

Evaluation evaluativer Prozesse von sprachlichen Lehr-Lernkontexten¹

Werden Lehr-Lern-Prozesse, Bildungsangebote und Fördermaßnahmen initiiert und damit zusammenhängend dann in institutionellen Kontexten auch finanziert, stellt sich immer wieder die Frage nach der Sinnhaftigkeit und nach dem Erfolg dieser Angebote und so werden mehr oder weniger umfangreiche Evaluationen unternommen. In meinem Vortrag möchte ich heute einen Blick auf genau diese Evaluationen werfen und damit in gewisser Weise über die Evaluation von Evaluationsprozessen reflektieren.

Etymologisch betrachtet geht es bei der Evaluation von Lehr-Lern-Prozessen um eine Bewertung oder Wertung (von lateinisch *valere* = Wert sein). Möchten wir evaluative Prozesse evaluieren, also eine Evaluation der Evaluation durchführen, stellt sich daher m. E. zunächst die Frage nach dem Bewertungsmaßstab. Woran mache ich eigentlich fest, ob eine Evaluation gelungen ist?

Dem Forschenden kommen hierzu vermutlich spontan die „klassischen“ Gütekriterien in den Sinn, als da wären: Objektivität, Reliabilität und Validität oder anders formuliert: Erstens: Wie objektiv sind die Ergebnisse, d. h. wie unabhängig ist das, was herausgefunden wurde, von der analysierenden oder evaluierenden Person?, zweitens: Wie reliabel oder zuverlässig sind die Ergebnisse? (was auch die Haltbarkeit der Ergebnisse bei Wiederholungen mit einschließt), und drittens: Wie valide sind die Ergebnisse, d. h. wird überhaupt das gemessen oder evaluiert, was vorgegeben wird?

Vor allem mit Blick auf den Aspekt der Validität stellt sich in Verlängerung auch die Frage, wie Lehr-Lernprozesse überhaupt evaluiert werden können. Nun ja: Um einen Lehr-Lernkontext zu evaluieren, gilt es diesen zunächst zu beforschen, wofür grundsätzlich verschiedene Methoden zur Verfügung stehen: So können wir einen Lehr-Lernprozess indirekt evaluieren, indem wir den Lernzuwachs oder „Output“ „messen“, bspw. indem vor und nach einer Lerneinheit ein bestimmter Test eingesetzt wird, um den Lernzuwachs zu ermitteln; methodisch könnten wir dies der Kategorie „Testen und Prüfen“ zuordnen. Eine weitere Möglichkeit besteht darin, den Lehr-Lernprozess mittels einer Beobachtung als solchen zu fokussieren und ihn auf dieser Grundlage zu evaluieren; außerdem ist es möglich, die am Prozess beteiligten Akteure, also die Lehrpersonen wie Erzieherinnen und Erzieher und die Lernenden, in diesem Fall die Kinder, oder möglicherweise auch deren Eltern, zu befragen und auf diese Weise Hinweise über die Lehr-Lernsituation zu erhalten.

Doch Vorsicht: Um sich Lehr-Lernkontexten auf einer oder sogar mehrerer dieser Wege zu nähern oder diese „unter die Lupe zu nehmen“, bedarf es für Forschende einiger Vorarbeiten oder Voraussetzungen. Beispielsweise ist es, ob wir die Situation nun beobachten, ob wir einzelne Akteure befragen oder ob wir diese testen, erforderlich, im

¹ Der vorliegende Beitrag basiert auf einem Vortrag, gehalten am 23.09.2014 in Flensburg. Der Vortragsstil wurde weitestgehend beibehalten. Zur Zitation: Ricart Brede, Julia (2014): Evaluation evaluativer Prozesse von sprachlichen Lehr-Lernprozessen. Vortrag im Rahmen der 6. Tagung „Am Phänomen lernen“, Europa-Universität Flensburg.

Vorfeld ein Einverständnis von den „Beforschten“ einzuholen. Dies mag vielleicht banal erscheinen, doch kann dieses Erfordernis in der Realität leicht zu einem ersten Stolperstein für evaluative Prozesse werden: Um bspw. Beobachtungen in einer Kindertageseinrichtung durchführen zu können, müssen Erzieherin, Kita-Leitung und Eltern darüber informiert sein UND schriftlich ihr Einverständnis gegeben haben. Nimmt man die Aufgabe des Informierens ernst, stößt man möglicherweise bereits auf einen zweiten Stolperstein, denken wir an Eltern, für die Deutsch keine Erstsprache ist, und für die u. U. Übersetzungen bereitgestellt werden müssen.

Des Weiteren gilt es zu differenzieren, ob allein das Testen, Befragen oder Beobachten genehmigt werden soll, oder ob erhobene Daten (wie Audio- oder gar Videoaufzeichnungen) später auch für Veröffentlichungen, Vorträge oder Workshops genutzt werden dürfen. Letzteres ist insbesondere für formative Evaluationsprozesse, bei denen Evaluationsergebnisse direkt an die Praxis rückgebunden werden, von Bedeutung. Für einen solchen Fall sollten Einverständniserklärungen nach Möglichkeit mehrstufig angelegt sein, um einerseits das Einverständnis für die Analyse und Evaluation des aufgezeichneten Materials durch die „Forschergruppe“ festzuhalten und andererseits das Einverständnis der weiteren Verwendung der Daten abzubilden. Insbesondere letzteres entpuppt sich für Forschende jedoch häufig nicht nur als Stolpersteinchen, sondern geradezu als Hinkelstein, sodass sich dem ein oder anderen die Frage stellen mag, ob es tatsächlich von derart hoher Bedeutung ist, dass Audio- oder gar Videoaufzeichnungen für Dritte freigegeben werden sind. Doch dazu später. Halten wir zunächst das Informieren (teilweise in verschiedenen Sprachen) und das auf jeden Fall schriftliche und evtl. sogar mehrstufige Einholen der Einverständnisse ALLER Beteiligten (was unter Umständen im Übrigen auch das Ministerium einschließen kann) als Grundvoraussetzungen für die Evaluation von Lehr-Lernsituationen fest.

Kommen wir zurück zu den Erhebungsformaten Befragung, Beobachtung sowie Test, die in dieser Form noch sehr weit gefasst sind. Aussagen über die Validität, die Objektivität oder die Reliabilität der Evaluation, sind damit gewiss noch nicht möglich. Je nach Evaluations- bzw. Erkenntnisinteresse würden wir uns demnach auf eine oder sogar mehrere Erhebungsmethoden festlegen und in einem nächsten Schritt die einzelnen Fragen für die Befragung oder den Test bzw. die Kategorien für die Beobachtung festlegen, um die Lehr-Lernsituation valide abzubilden und damit zu einer adäquaten Evaluation zu gelangen.

Führen wir diese Aufgabe einmal am Beispiel einer Beobachtung gedanklich fort: Als nächstes gälte es also, die Forschungsfrage oder auch die ForschungsfrageN zu konkretisieren. Dabei gilt es Dinge messbar zu machen und in diesem Zuge auch, latente, d. h. nicht direkt beobachtbare Variablen in manifeste zu überführen. Indem wir sorgfältig recherchieren, also eine Literatursichtung betreiben, können wir dabei zwar mitunter auf bereits Erprobtes zurückgreifen, doch gilt es durch Anwendung und Überprüfung sicherzustellen, dass das Gefundene für die eigenen Zwecke passgenau ist und falls es denn adaptiert werden muss, gilt es hiernach zu prüfen, ob es dann auch tatsächlich hält, was es verspricht, ob es von unterschiedlichen Personen anwendbar ist und dabei zu denselben Ergebnissen führt u. d. m. Um eine derartige Prüfung vorzunehmen, wird eine Situation häufig von zwei unabhängigen Beobachtenden analysiert. Die Ergebnisse werden anschließend über einen Abgleich auf ihre Konsistenz hin überprüft, wobei Maße wie Cohens-Kappa hierfür etabliert und weit verbreitet sind (vgl. auch Wirtz/ Kutschmann 2007, Ricart Brede et al. 2010). Eine Beobachtung, die mit Hilfe eines Beobachtungssystems durchgeführt wurde, für das der Kappa-Wert 0.7 oder höher ist (Cohens Kappa kann Werte bis zu 1.0 erreichen), kann für die Evaluation der Beobachtung damit als erstes positives Signal gewertet werden. Schön und gut.

Doch haben Sie darüber nachgedacht, wer eigentlich diese unabhängigen Beobachtenden sind, die die Probekodierung durchführen? Nehmen wir einmal an, in einem interdisziplinären Forschungsprojekt zur Evaluation vorschulischer Sprachfördersituationen, bei denen Kinder Versuche durchführen, arbeiteten Linguistinnen und Linguisten mit Naturwissenschaftlerinnen und Naturwissenschaftlern zusammen. Von Interesse wären in diesem Zusammenhang nun auch Erklärprozesse, die über Beobachtungskategorien zunächst identifiziert werden sollten. Mutmaßlich würden die Ohren der sprachwissenschaftlichen Fraktion u. a. dann hellhörig werden, wenn kausale Konnektoren wie „weil“ oder „da“ in den Kinderäußerungen zu hören wären; vermutlich würden sie sich hierin dann auch mit den Kolleginnen und Kollegen der naturwissenschaftlichen Disziplinen einig werden, da der Schluss vom Explanans auf das Explanandum nach dem Hempel-Oppenheim-Schema Teil einer Erklärung ist und derartige Schlüsse sprachlich häufig über Konnektoren realisiert werden (vgl. Poser 2001: 45, Klein 2009: 31, auch Neumeister und Vogt 2009: 565). Doch wie sieht es mit den Antescendenzbedingungen aus? So zählt nach dem in naturwissenschaftlichen Kontexten durchaus verbreiteten HO-Schema die Aufführung der Voraussetzungen, die notwendig sind, um den Schluss vom Explanans auf das Explanandum vorzubereiten, bereits zur Erklärung (vgl. Poser 2001: 45). Ist dies in den Sprachwissenschaften nun genau so? Keineswegs, wenn man „Erklären“ bspw. wie Vollmer als eine Diskursfunktion oder Sprachhandlung betrachtet, die sich dadurch auszeichnet, dass „im Unterschied zur Diskursfunktion DESCRIBING (also dem Beschreiben) beim EXPLAINING die Ursachen von Ereignissen, Situationen, Phänomenen oder Prozessen genannt werden“ (Vollmer 2011: 7) müssen. Als Beispiel für eine Erklärung führt Vollmer im Folgenden die Äußerung an: „Eisbären sind weiß, da ihre Haare nicht dunkel pigmentiert sind.“ (Vollmer 2001: 7) Möglicherweise würde der Naturwissenschaftler bzw. die Naturwissenschaftlerin dem, so könnte ich mir vorstellen, dann doch noch die Antescendenzbedingung vorangestellt wissen wollen, dass es eben die Pigmente sind, die für die dunkle Farbe der Haare sorgen.

Was wir daran sehen können: Nicht nur die Festlegung und Beschreibung der Beobachtungskategorien an sich ist bereits eine aufwändige Angelegenheit; weiterhin gilt es zu fragen, in welchen Kontexten und Fachdisziplinen diese angewandt werden sollen und wer die Kodierung vornimmt. Das heißt folglich: Statistische Kennwerte wie Cohens-Kappa und Co leisten lediglich EINEN Beitrag dazu, um evaluative Prozesse bewerten zu können. Mindestens ebenso wichtig sind aus meiner Sicht Transparenz und Nachvollziehbarkeit, indem Beobachtungskategorien bspw. nicht einfach aufgestellt, sondern indem diese in einem Handbuch auch beschrieben werden und indem angegeben wird, welche Personen welche Kodierungen vorgenommen haben.

Im Übrigen: Selbst wenn wir all das bislang Gesagte berücksichtigen und selbst wenn unser Beobachtungssystem im Probedurchlauf Kappa-Werte von 1.0 erzielt hat, Beobachtungen sind stets gebrandmarkt durch das sogenannte „Beobachterparadoxon“ (vgl. Labov 1971, auch Burdick 2008, Ricart Brede 2014: 138) und können sich demnach nicht davon befreien, dass die Beobachteten – um die Beobachtung wissend – sich möglicherweise anders verhalten: indem Lehrerinnen und Lehrer versuchen ihr bestes zu geben, indem Kinder den „Klassen- oder Gruppenclown“ markieren, oder einfach nur, indem sie schüchterner sind als gewöhnlich und deshalb weniger sagen oder sich anders beteiligen.

Sicherlich mag der Grad an Invasivität deutlich größer sein, wenn für die Beobachtungen Audio- oder gar Videogeräte zum Einsatz kommen – doch: ein Rest Invasivität wird immer bleiben, und auch Gewöhnungsprozesse scheinen nicht wirklich stattzufinden, d. h. auch nach dem fünften Mal videografieren, sind sich Lernende gleichermaßen wie Lehrende der

Aufnahmesituation sowie der Gerätepräsenz durchaus bewusst – und dies trotz möglichst unauffälligem Verhalten der Beobachtenden (vgl. Maak/ Ricart Brede 2014: 161). Stark abhängig ist der Grad der Invasivität hingegen – wie zu erwarten – von der Positionierung ggf. eingesetzter Aufnahmegeräte (wie einer Videokamera) (vgl. Maak/ Ricart Brede 2014: 165 f.). Auch hier greift daher wieder das Gütekriterium der Transparenz: Um getätigte Beobachtungen evaluieren zu können, sind Informationen über die Erhebungssituation hilfreich, sodass ein Kameraskript weiterhilft, das Informationen über das Verhalten der Kamerapersonen sowie in Form von Stellplänen Informationen über die Positionierung der Geräte im Raum enthält.

Sie ahnen vermutlich, worauf ich hinaus möchte: Das Informieren, das Einholen der Einverständnisse – mehrstufig und in mehreren Sprachen –, das Auswählen, Adaptieren, Aushandeln, Erproben und Prüfen von Beobachtungskategorien, von Testfragen und Interviewfragen, das Erstellen von Manualen, Kameraskripten und schließlich das Anwenden der entwickelten Kategorien, das Aufbereiten, das Transkribieren und das Auswerten der Daten: Wer – bitteschön – soll denn all dies leisten?

An dieser Stelle zeigt sich nun ein weiteres Gütekriterium, das im wissenschaftlichen Kontext für mein Gefühl allzu häufig außer Acht wird oder aus dem Gleichgewicht geraten ist: nämlich die Kosten-Nutzen-Relation bzw. die Frage, in welchem Verhältnis Aufwand und Nutzen zueinander stehen. Den Aufwand sauber angelegter Evaluationen von Lehr-Lernprozessen habe ich versucht im Rahmen meines Vortrags – vorrangig mit Blick auf (videographische) Beobachtungen – herauszustellen. Stellt sich nun die Frage nach dem Nutzen oder dem Ertrag derartiger Evaluationen. Und genau hieran wird nun, so meine ich, der besondere Mehrgewinn videographisch angelegter Beobachtungen deutlich, die unglaubliche Flexibilität bieten: Die Iterationsmöglichkeit erlaubt es, videographisch beobachtete Lehr-Lern-Situationen immer und immer wieder anzuschauen – mit veränderten Beobachtungskategorien, um Ergebnisse zu prüfen sowie um ein interdisziplinäres Arbeiten und Aushandeln zu ermöglichen (vgl. Ricart Brede 2011: 94, Helmke 2004: 179, Petko et al. 2003: 265, 279). Denkbar wäre es allerdings bspw. auch, das neue Forschergruppen bereits erhobene und vorliegende Videodaten unter anderen Gesichtspunkten und mit neuen Fragestellungen beleuchten – in der Realität ist dies allerdings leider insofern ein eher selten vorkommender Glücksfall, als Videodaten i. d. R. ausschließlich für die Forschergruppe freigegeben sind, die die Daten auch erhoben hat. Videodatenbanken wie unterrichtsvideos.ch, die Lehr-Lernbeispiele und damit möglicherweise auch „Best-Practice-Beispiele“ zeigen (nebenbei bemerkt: ein weiterer Nutzen von Videobeispielen), sind daher leider äußerst selten. In abgeschwächter Form stellen möglicherweise Transkriptdatenbanken sowie Transkriptbände, die die Kommunikation audio- und videographisch aufgezeichneter Beobachtungen immerhin schriftlich abbilden, eine Alternative dar (für ein Bsp. s. auch Ricart Brede 2014).

Ein hoher Nutzen bedeutet für mich neben Wissenschaftstransfer zudem vor allem ein Ertrag für die Praxis, auch für die Lehr-Lern-Situation, die evaluiert wurde. Und gerade hier können Videodaten – insbesondere im Zuge formativer Evaluationen – ebenfalls einen wichtigen Beitrag leisten, indem sie Ergebnisse anschaulich machen und es auf diese Weise ermöglichen, darüber zu diskutieren. „Videodaten bilden daher“, so Petko et al., „einen Kristallisationspunkt der fachdidaktischen Verständigung über Unterrichtsprozesse und bauen eine Brücke über die Kluft zwischen Theorie und Praxis“ (Petko et al. 2003: 278), die leider allzu oft zu konstatieren ist.

So können in Workshops und Fortbildungen anhand von Videobeispielen Merkmale oder Kriterien für gelungene Praxis herausgearbeitet werden. Für mich als besonders ertragreich

hat sich in diesem Zusammenhang die Methode der kontrastiven Analyse gezeigt, bei der zwei Videobeispiele miteinander verglichen werden – nicht, um diese wertend in eine Rangfolge zu stellen, sondern um das jeweils beispielhafte und besonders gelungene daran herauszustellen.

Videodaten können aber auch für die Konzept- und Materialentwicklung wichtige Einblicke in die Praxis bieten. Ein Beispiel zeigt das folgende Ablaufschema, das aufgrund der videographischen Analyse von 48 vorschulischen Sprachfördereinheiten ermittelt werden konnte.

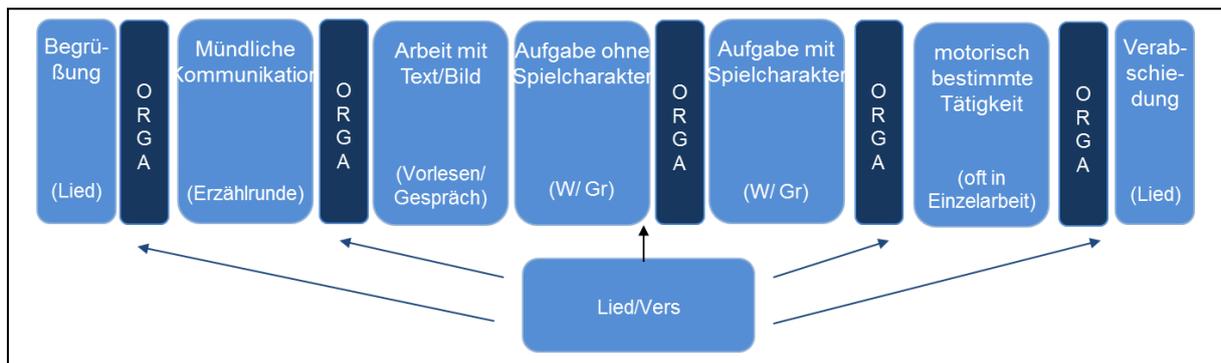


Abb. 1: Typisches Ablaufmuster einer vorschulischen Sprachfördereinheit (vgl. Ricart Brede 2011: 183).

Hierzu wurden die videographierten Sprachfördereinheiten je nach Aktivitäten in Handlungssequenzen eingeteilt, bevor diese – je nach sprachlichen Handlungen und Sozialformen weiter kodiert wurden. Interessanterweise folgten die Sprachfördereinheiten, obwohl diese an ganz unterschiedlichen Einrichtungen stattfanden und von unterschiedlichen Erzieherinnen oder Sprachförderpersonen abgehalten wurden, alle demselben Schema: Auf eine Begrüßung folgt ein Wechsel von organisatorischen Sequenzen und inhaltlichen Sequenzen oder sogenannten inszenierten Sprachlernsituationen, wobei für die „inszenierten Sprachlernsituationen“ zunächst ein Erzählkreis typisch ist, auf den in der Regel die Arbeit mit einem Bild oder Text folgt, bevor eher formale Aufgaben, dann eher spielerische Aufgaben (wie Brettspiele) und schließlich stark motorisch bestimmte Tätigkeiten wie Bastelsituationen oder Turnspiele initiiert werden, bevor eine Verabschiedung die Einheit beschließt.

Interessanterweise wird zudem in jeder Einheit ein Lied gesungen oder ein Vers gesprochen, wobei sich diese Aktivität nicht eindeutig im Ablauf verorten lässt; - möglicherweise deshalb, da Verse und Lieder zum integralen Bestandteil der Arbeit an Kindertageseinrichtungen zählen und ihr Einsatz deshalb nicht geplant wird, sondern dann erfolgt, wenn Kinder unruhig werden oder ähnliches.

Worin besteht nun der Nutzen einer Kenntnis eines derartigen Ablaufschemas? Ganz einfach: Es zeigt, wie Praxis abläuft und bietet damit Anschlussmöglichkeiten z. B. für Fortbildungen oder eben für die Konzept- und Materialentwicklung. Bspw. wird anhand des Ablaufschemas der hohe Anteil organisatorischer Sequenzen sichtbar, sodass darüber diskutiert werden kann, inwiefern auch in diesen Sequenzen ein aufmerksamer Sprachgebrauch sinnvoll ist und inwiefern auch organisatorische Sequenzen genutzt werden können, um sprachbildende Prozesse zu initiieren und Gesprächsanlässe zu schaffen.

Die nun hier herausgestellten Vorzüge videographisch angelegter Beobachtungen sollten nun allerdings keinesfalls zu dem Schluss führen, Beobachtungen oder gar videographisch angelegte Beobachtungen seien besser oder wertiger als Evaluationsstudien, die auf Fragebogen- und/oder Interviewdaten basierten. Keinesfalls! Lediglich sollte die

Herausstellung der Vorzüge deutlich machen, dass es, selbst wenn videographisch angelegte Beobachtungen einen immensen Aufwand bedeuten, doch auch auf der anderen Seite einige Argumente in die Waagschale zu werfen gilt und dass es stets ein Abwägen sein sollte, ob der Nutzen dem Aufwand gerecht wird.

Und diese Frage hängt nun wiederum davon ab, welches Ziel mit der Evaluation verfolgt werden soll, welche Fragen an die Lehr-Lern-Situation gestellt werden, was also das übergeordnete Erkenntnisinteresse ist. Die Passung des Evaluations- oder Forschungsdesigns zur Forschungsfrage stellt daher eine Art übergeordnetes Gütekriterium dar, das es immer wieder zu prüfen gilt.

Um auf den Ertrag oder Nutzen zurückzukommen, den Evaluationen – gerade im Kontext von Lehr-Lern-Prozessen – leisten: Eine gewisse Verantwortung und auch Verpflichtung haben Evaluierende in Bildungskontexten m. E. zudem gegenüber der Bildungspolitik: Es geht darum, aufzuzeigen, wo noch Optimierungsbedarf besteht, wo es Stellschrauben zu drehen gilt und letztendlich wo es weiter zu investieren gilt und wo eben nicht (zumindest ist das leider häufig eine Konsequenz). Als ein Beispiel dafür, welche Verpflichtung Evaluierende daher in Bezug auf die Datenpräsentation und -interpretation haben und wie sensibel, zurückhaltend und wohlüberlegt es daher mit Ergebnissen umzugehen gilt, möchte ich im Folgenden einige wenige Beispiele aus der wissenschaftlichen Begleitung zum Programm „Sag‘ mal was – Sprachförderung für Vorschulkinder“ anführen; ein Projekt, dem ich mich durch meine Promotionsarbeit in besonderer Weise verbunden fühle und aus dem auch die bereits gezeigten Ergebnisse der Videoanalysen stammen.

Mit dem Programm „Sag‘ mal was – Sprachförderung für Vorschulkinder“ wurden seit 2003 seitens der Baden-Württemberg Stiftung (bzw. ehemals der Landesstiftung Baden-Württemberg) insgesamt 39 Mio. € für Sprachförderprogramme bereitgestellt. Fast 90.000 Vorschulkinder in Baden-Württemberg konnten damit erreicht werden (vgl. <http://www.sagmalwas-bw.de/>). Darüber hinaus sah das Programm auch eine wissenschaftliche Begleitung durch die Standorte Heidelberg und Weingarten vor (vgl. Baden-Württemberg Stiftung 2011, Roos/ Polotzek/ Schöler 2010 sowie Gasteiger-Klicpera/ Knapp/ Kucharz 2010). Am Standort Weingarten gestaltete sich die Evaluation der Sprachfördermaßnahmen in einem umfangreichen Längsschnittprojekt, bei dem insgesamt 1.159 Kinder in zwei Kohorten (die zwei aufeinanderfolgenden Jahrgängen entsprachen) getestet und beobachtet wurden. Die beiden Kohorten umfassten dabei 864 Förderkinder mit Sprachförderbedarf und 295 Vergleichskinder ohne Sprachförderbedarf. Als Evaluationsinstrumente wurden neben den bereits angeführten videographischen Beobachtungen der Sprachfördersituationen, auch Dokumentenanalysen von Förderplänen durchgeführt, es wurden schriftliche sowie mündliche Befragungen mit Eltern und Erzieherinnen durchgeführt, und es wurden umfangreiche Testungen mit den Kindern vorgenommen – sowohl mit den Förderkindern als auch mit den Vergleichskindern – und zwar jeweils zu Beginn der Fördermaßnahmen, nach einem Jahr Förderung (bzw. eben Nicht-Förderung) sowie nach dem ersten Schuljahr.

Blicken wir in Abb. 2 auf die Ergebnisse einer Varianzanalyse zu zwei Untertests aus dem sog. „Sprachscreening für Vorschulkinder“ (kurz: SSV), links dem Untertest „Phonologisches Gedächtnis für Nichtwörter“ (kurz: PGN) sowie rechts dem Untertest „Satzgedächtnis“ (kurz: SG): T1 markiert jeweils den ersten Messzeitpunkt vor Förderbeginn, T2 markiert den zweiten Messzeitpunkt nach einem Jahr. In blauer Schriftfarbe gehalten sind die durchschnittlichen Fortschritte in der Sprachentwicklung für die Vergleichskinder; in roter Schriftfarbe sehen wir die durchschnittliche Entwicklung für die Förderkinder.

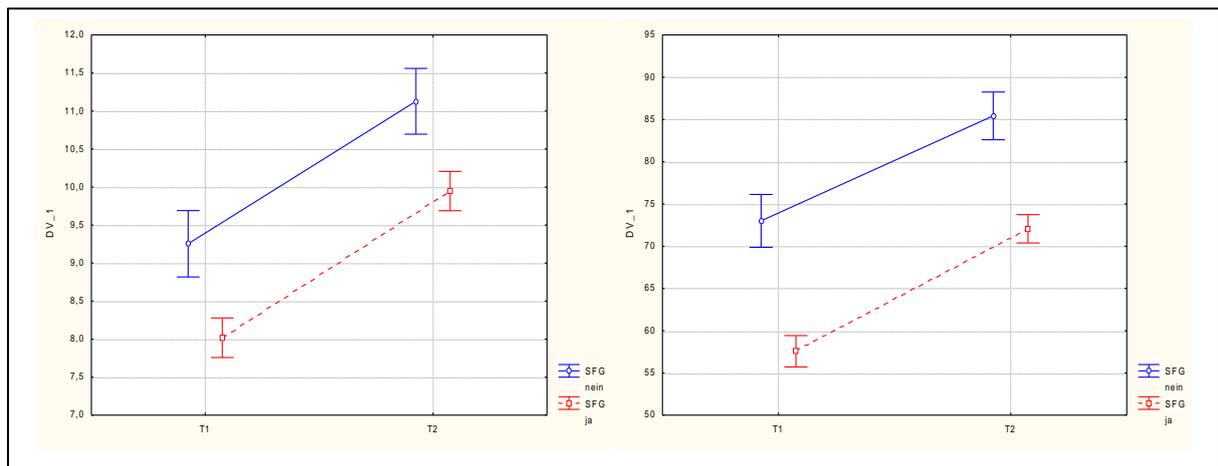


Abb. 2: Effekte der Gesamtgruppe in der wiss. Begleitung zu „Sag' mal was – Sprachförderung für Vorschulkinder“, aus: Gasteiger-Klicpera et. al 2010: 49, 52.

Ein möglicher Schluss wäre nun der, die Förderung sei erfolglos gewesen; schließlich scheinen die Linien jeweils beinahe parallel zu verlaufen, d. h. die Förderkinder holen die Vergleichskinder über den Förderzeitraum nicht ein. M. E. ein grober Fehler, denn wissen wir, wie die Entwicklung der Förderkinder ohne Förderung ausgesehen hätte? Möglicherweise hätte sich der Abstand zwischen den beiden Gruppen und damit die Schere ohne Förderung weiter vergrößert? Kurz: Was erwarten wir nach einem Jahr Sprachförderung eigentlich? Wie viel „Sprachzuwachs“ müssen die Kinder zeigen, damit wir von erfolgreicher Sprachförderung sprechen?

Mit Blick auf mögliche bildungspolitische Konsequenzen können derartige Interpretationen fatale Folgen haben. Glücklicherweise waren die Konklusionen der Baden-Württemberg Stiftung andere. So zeigten sich bei genauerem Betrachten insbesondere für solche Kinder Zuwächse, die zu Beginn der Förderung besonders jung waren. Auf diese Weise entstand in Verlängerung zu „Sag' mal was – Sprachförderung für Vorschulkinder“ das Programm „Sag' mal was – Sprachliche Bildung für Kleinkinder“ (vgl. <http://www.sagmalwas-bw.de/>).

Was für mich daher – eben gerade auch mit Blick auf bildungspolitische Konsequenzen – zwei weitere Gütekriterien für die Evaluation von Lehr-Lernprozessen sind: einerseits Vorsichtigkeit und Bescheidenheit in interpretativen Schlüssen und andererseits Selbstkritik – und damit die Bereitschaft aus der eigenen Forschung oder – erlauben Sie mir diese Formulierung: „am Phänomen“ – zu lernen.

Und damit – so möchte ich selbstkritisch schließen – habe ich nun vorrangig die Hindernisse, Stolpersteine und Hinkelsteine evaluativer Prozesse beleuchtet und weniger die Sonnenmomente und Sternstunden – doch sind Natur und Landschaften etwas wunderbares – ebenso wie Lehr- und Lernprozesse es mitunter sein können und ebenso erfreulich und lustvoll kann daher auch ihre Beforschung sein.

Literatur

- Baden-Württemberg Stiftung (Hrsg.) (2011): Sag' mal was - Sprachförderung für Vorschulkinder. Tübingen: Franke.
- Burdick, Colin M. (2008): Das Beobachterparadoxon in der Feldarbeit. Eine linguistische Perspektive. Saarbrücken: vdm.
- Gasteiger-Klicpera, Barbara/ Knapp, Werner/ Kucharz, Diemut/ Patzelt, Doreen/ Ricart Brede, Julia/ Schmidt, Barbara/ Vomhof, Beate (2010): Abschlussbericht der

- Wissenschaftlichen Begleitung des Programms „Sag’ mal was – Sprachförderung für Vorschulkinder“. Pädagogische Hochschule Weingarten [http://www.sagmalwas-bw.de/media/WiBe%201/pdf/PH-Weingarten_Abschlussbericht_2010.pdf, 22.09.2014].
- Helmke, Andreas (2004/ 2): Unterrichtsqualität erfassen, bewerten, verbessern. Seelze: Kallmeyer.
- Klein, Josef (2009): ERKLÄREN-WAS, ERKLÄREN-WIE, ERKLÄREN-WARUM. Typologie und Komplexität zentraler Akte der Welterschließung. In: Vogt, Rüdiger (Hrsg.): Erklären. Gesprächsanalytische und fachdidaktische Perspektiven. Reihe: Linguistik, Band 52. Tübingen: Stauffenburg, S. 25–36.
- Labov, William (1971): Das Studium der Sprache im sozialen Kontext. In: Klein, Wolfgang/ Wunderlich, Dieter (Hrsg.): Aspekte der Soziolinguistik. Frankfurt a. M.: Athenäum, S. 111-194.
- Maak, Diana/ Ricart Brede (2014): Empirische Erfassung von Invasivität in videografierten Lehr-Lernsituationen: Entwicklung und Erprobung eines Beobachtungssystems. In: Neumann, Astrid/ Mahler, Isabelle (Hrsg.): Empirische Methoden der Deutschdidaktik: audio- und videografierte Unterrichtsforschung. Baltmannsweiler: Schneider Verlag Hohengehren.
- Neumeister, Nicole; Vogt, Rüdiger (2009): Erklären im Unterricht. In: Becker-Mrotzek, Michael (Hrsg.): Mündliche Kommunikation und Gesprächsdidaktik. Reihe: DTP, Band 3. Baltmannsweiler: Schneider Verlag Hohengehren, S. 562–583.
- Petko, Dominik et al. (2003): Methodologische Überlegungen zur videogestützten Forschung in der Mathematikdidaktik. Ansätze der TIMSS 1999 Video Studie und ihrer schweizerischen Erweiterung. In: ZDM (Zentralblatt für Didaktik der Mathematik). Vol. 35, Nr. 6, S. 265-280.
- Poser, Hans (2001): Wissenschaftstheorie. Eine philosophische Einführung. Stuttgart: Reclam.
- Ricart Brede, Julia (2011): Videobasierte Qualitätsanalyse vorschulischer Sprachfördersituationen. Freiburg i. Br.: Fillibach.
- Ricart Brede, Julia (2014): Beobachtung. In: Settineri, Julia/ Demirkaya, Sevilen/ Feldmeier, Alexis/ Gültekin-Karakoç, Nazan/ Riemer, Claudia (Hrsg.): Empirische Forschungsmethoden für Deutsch als Fremd- und Zweitsprache. Eine Einführung. Paderborn: Schöningh, S. 137-145.
- Ricart Brede, Julia (2014): Vorschulische Sprachfördersituationen. Ein aufbereiteter und kommentierter Transkriptband aus dem Projektkontext von „Sag’ mal was – Sprachförderung für Vorschulkinder“. Reihe: Berichte und Materialien www.dazportal.de, Band 2 [http://www.dazportal.de/index.php?option=com_content&view=article&id=44&Itemid=39, 05.10.2014].
- Ricart Brede, Julia et al. (2010): Die Entwicklung von Beobachtungssystemen in der videobasierten Forschung am Beispiel von Qualitätsanalysen vorschulischer Sprachfördereinheiten. In: Aguado, K./ Schramm, K./ Vollmer, H. J. (Hrsg.): Fremdsprachliches Handeln beobachten, messen, evaluieren. Reihe: KFU, Band 37. Frankfurt a.M.: Peter Lang, S. 257-275.
- Roos, Jeanette/ Polotzek, Silvana/ Schöler, Hermann (2010): EVAS Evaluationsstudie zur Sprachförderung von Vorschulkindern. Abschlussbericht der Wissenschaftlichen Begleitung der Sprachfördermaßnahmen im Programm „Sag’ mal was – Sprachförderung für Vorschulkinder“. Unmittelbare und längerfristige Wirkungen

von Sprachförderungen in Mannheim und Heidelberg. Heidelberg: Pädagogische Hochschule Heidelberg, Projekt EVAS [http://www.sagmalwasbw.de/media/WiBe%201/pdf/EVAS_Abschlussbericht_Januar_2010.pdf, 10.07.2014].

Vollmer, Helmut Johannes (2011): Schulsprachliche Kompetenzen: Zentrale Diskursfunktionen [http://www.home.uni-osnabrueck.de/hvollmer/VollmerDF-Kurzdefinitionen.pdf, 22.09.2014].

Wirtz, Markus/ Kutschmann, Marcus (2007): Analyse der Beurteilerübereinstimmung für kategoriale Daten mittels Cohens Kappa und alternativer Maße. In: Rehabilitation, Jahrgang 46, Heft 3. Stuttgart: Thieme, S. 1-8.