähigkeit gleichermaßen zu fördern<sup>408</sup>?
- Gibt es Unterrichtsformen, die sowohl die Durchschnittsleistung steigern als auch die Leistungsunterschiede innerhalb der Klassen mindern<sup>409</sup>?
- Sind die unmittelbar gemessenen Effekte gute Indikatoren für langfristige Veränderungen?

Gegen die letzte Annahme sprechen Erfahrungen vor allem aus kompensatorischen Förderprogrammen im Vorschulbereich, dass einerseits IQ-Gewinne, die vor allem strukturierte Lehrgänge auszeichnen, in kurzer Zeit abnehmen<sup>410</sup>, dass andererseits erst verzögert sichtbar werdende sog. "sleepers effects" (als Folge veränderter Einstellungen und Grundqualifikationen, nicht als Addition einzelner Fertigkeiten und Kenntnisse) eher in offenen

---

menfassend Brügelmann (2007a+c)

<sup>406</sup> Samson u. a. (1984) kommen im Durchschnitt von 33 Studien nach 1949 auf eine Korrelation von .16; vgl. die Zusammenfassungen bei Fraser u. a. (1987, 162) und zur geringen Prognosekraft von Noten auf den späteren Schul-, Studien und Berufserfolg Arbeitsgruppe Primarstufe (2006).

<sup>407</sup> Ausführlicher dazu Petillon (1997).

<sup>408</sup> Eher nein: Giaconia/ Hedges (1982, 599); Arbinger/ Saldern (1984); möglich: Helmke/ Schrader (1990); Gruehn (1995); Schrader u. a. (1997);

<sup>409</sup> *Nein* für den Sekundarbereich: Treiber/ Weinert (1985); dazu methoden-kritisch: Beck u. a. (1988a+b); *in Ausnahmefällen*: Baumert u. a. (1986a+b); Helmke (1988); *ja* für die Grundschule: Treinies/ Einsiedler (1996); vgl. auch die kritische Diskussion der forschungsmethodischen Probleme bei Schrader u. a. (1997) und Baumert (1997).

<sup>410</sup> Vgl. Miller/ Dyer (1975).

Programmen entstehen oder beiden Förderkonzepten (im Vergleich mit ganz fehlender Frühförderung) gemeinsam sind<sup>411</sup>.

Insofern wird das "Baustein"-Modell des Unterrichts ("Was einmal gelernt wurde, bleibt als Element intakt erhalten, auf dem weiterer Unterricht kumulativ aufbauen kann") weder der Alltagserfahrung noch den Befunden zur Umorganisation früheren Wissens durch neue Erfahrung gerecht<sup>412</sup>. Also müssten für aussagekräftige Urteile auch längerfristige Wirkungen von unterschiedlichen didaktischen Konzeptionen untersucht werden, was aber erhebliche Probleme der Forschungsorganisation und der Zurechenbarkeit von Effekten zu bestimmten (dann kaum mehr kontrollierbaren) Bedingungen oder Maßnahmen aufwirft.

Unabhängig von diesen empirischen Fragen stellt sich die immer wieder kontrovers diskutierte normative Frage, wie die konkurrierenden Gesichtspunkte bei der Beurteilung von Befunden jeweils gewichtet werden sollen.

**Zwischenfazit:** Bei der Auswertung von vorliegenden Studien ist zu spezifizieren, welche Effekte untersucht wurden. Zukünftige Studien sollten ein breites Spektrum an Leistungen und unterschiedliche Niveaus der diskutierten Leistungen erfassen; sie sollten nicht nur das Niveau der Mittelwerte, sondern auch die Breite der Leistungsstreuung im Auge behalten und über unmittelbare Lerneffekte auch längerfristige Wirkungen zu erfassen suchen. Zum Verständnis der tatsächlichen Aktivitäten im offenen Unterricht sind neben kontrollierten Wirkungsstudien allerdings auch qualitative Prozess-Untersuchungen notwendig. Wie sich z.B. diese Konzeption im Schulalltag auf verschiedene Kinder auswirkt, haben Garlichs u. a. (1990) eindrucksvoll in ihrer sorgfältigen Beobachtungsstudie dokumentiert.

### 3.6 Die Stichproben werden aus ganz unterschiedlichen Lernergruppen gezogen

In dieser Hinsicht sind drei Merkmale von besonderer Bedeutung

---

<sup>411</sup> Vgl. Rauer/ Valtin (1985, 244-5) und die Beiträge in Brown (1978).

<sup>412</sup> Vgl. zur Illustration nicht-linearer Lernprozesse die Rechtschreibstudie von Brinkmann (2003).

- *Alter* der einbezogenen SchülerInnen (z. B. Kindergarten vs. Oberstufe);
- *Leistungsniveau* (z. B. Grund- vs. SonderschülerInnen) und Persönlichkeitsmerkmale (z. B. ängstliche vs. selbstsichere Kinder);
- *Herkunftsmilieu* bzw. Lebenswelt der Population (z. B. Unter- vs. Mittelschicht; Stadt vs. Land).

Unter dem *Alters*-Aspekt sind gegenläufige Hypothesen möglich. So lässt sich vermuten, dass ältere SchülerInnen die in offenen Situationen angebotene (bzw. geforderte) Selbstständigkeit eher erbringen können als jüngere Kinder. Andererseits erschwert die mit den Schulstufen zunehmende Fach- und Prüfungsorientierung eine Öffnung des Unterrichts.

Für Kleinkinder stellten B. Palmer/ Siegel (1977)<sup>413</sup> fest, dass die freie Wahl der Aufgaben genauso effektiv war wie die Zuweisung nach Entwicklungsstand.

Andere Studien wurden auf unterschiedlichen Schulstufen durchgeführt, z. T. in der Sekundarstufe I<sup>414</sup> bzw. Sekundarstufe II oder sogar mit StudentInnen und zunehmend auch in der Primarstufe<sup>415</sup>. Eine zusammenfassende Auswertung der Wirkungen über die Schulstufen hinweg würde eine Gleichartigkeit der Arbeitsformen offenen Unterrichts unterstellen, die schon der Alltagserfahrung widerspricht.

Dies kann schon deshalb nicht sein, weil die Altersstufen unterschiedlichen Formen der Unterrichtsorganisation korrespondieren: Fachunterricht in der Sekundarstufe schafft ganz andere Rahmenbedingungen für eine Öffnung als das Klassenlehrerprinzip in der Grundschule oder die Unabhängigkeit des Kindergartenvormittags vom 45-Minuten-Rhythmus der Schulstunden. Auch die größere Heterogenität der Primar- vs. Sekundarstufe lässt andere Befunde erwarten<sup>416</sup>. Eine Untersuchung, die ausdrücklich dieselben Arbeitsformen in verschiedenen Altersgruppen verglichen hat (Beck u. a. 1991; 1995), stellte z.B. fest, dass deren Effekte je nach Schulstufe unterschiedlich (stark) sind.

<sup>413</sup> Zit. nach: Beller (1987, 813).

<sup>414</sup> Z. B. 5. Klasse: Helmke/ Schrader (1990); 7. Klasse: Gruehn (1995)

<sup>415</sup> Vgl. Günther (1988), Hanke (2001; 2005) Peschel (2003), Hardy u. a. (2006).

<sup>416</sup> Vgl. das oben genannte Beispiel: Treinies/ Einsiedler 1996 vs. Treiber/ Weinert 1985).

Stahl/ Miller (1989) fanden, dass "whole-language"-Programme lehrgangsorientiertem Unterricht zwar bei Vorschulkindern, aber nicht mehr bei Erstklässlern überlegen waren. Ähnliche Befunde berichtet Walter (1996) aus einer Zusammenfassung verschiedener Studien in den USA: Während im Vorschulbereich der (dort: ganzheitliche) Spracherfahrungsansatz erfolgreicher war, ließen sich in der ersten Klasse kaum mehr Unterschiede zu (stärker synthetisch ausgerichteten) Lehrgängen feststellen.

Weitere bedeutsame Faktoren sind Leistungs-, Persönlichkeits- und Schichtmerkmale. In der Tendenz wird (übrigens auch von vielen LehrerInnen) argumentiert, dass leistungsschwache und sozial unterprivilegierte SchülerInnen "mehr Struktur" (als im offenen Unterricht) brauchen, leistungsstarke Kinder, vor allem aus der Mittelschicht, dagegen von dem gewährten Freiraum besonders profitieren. Da andererseits gerade in Konzeptionen offenen Unterrichts differenzierte Formen der Strukturierung entwickelt worden sind<sup>417</sup>, überrascht es nicht, dass die Befunde nicht so eindeutig sind.

Bei einer Differenzierung nach *Leistungsniveau* werden negative Effekte für leistungsschwache SchülerInnen berichtet von Ausubel (1974), Simons u. a. (1975); Baumert u. a. (1986a+b) - speziell für Jungen dieser Gruppe dagegen Vorteile bei Bennett u. a. (1976) und generell für den Einstellungsbereich positive Wirkungen von Burg (1987) und von Trudewind u. a. (1979). Generell positive Effekte berichtet Peschel (2003, Kap. 13-15) aus seinem Längsschnitt seiner radikal geöffneten Klasse<sup>418</sup>.

Auch für Kinder mit bestimmten Verhaltensproblemen oder *Persönlichkeitsmerkmalen* werden unterschiedliche Wirkungen einer offenen Situation angenommen<sup>419</sup>. Insbesondere wird bei ängstlichen Kindern eine Überforderung durch wenig strukturierte Freiräume befürchtet; andererseits werden positive Wirkungen der Öffnung von Unterricht auf hyperaktive Kinder berichtet<sup>420</sup>.

<sup>417</sup> Vgl. dazu Kap. 4 und 5 in meinem Beitrag „Die Öffnung des Unterrichts muss radikaler gedacht werden...“ i.d.B.

<sup>418</sup> Vgl. speziell zur „wider Erwarten“ positiven Entwicklung leistungsschwacher Kinder a. a. O., Kap. 16 und 17.

<sup>419</sup> vgl. die erhellende Fallstudie von Wittoch (1991)

<sup>420</sup> Vgl: Flynn/ Rapoport (1976); Goetze/ Jäger (1991); Goetze (1992).

Häufig werden negative Effekte vor allem auf Unterschichtkinder erwartet, die durch die Anforderungen der Offenheit an ihre Fähigkeit zur Selbstorganisation überfordert würden<sup>421</sup>. Speziell für diese Gruppe berichten aber Gilkeson u. a. (1981) positive Wirkungen, Traub/ Weiss (1974) fanden keine Unterschiede. Speziell für Effekte auf die Einstellungen fand Burg (1987) positive Effekte, Traub/ Weiss (1974) ebenso für Selbständigkeit, dagegen nicht für andere Einstellungsbereiche.

Die Unübersichtlichkeit der Ergebnisse - selbst bei dieser Differenzierung von Untergruppen - mag daran liegen, dass diese Gruppen immer noch zu grob gegliedert sind. Dafür sprechen die Unterschiede, die Solomon/ Kendall (1976)<sup>422</sup> fanden, als sie Kinder differenzierter nach "Grundeinstellungen, Motive[n] und Schulleistungen" zu sechs Typen gruppierten.

Unterschiede in der Einschätzung offenen Unterrichts können also damit zusammenhängen, dass einzelne Stichproben aus dem Rahmen fallen oder dass sie für den Vergleich mit anderen Studien nicht zureichend definiert sind. Ein Beispiel: Erhebungen in Sonderschulen stehen bei der Interpretation von Wirkungen offenen Unterrichts vor dem Problem, ob hier wegen der schwächeren Lernvoraussetzungen und ungünstigeren Motivationslage bei jeder Veränderung des Unterrichts ein besonders starker Lernzuwachs zu erwarten oder eher Einschränkungen für die Entwicklung der Kinder (im Vergleich zu Regelschulen) zu unterstellen sind.

**Zwischenfazit:** Die vorliegenden Studien sind nach der Zusammensetzung der Stichproben differenziert zu bewerten. Zukünftige Studien sollten die erhobenen Effekte möglichst aufgeschlüsselt nach Leistungs- und Schichtmerkmalen berichten. Für den Grundschulbereich empfiehlt sich Vorsicht bei der Argumentation mit Befunden aus dem Vorschul- bzw. dem Sekundarbereich. Zukünftige Studien sollten diesen Faktor bewusst variieren.

### 3.7 Die Untersuchungen werden in unterschiedlichen kulturellen Kontexten durchgeführt

<sup>421</sup> Vgl. etwa Sertl (2007).

<sup>422</sup> Zit. nach: Wagner/ Schöll (1992, 11).

Kulturelle Kontexte sind bedeutsam für die Wirkung schulischer Maßnahmen<sup>423</sup>. Sie variieren in zwei Dimensionen:

- Befunde aus Studien im einen Schulsystem können nicht ohne weiteres auf ein anderes übertragen werden.
- Auch innerhalb einer Gesellschaft und ihres Bildungssystems gibt es bedeutsame Veränderungen über die Zeit hinweg, so dass Befunde aus den 1930er, 1970er oder selbst den 1990er Jahren nicht auf analoge Formen des Unterrichts in 2010 zutreffen müssen<sup>424</sup>.

Die Frage, in welchem Schulsystem offener Unterricht untersucht wurde, ist wichtig, weil nicht selbstverständlich angenommen werden kann, dass Befunde über verschiedene Kulturen hinweg gelten. In jedem Schulsystem gibt es unterschwellige Unterrichtstraditionen (wie z. B. das lediglich reproduktive Abschreiben themenspezifischer Ausschnitte aus Büchern in häufigen Realisierungen des englischen "Project Work") und organisatorische Rahmenbedingungen wie z. B. unterschiedliche Auslesesysteme, die beeinflussen, wie sich eine Unterrichtsform entfalten kann, wie sie von den Beteiligten wahrgenommen und ggf. uminterpretiert wird.

Am Erziehungsstil von Eltern ist der Einfluss solcher Randbedingungen untersucht worden:

"Wärme als Beziehungsdimension steht zum kulturellen Kontext in Beziehung (Rohner 1986), so dass streng kontrollierendes und disziplinierendes Verhalten z. B. in asiatischen Kulturen eher als Sorge (und damit doch als warme Beziehung) interpretiert wird und nicht als Ablehnung" (Oerter 1995, 116).

<sup>423</sup> Vgl. das ökologische Modell von Bronfenbrenner (1974; 1981)

<sup>424</sup> Vgl. zur Kontextabhängigkeit sozialwissenschaftlicher Befunde generell meine Beispiele in Brügelmann (2005c). Solche Differenzen können auch mit Unterschieden zwischen Pionier- und Alltagssituation zusammenhängen, vgl. etwa die Feststellung von Walter (1996), dass die Überlegenheit des Spracherfahrungsansatzes von früheren zu neueren Untersuchungen hin abnimmt.

Viele der berichteten Untersuchungen zum offenen Unterricht stammen aus dem angelsächsischen Raum. Dort gibt es andere Traditionen und entsprechend auch andere "Kräftefelder", in denen sich offene Formen entwickeln und abzugrenzen haben<sup>425</sup>.

Wer englische "primary schools" besucht hat, weiß, dass "informal classrooms" sich nicht selten durch ein zwar individualisiertes, aber mehr oder weniger durchorganisiertes Abarbeiten vorgegebener Arbeitsblätter von einem Frontalunterricht abheben<sup>426</sup>. Die Offenheit des Unterrichts beschränkt sich also oft auf die eine formale "Dezentralisierung". Qualitativ bedeutet dies, dass Kinder möglicherweise nur etwas a) Vorgeschriebenes und b) Stumpfsinniges (Abkopieren) tun, so dass Öffnung auf die methodisch-organisatorische Ebene begrenzt bleibt. Ergebnisse eines solchen Unterrichts können deshalb für eine Bewertung anspruchsvollerer Formen der Öffnung nur mit Vorbehalt herangezogen werden kann.

Es ist ebenso problematisch, wenn Befunde aus Studien vor 20 oder 40 Jahren ohne weitere Überlegung auf die heutige Zeit übertragen werden. Zu welcher Zeit eine Untersuchung durchgeführt wurde, wird nur selten thematisiert. Doch schon elementare äußere Bedingungen wie Klassengrößen haben sich in wenigen Jahrzehnten deutlich verändert (zunächst Abnahme von über 30 in den 1950er und 1960er Jahren hin zu 20 Kindern in den 1970er und 1980er Jahren, mit einem erneuten Anstieg in den 1990er Jahren hin zu häufig 25-30 Kindern).

Noch bedeutsamer können Veränderungen in den Werten und Denkweisen von Eltern und LehrerInnen, im "pädagogischen Klima" einer Zeit sein.

Am Beispiel elterlicher Erziehungsstile hat Bronfenbrenner (1985) das Problem verdeutlicht<sup>427</sup>:

---

<sup>425</sup> Vgl. zum Einfluss fachdidaktischer Traditionen oben Kap. 3.4.

<sup>426</sup> So auch Bennett u. a. (1984, 3, 215: "Children predominantly work on their own tasks assigned and structured by their teachers. [...] From the teachers' point of view, children were busy, and busy work was equated with appropriate demands"; s. deshalb auch die Vorbehalte von Kasper (1992) gegen die Nutzung der Bennett-Studie, s. oben Kap. 3.0.

<sup>427</sup> Auf grundsätzliche Unzulänglichkeiten dieser Vergleichsstudien zum elterlichen Erziehungsverhalten wegen der fehlenden Berücksichtigung von Ansätzen partnerschaftlicher Erziehung hat Schimpke (2007;2008) hingewiesen

"In verschiedenen Jahrzehnten wirkte sich permissives Erziehungsverhalten der Eltern ganz unterschiedlich aus. In den dreißiger und noch vierziger Jahren zeigte sich ein eher positiver Einfluss auf die Entwicklung der Kinder und Jugendlichen, weil das außerfamiliäre Stützsystem intakt und stabil war.

[...] Praktisch alle Untersuchungen über Risikofaktoren im Jugend[!]alter zeigen, dass elterliches Kontrollverhalten positive Auswirkungen hat. [...]

Weiterhin zeigt sich in Längsschnittstudien die Stützfunktion anderer gesellschaftlicher Einrichtungen, wie Kirche und Militärdienst. [...] Dies liegt weniger an den Inhalten dieser Institutionen als vielmehr an deren Stützfunktion, die sie der Risiko[!]gruppe in einer Epoche der Desorientierung und Gefährdung gewährt (Untersuchungen hierzu s. Werner, 1990)." (Oerter 1995, 116-117)

Dieses Beispiel illustriert nicht nur den geäußerten Vorbehalt, sondern auch die erforderliche Relativierung bei der Darstellung und Interpretation von Befunden: Die Aussagen sind redlich auf eine bestimmte Altersphase und auf eine Teilgruppe eingeschränkt (s. oben Kap. 3.6).

An der zunehmenden Verbreitung freien Schreibens im Anfangsunterricht seit Ende der 80er Jahre läßt sich zeigen, wie stark sich bereits in 5-10 Jahren Unterrichtsformen verändern.

Zum einen hat das Konsequenzen für die Einschätzung der quantitativen Anteile offener Unterrichtsformen. Dies ist bedeutsam für die Frage, ob (noch) für eine stärkere Öffnung gestritten oder einer (evtl. schon) breiten Verwendung und damit einer selbstverständlich scheinenden *political correctness* gegengesteuert werden müsse<sup>428</sup>.

Zum zweiten bedeutet eine unterschiedliche Verbreitung, dass es zu verschiedenen Zeiten andere Gruppen von LehrerInnen sind, die "dieselbe" Unterrichtsform praktizieren. Das kann die Art und die Qualität dieser Unterrichtsform verändern.

Beispielsweise können Schwierigkeiten offener Unterrichtsformen bei weniger engagierten und kompetenten LehrerInnen der zweiten oder dritten "Mitläufer"-Generation zunehmen.

---

<sup>428</sup> So Dierkes (1994).

Es wird gelegentlich berichtet, dass neue Methoden (z. B. in der Psychotherapie) durch einen "Pionier"-Effekt anfänglich große Erfolge haben, die aber mit zunehmender Verbreitung verschwinden<sup>429</sup>.

Dass für die Ausprägung von Offenheit auch die Kontexte auf der Mikroebene bedeutsam sind, zeigen die Befunde von Köttl/ Sauer (1980): Der Grad der sozialen Direktivität von LehrerInnen variiert je nach dem sozialen Klima in Klassen. Die Haltung und die Kompetenz von LehrerInnen entscheiden also nicht allein, wieviel Aktionsraum sie ihren SchülerInnen gewähren.

**Zwischenfazit:** Die Auswertung von Studien aus anderen Ländern sollte zumindest heute schon bekannte Einflussfaktoren prüfen, die ihre Übertragbarkeit einschränken könnten. Studien aus früheren Jahrzehnten sollten sorgfältig daraufhin geprüft werden, ob sich bedeutsame Bedingungen verändert haben könnten, so dass heute andere Ergebnisse zu erwarten sind. Insgesamt sollten die Kontexte von Unterricht bei dessen Einschätzung stärker berücksichtigt werden<sup>430</sup>.

### 3.8 Die Stichproben sind zahlenmäßig sehr klein und/oder sie werden nicht durch eine Kontrollgruppe ergänzt

Gerade im deutschsprachigen Bereich beschränkt sich eine Reihe von Untersuchungen auf einzelne Klassen<sup>431</sup>. Dieses Verfahren weist bedeutsame Vorteile auf, z. B. was die Intensität der Beobachtung und die Breite der erfassten Merkmale betrifft<sup>432</sup>. Es schränkt andererseits den Geltungsbereich der Aussagen zunächst auf die beobachtete Gruppe ein. Ihre Überzeugungskraft leidet weiter, wenn keine Kontroll- oder zumindest Vergleichsgruppen einbezogen werden.

<sup>429</sup> Vgl. analog die Diaspora-Effekte beim Vergleich von synthetischen und ganzheitlichen Leselehrrichtungen in den 60er Jahren bei Ferdinand (1965), zusammengefasst in Schmalohr (1971, 82 f.).

<sup>430</sup> So auch Wolf (1997) in seiner Kritik an der SCHOLASTIK-Studie von Weinert/ Helmke (1997a).

<sup>431</sup> Vgl. etwa Ramseger (1977); Benner/ Ramseger (1981); Goetze/ Jäger (1991); Wagner/ Schöll (1992); Laus/ Schöll (1995); Peschel (2003).

<sup>432</sup> S. dazu die Kritik an der Aussagekraft großer Stichprobenuntersuchungen oben Kap. 3.3).

Dieses Problem fällt allerdings nicht nur bei Fallstudien auf<sup>433</sup>, sondern auch bei größeren Erhebungen wie Günther (1988). Umgekehrt trifft die Kritik auch nicht alle Fallstudien<sup>434</sup>. Insbesondere der Wechsel von Unterrichtsformen in derselben Klasse steigert die Aussagekraft einer Untersuchung<sup>435</sup>.

Ein weiteres Problem ist die mangelnde Vergleichbarkeit von Stichproben, da diese bei persönlich so bedeutsamen Positionen wie offenem Unterricht nicht durch Zufallszuweisung von Programmen an LehrerInnen und allenfalls auf der methodisch- organisatorischen Ebene durch eine systematische Variation des Stils in derselben Klasse (Freiarbeit vs. Frontalunterricht) erreicht werden kann.

**Zwischenfazit:** Produktiv wäre eine Verzahnung von Fallstudien mit Repräsentativerhebungen über gemeinsame Fragestellungen und ihre Wiederholung in verschiedenen Klassen/ Kontexten (z. B. über Qualifikationsarbeiten von Studierenden), so dass sie "kumulativ" zu einer Absicherung der Befunde beitragen. Überdies sollten auch Fallstudien künftig kontrastiv angelegt sein.

### 3.9 Die Auswahl von Methoden und Instrumenten in einer Untersuchung ist nicht kritisch nachvollziehbar

Dieser Vorwurf richtet sich meist gegen Fallberichte. Er trifft aber auch Stichprobenuntersuchungen. Auch zum offenen Unterricht gibt es quantitative Studien, deren Instrumente und methodische Vorgehensweise nicht ausreichend dokumentiert sind<sup>436</sup>.

Wünschenswert wäre im übrigen eine Verbindung verschiedener Forschungstraditionen, um ihre spezifischen Vorteile nutzen zu können<sup>437</sup>. Beispielsweise haben nicht-standardisierte Verfahren eine wichtige Funktion, um die vermittelnden Prozesse bei der Umsetzung einer Konzeption, zumal bei einer so komplexen wie offenem Unterricht, zu doku-

<sup>433</sup> Vgl. Benner/ Ramseger (1981); Wagner/ Schöll (1992).

<sup>434</sup> Vgl. Ramseger (1977); Laus/ Schöll (1995).

<sup>435</sup> Vgl. Wagner (1978); Goetze/ Jäger (1991); Laus/ Schöll (1995).

<sup>436</sup> Vgl. zur Studie von Günther (1988) die methodische Kritik bei Brügelmann (1989).

<sup>437</sup> S. a. oben 3.8, ausführlicher: Reichardt/ Cook (1979); Brügelmann (1991); ebenso für Metaanalysen: Light/ Pillemer (1982).

mentieren und zu verstehen (s. oben 3.3). Für einen Vergleich von Effekten sind Daten aus diesen Verfahren andererseits nur geeignet, wenn sie verrechenbar gemacht werden können.

**Zwischenfazit:** Um die Basis für Vergleiche zu stärken, müssen die Methoden nachvollziehbar dokumentiert und die Daten in einer Weise dargestellt (und ggf. statistisch geprüft) werden, dass sie die Folgerungen tragen können. Wünschenswert ist die Verwendung gleicher "Brücken-Instrumente" in verschiedenen Untersuchungen, um die Befunde aus Teilstudien mit spezifischem Fokus kumulativ auswerten zu können.

#### 4 Perspektiven für die Forschung und für die Praxis

Viele kritische und einschränkende Anmerkungen - und wo bleibt das "Positive"?

Zunächst zur Forschung: Die Kritik der vorangehenden Abschnitte lässt sich in zukünftigen Untersuchungen durchaus konstruktiv nutzen. Ich schlage deshalb ein koordiniertes und nach verschiedenen Ebenen gestuftes Forschungsprogramm vor, das bereits auf eine Reihe von Vorarbeiten zurückgreifen könnte. Es sollte sich an vier Leitfragen orientieren:

1. Was *verstehen* verschiedene Beteiligte in Forschung und Praxis unter offenem Unterricht? Dazu zählen zum einen Literaturanalysen von Konzeptionen wie sie in Kap. 3.1 kurz vorgestellt wurden, aber auch Befragungen von Lehrerinnen<sup>438</sup>.
2. Welche dieser Ziele und Prinzipien offenen Unterrichts werden in der Praxis als *wichtig* für den Unterricht angesehen<sup>439</sup>?
3. In welchem Maße werden die Konzepte tatsächlich umgesetzt<sup>440</sup>?
  - Unter welchen Bedingungen ist Offener Unterricht in der Regelschule überhaupt machbar<sup>441</sup>?
  - Ist er mit vergleichbarem Aufwand realisierbar wie lehrerzentrierter Unter-

<sup>438</sup> Vgl. z. B. Kap. 3.1 in meinem Beitrag „Wie verbreitet...“ in diesem Band; Müermann (1997).

<sup>439</sup> Vgl. z. B. Kap. 3.2 in meinem Beitrag „Wie verbreitet...“ in diesem Band und Brodbeck (1992); Hanke (1996); Nohn (1997).

<sup>440</sup> Vgl. z. B. Kap. 3.3 und 3.4 in meinem Beitrag „Wie verbreitet...“ in diesem Band und Gövert u. a. (1989); Richter (1992); Könecke/ May (1994); Hanke (1998); Holtappels (1998).

<sup>441</sup> Vgl. Hopf (1990).

richt<sup>442</sup>?

4. In welchem Umfang erreichen Konzeptionen offenen Unterrichts
  - ihre eigenen und
  - weitere - im öffentlichen Schulwesen als wichtig angesehene - Ziele<sup>443</sup>?

In unserem eigenen Projekt OASE ("Offene Arbeits- und Sozialformen entwickeln") haben wir diese Fragen in Teilschritte übersetzt, denen unterschiedliche Untersuchungsformen zugeordnet werden:

- A Befragungen von LehrerInnen mit offenen und standardisierten Fragen
  - zu ihrem Begriff/ Verständnis von "Offenheit"
  - zu ihren Prioritäten im Blick auf Teilprinzipien
  - zum Grad der Realisierung dieser Teilprinzipien in ihrem Unterricht
  - zu Gründen/ Bedingungen, die die Umsetzung ihrer Ansprüche erschweren
- B Interviews zur Vertiefung dieser vier Aspekte und zusätzlich Erhebung von
  - Motiven/ Gründen für eine Entwicklung zu größerer Offenheit
  - Erfahrungen mit entsprechenden Versuchen
- C Unterrichtsbeobachtungen zu
  - Art und Grad der Realisierung von Offenheit .
  - ihren jeweiligen Rahmenbedingungen
- D Schülerbeobachtungen und Leistungstests zur Erhebung
  - der Entwicklung der Fachleistungen
  - der Entwicklung von Grundqualifikationen.

Zur Festigung und gezielten Erweiterung der Forschungsbasis schlagen wir vor, bereits existierende Untersuchungen mit einem gemeinsamen Klassifikationsschema zu erschließen, für das wir einen Vorschlag erarbeitet und erste Beispiele analysiert haben. Eine solche nach gemeinsamen Kriterien strukturierte Sammlung von Studien kann helfen, dass die Diskussion über offenen Unterricht sich von Globalurteilen löst und auf spezifischere Fragen einlässt, für die empirische Befunde genutzt werden.

<sup>442</sup> Vgl. Hippenstiel (1997)

<sup>443</sup> S. die in Kap 2. diskutierten Studien.

Forschungsmethodisch ist eine stärkere Verknüpfung erforderlich von

- verschiedenen *methodischen* Ansätzen, d.h. von
  - standardisierten und offenen Verfahren
  - Einzelfallstudien und repräsentativen Stichproben;
- Pre-/ Post-Test-*Designs* mit Prozess-Studien
- verschiedenen *Analyseebenen*, insbesondere die Berücksichtigung unterschiedlicher Kontexte von Unterricht bzw. von individuellen Lernentwicklungen.

Die Bedeutung des oben skizzierten Forschungsprogramms wird deutlich, wenn man den Druck bedenkt, unter den Reformbemühungen in den Schulen zunehmend geraten. So wird der Öffnung des Unterrichts, vor allem in der Grundschule, vorgeworfen, sie gefährde die Leistungsfähigkeit des Schulwesens<sup>444</sup> - und das, obwohl die Grundschule im Vergleich zur Sekundarstufe in den internationalen Vergleichsstudien durchweg besser abschneidet<sup>445</sup>

Die Kritik behauptet bzw. unterstellt:

- LehrerInnen entwickelten - gestützt durch neue Richtlinien bzw. durch ihre Aus- bzw. Fortbildung - *einseitige Unterrichtskonzeptionen*. In ihnen dominiere die individuelle Freiheit der Kinder und die notwendige inhaltliche und methodische Strukturierung durch die LehrerInnen werde vernachlässigt. Das "Pendel" müsse wieder zurückschwingen.
- Diese "einseitigen" Konzeptionen seien in solcher Breite und derart *nachhaltig im Schulalltag umgesetzt*, dass sie für eine beobachtete (oder unterstellte) Veränderung des Niveaus schulischen Lernens und seiner Erträge verantwortlich zu machen seien.
- Kontrollierte Wirkungsstudien hätten gezeigt, dass offener Unterricht zu schlechteren Fachleistungen und anderen ungünstigen Effekten in der Entwicklung von SchülerInnen führe.

<sup>444</sup> Öffentlichkeitswirksam vor allem Günther (1988; 1996); s. aber auch Giesecke (1995/96); Weinert in: Nöh (1997) und Schmoll (1997).

<sup>445</sup> Vgl. zuletzt Bos u. a. (2003; 2007) vs. Prenzel u. a. (2004; 2007)

Die unten<sup>446</sup> und die in diesem Beitrag berichteten Befunde stellen diese Behauptungen nachdrücklich in Frage:

- Die Idee des "offenen Unterrichts" hat *nicht* zu einseitigen Vorstellungen in den Köpfen der LehrerInnen geführt: drei Viertel von ihnen bevorzugen ein ausgewogenes Konzept, in dem ganz verschiedene didaktisch-methodische Elemente mit je 15-25% Zeitanteil vertreten sind. Im durchschnittlichen "Ideal-Stundenplan" der Befragten dominieren von der Lehrerin inhaltlich bestimmte Aktivitäten sogar zwei Drittel der Unterrichtszeit.  
Selbst unter LehrerInnen, die noch in der Ausbildung sind, werden Konzepte, in denen gemeinsam entwickelte Projektarbeit oder individuell gewählte Aufgaben mehr als 1/3 der Unterrichtszeit ausmachen, nur von 10% bzw. 5% der Befragten vertreten.
- Evtl. Mängel in Fachleistungen oder anderen Unterrichtseffekten können *nicht* dem "offenem Unterricht" angelastet werden. In der Realität stellen seine verschiedenen Varianten immer noch die Ausnahme dar. Selbst traditionelle und auch von den Kritikern geforderte Formen der Differenzierung werden von der Mehrheit der LehrerInnen noch nicht einmal täglich realisiert; inhaltliche Formen der Selbst- und Mitbestimmung von Kindern werden von der Mehrheit der LehrerInnen nicht einmal wöchentlich angeboten.

Insgesamt wird eine konsequente "Öffnung des Unterrichts" (je nach Kriterium) allenfalls von 5-15% der LehrerInnen realisiert.

Zum dritten Vorwurf sind die in diesem Beitrag berichteten kontrollierten Wirkungsstudien bedeutsam. Sie zeigen im Durchschnitt *keine generelle Überlegenheit* eines lehrerInnen-zentrierten Unterrichts. Selbst für fachliche Leistungen lassen sich statistisch allenfalls marginale, d.h. pädagogisch nicht bedeutsame Vorteile nachweisen.

**Zwischenbilanz** : Es gibt weder Anzeichen für eine didaktische Revolution in den Köpfen der LehrerInnen, noch Belege für eine Umwälzung der Praxis. Es gibt aber auch keinen

<sup>446</sup> Vgl. Kap. 3.2 und 3.3 in meinem Beitrag „Wie verbreitet...“ in diesem Band.

Grund, Angst vor den Auswirkungen einer stärkeren Öffnung des Unterrichts zu haben. Sie scheint sogar notwendig, um angesichts der geringen Verbreitung offenen Unterrichts das sog. "Pendel" überhaupt erst in Schwung zu bringen.

Für die Unterrichtspraxis bedeuten die referierten Befunde - in aller Kürze und mit entsprechender Vorsicht: Wer Kindern aus *normativen Gründen* mehr Selbständigkeit zutrauen und ihnen mehr Verantwortung für ihr Lernen zumuten will,

- hat keine *empirischen Befunde* zu fürchten, die grundsätzlich dagegen sprechen;
- sollte sich als erstes klar machen, welche konkreten Schritte er oder sie zur Öffnung des Unterrichts beschreiten will und welche Befunde zu diesen *spezifischen* Formen vorliegen;
- braucht *generell*/keine Sorgen zu haben, dass eine Öffnung des Unterrichts auf Kosten der fachlichen Leistungen geht, da diese im Mittel der Studien gleich gut oder nur geringfügig schwächer ausgefallen sind als in Regelklassen - in einer Reihe von Klassen sogar deutlich besser;
- kann andererseits positive, aber im Mittel *nicht dramatische* Vorteile im Persönlichkeitsbereich und bei Einstellungen zur Schule, zu den Fächern und zum Lernen erwarten;
- muss allerdings mit *anderen Lernkurven* rechnen, als sie ihm oder ihr aus dem zentral gesteuerten Lehrgangsunterricht vertraut sind<sup>447</sup>;
- sollte nicht vergessen, dass die *Streuung* innerhalb der untersuchten Gruppen immer deutlich größer war als die Differenz der Mittelwerte zwischen ihnen.

---

<sup>447</sup> Vgl. den aus der Forschung zu kompensatorischen Vorschulprogrammen bekannten „sleepers-effect“ und speziell zu nicht-linearen Lernverläufen am Beispiel der Rechtschreibentwicklung Brinkmann (1997; 2003) und die Befunde zu verschiedenen Ansätzen des Anfangsunterrichts im "Schreibvergleich BRDDR" bei Brügelmann u. a. (1994b).

Die letzte Anmerkung heißt: Wie der jeweilige Ansatz umgesetzt wurde, was die Lehrerin spezifisch "dazugetan" hat, spielte eine erhebliche Rolle für die Wirkungen<sup>448</sup>. Von besonderer Bedeutung ist dabei eine Perspektive, die bisher vor allem in der Forschung, weniger in der Praxis zu kurz gekommen ist: Wie entwickelt sich die Selbständigkeit, die offener Unterricht voraussetzt - und in welchen konkreten Formen kann offener Unterricht am besten dazu beitragen, sie zu stützen und weiter zu entwickeln<sup>449</sup>?

Für die Forschungs- wie für die Unterrichtsmethode gilt insofern gleichermaßen: Der Teufel steckt wie so oft im Detail.

Oder anders gesagt: Wenn der methodische Teufel im Detail steckt, hängt das pädagogische Heil zu einem guten Teil immer noch an der (Lehr-)Person - gerade im offenen Unterricht. Dies bedeutet: Wir brauchen eine zweite, situationsbezogene Ebene der Forschung, nämlich durch LehrerInnen selbst, die ihren Unterricht dokumentieren, mit KollegInnen Erfahrungen austauschen und - zumindest punktuell - an der Systematisierung solcher Erfahrungen durch Hochschulprojekte beteiligt werden<sup>450</sup>.

---

<sup>448</sup> Ebenso - allerdings in stärker fachdidaktischer Akzentuierung - Lipowsky (2007b).

<sup>449</sup> Vgl. Weinert (1982); Paris/ Newman (1990); Zimmerman (1990); Lompscher (1997).

<sup>450</sup> Vgl. zur grundsätzlichen Notwendigkeit einer solchen kontextsensiblen Forschung meine Überlegungen in Brügelmann (2005a, Kap. 56-63; 2005b; 2007, 23ff., 32ff.).

