

Erläuterungen (Teil 1) zur Bekanntmachung des Bundesministeriums für Familie, Senioren, Frauen und Jugend (BMFSFJ) sowie des Bundesministeriums für Bildung und Forschung (BMBF) von Richtlinien zur Förderung von Vorhaben zur Evaluation von Konzepten und Maßnahmen der Sprachförderung, Sprachdiagnostik und Leseförderung – Evaluationsvorhaben der zweiten Runde – im Rahmen der Bund-Länder-Initiative „Bildung durch Sprache und Schrift (BiSS)“ vom 12. Januar 2015

URL: http://www.bmbf.de/pubRD/BISS_Evaluationskonzept_12_2014.pdf

Evaluationskonzept der Bund-Länder-Initiative „Bildung durch Sprache und Schrift (BiSS)“

Stand: Dezember 2014

Das Evaluationskonzept wurde vom Trägerkonsortium der Bund-Länder-Initiative „Bildung durch Sprache und Schrift“ erarbeitet.¹ Der Lenkungsausschuss der gemeinsamen Initiative nahm das Evaluationskonzept am 25. Februar 2014 zustimmend zur Kenntnis und beschloss, dem Text eine Präambel voranzustellen. Die hier wiedergegebene, aktualisierte Fassung des Evaluationskonzepts wurde vom Lenkungsausschuss am 12. Dezember 2014 gebilligt.

Präambel

„Insbesondere für Konzepte und Verbünde, die im Rahmen der ersten und zweiten Runde in keine externe Evaluation einbezogen werden konnten, wird gemeinsam mit dem Lenkungsausschuss unter erneuter politischer Beratung nach Möglichkeiten gesucht, um auch diesen die Teilnahme an einer externen Evaluation zu ermöglichen.“

1. Einleitung

In den vergangenen Jahren wurde in allen Ländern – insbesondere in Reaktion auf die Ergebnisse der ersten PISA- und IGLU/PIRLS-Studien – die Konzeption und Umsetzung von Maßnahmen zur Sprachförderung², Sprachdiagnostik und Leseförderung im Bereich der institutionellen Bildung angestoßen. Wie weit die dabei entstandenen und im pädagogischen Alltag umgesetzten Konzepte entwickelt sind, ist jedoch in vielen Fällen unklar. Ebenso steht eine systematische Überprüfung der Durchführungsqualität und Wirksamkeit der Ansätze in der Regel noch aus. Ein zentrales Ziel des Forschungs- und Entwicklungsprogramms „Bildung durch Sprache und Schrift“ (BiSS) besteht deshalb in der Weiterentwicklung und Erprobung von Maßnahmen zur Sprachförderung, Sprachdiagnostik und Leseförderung sowie in der wissenschaftlichen Überprüfung der Wirksamkeit dieser Ansätze (vgl. Steuerungsgruppe Feststellung der Leistungsfähigkeit des Bildungswesens im internationalen Vergleich, 2012). Die Grundlage dafür bildet ein Evaluationskonzept, das ausgehend von den Möglichkeiten wissenschaftlicher Evaluationen ein Vorgehen für die Prüfung der in den Verbänden umgesetzten Maßnahmen vorschlägt. Dieses Evaluationskonzept wird damit maßgeblich zur Qualitätsentwicklung und -sicherung von Konzepten zur Sprachförderung, Sprachdiagnostik und Leseförderung der Bund-Länder-Initiative beitragen.

1.1 Möglichkeiten und Ziele wissenschaftlicher Evaluation

Evaluationen (lat. *valere*: wert sein) beschreiben Gegenstände oder Sachverhalte mit dem Ziel einer Bewertung. Im Gegensatz zu alltäglichen Wertungen (z.B. „Dieses Restaurant ist gut.“) werden bei wissenschaftli-

¹ An der Erarbeitung des Evaluationskonzepts waren folgende Wissenschaftlerinnen und Wissenschaftler beteiligt: Dr. Sofie Henschel, Prof. Dr. Petra Stanat, Prof. Dr. Michael Becker-Mrotzek, Prof. Dr. Marcus Hasselhorn, Prof. Dr. Hans-Joachim Roth

² Der Begriff der Sprachförderung ist im Elementar- und im Schulbereich unterschiedlich konnotiert. Im Zusammenhang mit BiSS umfasst der Begriff jeweils sowohl die alltagsintegrierte sprachliche Bildung als auch die kompensatorische Zusatzförderung für Kinder mit besonderen Bedarfen.

chen Evaluationen sozialwissenschaftliche Forschungsmethoden zur systematischen Untersuchung der Konzeption, Ausgestaltung, Umsetzung und des Nutzens oder der Güte eines Evaluationsgegenstandes angewendet (Mittag & Hager, 2000). Die erzielten Ergebnisse, Schlussfolgerungen und Empfehlungen müssen nachvollziehbar auf empirisch gewonnenen qualitativen und/oder quantitativen Daten beruhen (DeGEval, 2002). Evaluiert werden kann nahezu alles (z.B. Produkte, Methoden, Systeme). Im Bereich der pädagogisch-psychologischen Evaluationsforschung sind häufig Förderkonzepte von Interesse, die Bündel von Maßnahmen umfassen und bestimmte Zielzustände (z.B. eine Leistungsverbesserung) anstreben. Die Grundlage für eine erfolgreiche Evaluation von Maßnahmen besteht zunächst darin, während der Konzeption der Maßnahme festzulegen, welcher Bereich gefördert werden soll, welche Ziele mit der Förderung angestrebt werden, wie die Maßnahme inhaltlich ausgestaltet werden kann, mit welchen Strategien die Umsetzung erfolgen sollte und welche diagnostischen Verfahren geeignet sind, um den Erfolg der Maßnahme valide abzuschätzen (vgl. Hager & Hasselhorn, 1995).

Prinzipiell können Evaluationsvorhaben entweder absolut betrachtet oder im Vergleich zu einer Alternativmaßnahme Erkenntnisse über Eigenschaften, Stärken und Schwächen in der Umsetzung, Effektivität (wie wirksam ist die Maßnahme?) oder Effizienz (welche Kosten-Nutzen-Bilanz besteht?) eines Konzeptes liefern. Die ermittelten Erkenntnisse können dazu genutzt werden, einen Entscheidungsprozess bezüglich der Auswahl von alternativen Vorgehensweise zu erleichtern (Wottawa & Thierau, 1998), um beispielsweise die Ausgestaltung und Umsetzung zu optimieren, zu legitimieren (z.B. gegenüber Geldgebern) oder darüber zu entscheiden, ob eine Maßnahme fortgeführt oder in größerem Umfang implementiert werden sollte (Bortz & Döring, 2006).

Bei der Planung eines Evaluationsvorhabens ist grundsätzlich festzulegen, (a) welcher *Bereich* eines Konzeptes, (b) mit welchem *Ziel* und zu welchem *Zeitpunkt*, in (c) welchem *Umfang*, im (d) *Vergleich* wozu bewertet wird und (e) welche Rolle dem *Evaluator* dabei zukommt (vgl. Benit, 2013; Bortz & Döring, 2006; Hager & Hasselhorn, 1995; Mittag & Hager, 2000; Stufflebeam, 1983):

- a) **Bereich.** Bei einer Evaluation kann schwerpunktmäßig die Bewertung des *Konzeptes* an sich (entspricht es dem wissenschaftlichen Kenntnisstand?), die *Erprobung und Durchführung* des Konzepts (wird es planmäßig und vollständig entsprechend der Konzeption umgesetzt?) oder die *Wirksamkeit* des Konzepts (ist es effektiv und effizient?) im Vordergrund stehen.
- b) **Ziel und Zeitpunkt.** *Formativ* angelegte Evaluationen werden während der Entwicklung bzw. Durchführung der Maßnahme vorgenommen und überprüfen, ob sich die Anwendung unter Alltagsbedingungen bewährt. Das Ziel besteht darin, Schwächen in der Maßnahmenkonzeption oder in deren Umsetzung zu identifizieren und das Konzept auf Basis dieser Ergebnisse für die Zielerreichung zu optimieren. *Summative* Evaluationen bewerten ein Konzept dagegen nach dessen Umsetzung mit dem Ziel, Aussagen darüber zu treffen, wie effektiv und effizient dieses insgesamt ist. Folglich sind Prozessevaluationen in der Regel *formativ* angelegt, *Wirksamkeitsevaluationen* hingegen *summativ*.
- c) **Umfang.** Bei *globalen* Evaluationen werden Maßnahmen als Ganzes einer Bewertung bezüglich ihrer Durchführbarkeit und/oder Wirksamkeit unterzogen. Werden nur einzelne Elemente der Maßnahme (z.B. Vermittlungsstrategien, Materialien, Aufgaben) bewertet, handelt es sich um eine *analytische* Evaluation, mit der auch bestimmt werden kann, welche Komponenten für die Wirksamkeit verantwortlich sind.
- d) **Vergleich.** Bei *Wirksamkeitsevaluationen* kann weiterhin eine *isolierte* Bewertung im Vergleich mit einer Kontrollgruppe erfolgen, die nicht an einer Maßnahme mit dem gleichen Ziel teilnimmt. Um zu entscheiden, welches von verschiedenen Konzepten effektiver ist, kann zudem eine *vergleichende* Be-

wertung der Wirksamkeit gegenüber einer Alternativmaßnahme sinnvoll sein, das auf denselben Bereich abzielt (z.B. Förderung von Leseflüssigkeit) und unter vergleichbaren Randbedingungen (z.B. Dauer, Durchführungsmodus) stattfindet.

- e) **Evaluator.** Evaluationen werden als *interne Selbstevaluation* bezeichnet, wenn sie von den Entwicklern oder Anwendern einer Maßnahme selbst durchgeführt werden. *Externe Fremdevaluationen* sind hingegen durch die Unabhängigkeit der Evaluatoren (die nicht Mitglied bzw. Mitarbeiter der Organisation sind) vom Untersuchungsgegenstand gekennzeichnet.

Diese fünf Dimensionen sind nicht vollständig unabhängig voneinander. So können Evaluationen mit einem formativen Charakter sowohl als Selbst- als auch als Fremdevaluation durchgeführt werden. Das Ziel besteht bei formativ ausgerichteten Evaluationen stets darin, Prozess- und Steuerungswissen zu generieren, um die Ausgestaltung und Durchführung des Konzeptes im Hinblick auf die Zielsetzung der Maßnahme und die Anwendbarkeit (unter alltagspraktischen Bedingungen) zu optimieren. Evaluationen, die primär auf die Wirksamkeit und Effizienz eines Konzeptes abzielen, erfolgen häufig nach Abschluss der Maßnahme als summative Bewertung (Bortz & Döring, 2006).

Zu berücksichtigen ist, dass die Identifikation von Wirksamkeitseffekten besondere Herausforderungen an das Evaluationsdesign stellt. Ob eine Maßnahme entsprechend der theoretischen Erwartung wirksam ist und der identifizierte Effekt tatsächlich aufgrund der durchgeführten Maßnahme und nicht aufgrund anderer Einflussgrößen verursacht wurde, erfordert die Umsetzung eines Kontrollgruppendesigns. Dabei werden Personen idealerweise zufällig (randomisiert) einer Untersuchungsbedingung zugewiesen, in der die Maßnahme angewendet oder nicht angewendet wird (Kontrollgruppe). Durch ein solch streng experimentelles Vorgehen werden personenbezogene Störvariablen minimiert und die eindeutige Interpretierbarkeit der Ergebnisse (interne Validität) erhöht, d.h. Unterschiede zwischen den Untersuchungsbedingungen lassen sich mit hoher Sicherheit auf die evaluierte Maßnahme zurückführen. Allerdings ist man bei einem solchen Vorgehen oft gezwungen, die Studie in einem unnatürlichen Setting durchzuführen, wodurch die Generalisierbarkeit der Ergebnisse (externe Validität) auf natürliche Situationen eingeschränkt ist. Wenn eine randomisierte Gruppenzuweisung unter natürlichen Bedingungen nicht möglich oder extrem aufwändig ist, stellen quasiexperimentelle Kontrollgruppendesigns eine angemessene Alternative dar, bei der natürliche Gruppen (z.B. Schulklassen, Kitagruppen) verglichen werden.

Wie nachhaltig sich eine Maßnahme auf die Entwicklung von Handlungskompetenzen auswirkt, wird durch Hypothesen über die Dauerhaftigkeit der Wirksamkeit (*Transferweite*) überprüft. In der Regel wird dazu ein Prä-, Post- Follow-Up Design umgesetzt. In einem *Prätest* wird zunächst der Entwicklungsstand vor Beginn der Maßnahme diagnostiziert, im *Posttest* erfolgt eine Messung direkt nach Abschluss der Maßnahme und in einem *Follow-Up Test* eine erneute Messung zu einem späteren Zeitpunkt nach Beendigung der Maßnahme. Der Vergleich von Prätest und Posttest erlaubt lediglich Aussagen über die kurzfristige Wirksamkeit, während der Vergleich von Prätest und Follow-Up auf längerfristige Effekte schließen lässt. Der Vergleich von Posttest und Follow-Up liefert schließlich Hinweise darüber, ob die Maßnahme auch nachhaltig wirksam ist und einen Entwicklungsanschub leistet.

Darüber hinaus ist es sinnvoll, Hypothesen über die *Transferbreite* der Wirksamkeit auf nahe oder ferne Leistungs- oder auch Alltagsbereiche (*Situationstransfer*) bzw. auf die Bewältigung ähnlicher oder unähnlicher Aufgabenanforderungen (*Anforderungstransfer*) zu prüfen. Dabei kann es aufschlussreich sein, neben Leistungsmerkmalen auch Einstellungs- und Verhaltensmerkmale als Kriteriumsvariablen zu berücksichtigen und zu untersuchen, inwiefern die jeweilige Maßnahme bestimmte beabsichtigte oder unbeabsichtigte Nebenwirkungen hervorruft, die praktisch bedeutsam sein könnten. So wird mit der Förderung des Lesever-

ständnisses zum Beispiel häufig auch das Ziel verbunden, die intrinsische Lesemotivation oder das Leseverhalten in der Freizeit zu verbessern (vgl. Gold, Trenk-Hinterberger & Souvignier, 2009).

1.2 Standards der Evaluation

Um sicherzustellen, dass die Interessen der Auftraggeber und Nutzer einer Evaluation angemessen berücksichtigt werden, müssen Evaluationen wissenschaftlichen Standards genügen, die von der Deutschen Gesellschaft für Evaluation auf Basis der *Standards des Joint Committee* explizit geregelt wurden (DeGEval, 2002; Sanders, 2000). Die Standards der Evaluation, die auch im Rahmen des BiSS-Programms die Grundlage für alle Evaluationsvorhaben bilden, umfassen die Aspekte (a) *Nützlichkeit*, (b) *Durchführbarkeit*, (c) *Korrektheit/Fairness* und (d) *Genauigkeit*.

- a) **Nützlichkeit.** Evaluationen sollen den tatsächlichen Informationsbedürfnissen der Adressaten gerecht werden, die Ergebnisse sollen verständlich und klar aufbereitet sein.
- b) **Durchführbarkeit.** Evaluationen sollen realistisch, gut durchdacht, diplomatisch und wirtschaftlich sowie konfliktvermeidend angelegt sein.
- c) **Korrektheit/Fairness.** In einer Evaluation soll respektvoll und fair mit den betroffenen Personen und Gruppen umgegangen werden, d.h. dass rechtliche und ethische Standards eingehalten, Ergebnisse vollständig, verständlich und fair dargestellt sowie allen Beteiligten zugänglich gemacht werden.
- d) **Genauigkeit.** Evaluationen sollen präzise (valide, reliable) Informationen und Ergebnisse zu Evaluationsfragestellungen hervorbringen und vermitteln. Dies beinhaltet auch die Analyse qualitativer und quantitativer Informationen sowie die systematische Bestimmung von Schätzfehlern.

2. Übergeordnete Ziele der Evaluationen in BiSS

Das Ziel der Evaluationen in BiSS besteht darin, eine Abstimmung, gemeinsame Weiterentwicklung und Optimierung von Ansätzen der Sprachförderung zu unterstützen, die als potentiell wirksam einzuschätzen sind. Es sollen Anhaltspunkte darauf gewonnen werden, welche Konzepte der Sprachförderung in der pädagogischen Praxis erfolgreich umsetzbar und effektiv sind. Den Evaluationsgegenstand bilden die Konzepte der Sprachförderung. Dabei handelt es sich um aufeinander abgestimmte Maßnahmen zur Professionalisierung, Diagnostik und Förderung, von denen begründet anzunehmen ist, dass sich mit ihnen bestimmte Ziele erreichen lassen (z.B. Förderung: Steigerung der Leseflüssigkeit mit Lesetandems). Idealerweise sind die in den Verbänden umgesetzten Konzepte theoretisch fundiert, liegen gut beschrieben vor und werden systematisch umgesetzt, sodass die Wirksamkeit der Ansätze aufgrund alltagspraktischer Erfahrungen naheliegend ist bzw. sich das Gesamtkonzept oder Bestandteile des Ansatzes bereits unter kontrollierten Bedingungen empirisch bewährt haben (vgl. Steuerungsgruppe, 2012). Im Verlauf des Programms hat sich gezeigt, dass diese Erwartungen an den Entwicklungsstand der Konzepte nur teilweise erfüllt werden. Aus dem sehr heterogenen Entwicklungsstand der Konzepte ergeben sich Konsequenzen für die Art der Evaluation der Ansätze, um nutzbare Erkenntnisse für die Weiterentwicklung und Umsetzung von Sprachförderkonzepten in Deutschland zu gewinnen.

3. Entwicklungsstand der Konzepte in den Verbänden und Konsequenzen für die Evaluation in BiSS

Die Konzepte, die in BiSS entwickelt und umgesetzt werden, lassen sich drei Entwicklungsstufen zuordnen. Diese sind zwar nicht vollständig trennscharf voneinander, sie verdeutlichen aber die Heterogenität des Entwicklungsstands:

I. Neuentwicklung: Der Verbund hat seine Arbeit erst kürzlich aufgenommen. Es liegen noch sehr vage und primär einrichtungsspezifische Vorstellungen über die Ziele und die Ausgestaltung (Auswahl und Anwendung von Tools) des Konzeptes vor.

Das Ziel besteht auf dieser Entwicklungsstufe darin, während der Programmlaufzeit eine gemeinsame und verbindliche Ausarbeitung, Abstimmung und Beschreibung der Konzeption im Verbund vorzunehmen (z.B. Konkretisierung der Ziele, systematische Weiterentwicklung und Anwendung der Tools), Bausteine des Konzeptes in der pädagogischen Praxis zu erproben und ausgewählte Teilziele der Diagnose und Förderung zu überprüfen. Auf Basis der dabei gewonnenen Erkenntnisse sollten konzeptuelle Schwächen identifiziert und eine erste Optimierung des Konzeptes vorgenommen werden, bevor eine erneute Phase der Zielüberprüfung erfolgt.

II. Weiterentwicklung: Der Verbund knüpft an Vorgängerprojekte an oder arbeitet seit längerem zusammen. Es liegen weitgehend konkrete Vorstellungen über die Ausgestaltung und die Umsetzung des Konzeptes vor, wobei sich die einzelnen Verbundeinrichtungen häufig noch auf ein gemeinsames Vorgehen einigen müssen.

Für Konzepte auf dieser Entwicklungsstufe besteht das Ziel im BiSS-Programm darin, eine gemeinsame und verbindliche Abstimmung der Ziele und Umsetzungsstrategien im Verbund vorzunehmen. Neben der Erprobung zentraler Bausteine des Ansatzes und einer darauf abgestimmten Optimierung, kann ein weiteres Ziel darin bestehen, Bedingungsfaktoren der Implementationsqualität während der Durchführung des Gesamtvorhabens prozessbegleitend zu untersuchen und dabei auch eine Überprüfung des unmittelbaren Maßnahmenverlaufs vorzunehmen (z.B. werden mit der vorangegangenen Professionalisierung Verhaltensänderungen bei pädagogischen Fachkräften erzielt?).

III. Optimierung: Das Konzept liegt weitgehend ausgearbeitet vor und zentrale Bestandteile des Ansatzes wurden in Vorgängerprojekten unter kontrollierten Bedingungen untersucht bzw. die systematische Umsetzung des Ansatzes ist im Verbund mindestens in Teilen etabliert. Häufig werden einzelne Bereiche des Ansatzes um neue Bausteine erweitert (z.B. sollen spezifische Tools zur Diagnostik und Förderung bei Kindern mit nicht-deutscher Herkunftssprache systematisch angewendet werden).

Das Ziel besteht darin, das Gesamtkonzept unter alltagspraktischen Bedingungen systematisch und routiniert umzusetzen. Neue Konzeptbausteine (z.B. Tools für bestimmte Zielgruppen) sollten im Zuge dessen erprobt werden und zu einer Feinabstimmung der Konzeption führen. Bei der formativen wissenschaftlichen Begleitung der Durchführung sollten während des Maßnahmenverlaufs Erkenntnisse über die unmittelbare Wirksamkeit und soweit wie möglich auch über die mittelbare Wirksamkeit (z.B. Entwicklung von sprachlichen und fachlichen Kompetenzen der Kinder bzw. Jugendlichen) des Ansatzes gewonnen werden.

In Abbildung 1 ist dargestellt, welchen Entwicklungsstufen die Konzepte anhand der vorliegenden Dokumente (Antrag: Februar 2014, Fragebogen: Juli-November 2014) zuzuordnen sind. Dabei wird deutlich, dass die Konzepte im Elementarbereich tendenziell einen etwas höheren Entwicklungsstatus aufweisen als die Konzepte im Schulbereich. Während in der Primarstufe viele Ansätze vollständig neu entwickelt werden, erfolgt in der Sekundarstufe häufig eine umfassende Weiterentwicklung auf Basis von Vorgängerprojekten.

Um die Verbünde in ihren aktuellen Entwicklungsprozessen zu unterstützen, sind zwei Formen der Evaluation vorgesehen, die nachfolgend beschrieben werden.

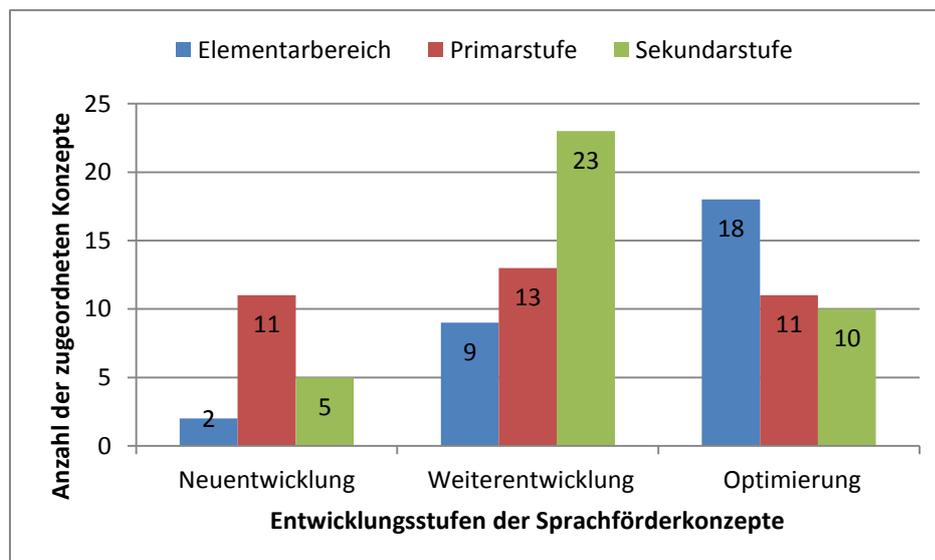


Abbildung 1. Zuordnung der Sprachförderkonzepte zu den Entwicklungsstufen; Stand: 26.11.2014.

3.1 Selbstevaluation

Das Ziel der Selbstevaluation besteht darin, die pädagogische Praxis möglichst unmittelbar aufgrund systematisch gewonnener Daten zu verbessern. Dazu werden ausgewählte Ziele eines Gesamtkonzeptes überprüft, die von den Durchführenden selbst als wichtig für ihre pädagogische Praxis erachtet werden. Ein großer Vorteil der Selbstevaluation ist darin zu sehen, dass sie sehr praxisorientiert ausgerichtet ist, d.h. die Evaluationsziele, die Fragestellungen, das Design, die Umsetzung und die Verwendung von Evaluationsergebnissen werden durch die professionell Handelnden selbst gesteuert. Die Anknüpfung am Wissen und an den Erfahrungen der in den Verbänden professionell Tätigen kann als Grundlage für die Planung und die Durchführung der Evaluation dienen. Selbstevaluationen können zentraler Bestandteil eines kontinuierlichen Qualitätsentwicklungsprozesses sein, wenn aus den Ergebnissen Handlungskonsequenzen für die Praxis abgeleitet werden und diese erneut als Ausgangspunkt für eine Zielüberprüfung dienen. Dies wird durch die aktive Beteiligung und die Identifikation der Durchführenden mit dem Selbstevaluationsprozess unterstützt. Der Nutzen einer Selbstevaluation besonders hoch, wenn die Entwicklung und Erprobung des Konzeptes noch nicht abgeschlossen sind (Entwicklungsstufen I und II). Auf diese Weise können Faktoren identifiziert werden, die die Wirksamkeit des Ansatzes einschränken und eine Modifikation des Konzeptes oder der damit zusammenhängenden Umsetzungsprozesse oder Kontextfaktoren erfordern. Bei sehr weit entwickelten Konzepten, die bereits umfassend erprobt oder sogar unter kontrollierten Bedingungen untersucht wurden, können Erkenntnisse aus der Selbstevaluation hilfreiche Hinweise für die Feinjustierung liefern. Dies gilt vor allem dann, wenn ein Konzept um neue Bausteine erweitert wird, mit denen Ziele verbunden sind, deren Erreichung zuvor noch nicht eingehender überprüft wurde.

Die bei der Selbstevaluation eingesetzten Verfahren und Instrumente sollten in ihrer Entwicklung, Anwendung und Auswertung mit möglichst wenig Aufwand flexibel einsetzbar sowie kosteneffizient sein und damit gut in die Alltagsroutine der professionell Handelnden integriert werden können. Die Befähigung zur Planung und Durchführung einer Selbstevaluation und zur Ableitung und Umsetzung konkreter Handlungs-

möglichkeiten aus den Ergebnissen, die für die Optimierung der eigenen Arbeit genutzt werden können, setzt eine Professionalisierung der Akteure voraus.

Die Selbstevaluation wird in BiSS als Chance und Herausforderung für alle Verbünde betrachtet. Bei der Planung, Durchführung und Auswertung ihres Selbstevaluationsvorhabens werden die Verbünde deshalb in regelmäßigen Workshops im Themencluster „Qualitätsmonitoring“ unterstützt. Das Ziel des Clusters besteht darin, den Entwicklungsprozess der Sprachförderkonzepte zu begleiten, ihn in länderübergreifenden Netzwerkgruppen gemeinsam zu reflektieren und die Teilnehmenden darin zu befähigen, eine systematische Selbstevaluation zukünftig selbstständig im Verbund durchzuführen und die erworbenen Kompetenzen zu multiplizieren.

3.2 Externe Evaluation

In die externen Evaluationen werden in BiSS ausgewählte Konzepte einbezogen, die einen fortgeschrittenen Entwicklungsstand erreicht haben. Um eine hohe Reichweite der Evaluationen und ihrer Ergebnisse zu erzielen, beziehen sich diese auf Module bzw. thematische Schwerpunkte, zu denen möglichst viele Verbünde Konzepte entwickeln und umsetzen. Die Schwerpunkte liegen im Elementarbereich auf dem Modul E1 (Gezielte alltagsintegrierte sprachliche Bildung), in der Primarstufe auf den Modulen P1 (Gezielte sprachliche Bildung in alltäglichen und fachlichen Kontexten), P3 (Diagnose und Förderung der Leseflüssigkeit und ihrer Voraussetzungen) und P4 (Diagnose und Förderung des Leseverständnisses) sowie in der Sekundarstufe auf den Modulen S2 (Lese- und Schreibstrategien im Verbund vermitteln), S4 (Sprachliche Bildung in fachlichen Kontexten) und auf der sprachlichen Integration von Seiteneinsteigern ins deutsche Schulsystem. Durch die Fokussierung der Evaluation auf zentrale Themenschwerpunkte können Tools, die sich während der Evaluation als praxistauglich erweisen und in die Tooldatenbank überführt werden, von vielen Verbänden genutzt werden, die vergleichbare Ziele verfolgen. Eine möglichst hohe Reichweite der Evaluation soll durch länderübergreifende Untersuchungsgruppen sichergestellt werden, deren Zusammensetzung mitunter heterogen ausfällt. Denn es werden zwar mit allen Konzepten einer Untersuchungsgruppe die gleichen Ziele verfolgt (z.B. Verbesserung des Leseverständnisses), häufig werden aber verschiedene Tools (z.B. Instrumente, Materialien, Vermittlungsstrategien) unter variierenden Kontextbedingungen in unterschiedlicher Art und Weise angewendet.

In der Evaluation wird der Prozess der Durchführung der Sprachförderansätze formativ durch externe Evaluationsteams begleitet. Im Mittelpunkt steht die entwicklungsorientierte Untersuchung der Implementationsqualität und ihrer Bedingungsfaktoren. Dies betrifft u.a. Fragestellungen zur Ausführungsintegrität (z.B. Werden Verhaltensweisen, Materialien und Organisationsstrukturen systematisch übernommen bzw. angepasst?), der Zielgruppenreichweite (z.B. Wird die Zielgruppe vollständig erreicht?) oder der unmittelbaren Wirksamkeit (z.B. Werden durch eine vorangegangene Professionalisierung Verhaltensänderungen bei pädagogischen Fachkräften erzielt?). Soweit dies unter Berücksichtigung der Heterogenität der einbezogenen Konzepte möglich ist, sollen auch Anhaltspunkte auf die mittelbare Wirksamkeit des Sprachförderansatzes gewonnen werden (z.B. Wie entwickeln sich die Handlungskompetenzen von Kindern nach der Professionalisierung des pädagogischen Fachpersonals?). Zudem sollen Gelingensbedingungen identifiziert werden, die beispielsweise die systematische Anwendung von Tools im pädagogischen Alltag begünstigen. Ebenso sollen die Evaluationen Hinweise darauf liefern, wie der Ansatz in die Organisationsstruktur einer Schule, Kita oder im Verbund integriert und damit verstetigt werden kann. Durch die regelmäßige Rückmeldung von Ergebnissen aus der Prozessbegleitung sollen in engem Austausch zwischen Wissenschaft und Praxis gemeinsam konzeptuelle Entwicklungspotenziale identifiziert, Handlungskonsequenzen für die Opti-

mierung der Ansätze erarbeitet und in die pädagogische Praxis implementiert werden. Dies setzt gut funktionierende Rückmeldeformate und -strategien voraus. Erst wenn die gewonnenen Daten den Bedürfnissen der beteiligten Akteure entsprechend aufbereitet sind, können sie in Prozesswissen umgewandelt und in die Qualitätsentwicklung der Konzepte einfließen.

Tools, die sich in der Evaluation empirisch bewähren, sollen in die BiSS-Tooldatenbank überführt werden und sowohl für die Weiterentwicklung der Sprachförderansätze in den BiSS-Verbänden nutzbar gemacht werden als auch Anknüpfungspunkte für Wissenschaft und Praxis darstellen. Überdies sollen mit den Evaluationsprojekten soweit wie möglich auch Anhaltspunkte auf die Wirksamkeit der Ansätze als Entscheidungsgrundlage für den Transfer von Konzepten in die Fläche genutzt werden.

Vorhaben der externen Evaluation in BiSS werden für den Elementarbereich vom Bundesministerium für Familie, Senioren, Frauen und Jugend (BMFSFJ) und für den Schulbereich vom Bundesministerium für Bildung und Forschung (BMBF) auf der Grundlage gemeinsamer Bekanntmachungen von Förderrichtlinien gefördert.

4. Vorbereitung der externen Evaluationen

Um eine möglichst große Anzahl von Konzepten und Verbänden in die Evaluationen einbeziehen zu können, ist für die öffentliche Bekanntmachung der Förderrichtlinien und die Umsetzung der Evaluationsprojekte ein zweistufiges Vorgehen vorgesehen, das eine erste und eine zweite Evaluationsrunde umfasst.

In der ersten Runde werden auf der Grundlage der Bekanntmachung der genannten Bundesressorts vom 26. Mai 2014 (<http://www.bmbf.de/foerderungen/23874.php>) diejenigen Konzepte in etappenspezifischen Evaluationsprojekten prozessbegleitend untersucht, die zu Beginn des Jahres 2014 einen möglichst fortgeschrittenen Entwicklungsstand aufweisen (mind. Entwicklungsstufe II) bzw. einem Themenschwerpunkt zugeordnet werden können, das eine hohe bildungspolitische Relevanz aufweist. Dies betrifft im Elementarbereich Konzepte zur Professionalisierung pädagogischer Fachkräfte für eine gezielte alltagsintegrierte Sprachbildung (Modul E1), in der Primarstufe Konzepte der Sprachförderung im Fachunterricht (Modul P1) und in der Sekundarstufe Sprachförderkonzepte für Schülerinnen und Schülern mit sehr geringen oder fehlenden Deutschkenntnissen (sog. Seiteneinsteiger). Es wird in jeder Bildungsetappe zu den genannten thematischen Schwerpunkten ab dem 1.1.2015 ein Evaluationsprojekt mit einer dreijährigen Laufzeit gefördert. Im Mittelpunkt dieser Evaluationsvorhaben steht die entwicklungsorientierte Prozessbegleitung bei der Durchführung von Sprachförderkonzepten mit denen vergleichbare Ziele erreicht werden sollen. Das Ziel der Evaluationen besteht darin, Bedingungsfaktoren der Implementationsqualität dieser Konzepte zu untersuchen, um die Ansätze für eine erfolgreiche Anwendung in der pädagogischen Praxis zu optimieren. Wenn die Effektivität der Ansätze untersucht wird, betrifft dies vorrangig unmittelbare Wirksamkeit (z.B. Ändern sich professionelle Verhaltensweisen von pädagogischen Fachkräften aufgrund einer vorangegangenen Qualifizierung?). Die Evaluationen der ersten Runde werden im Zeitplan als *Prozessevaluationen* bezeichnet.

In der zweiten Runde sollen mindestens je zwei weitere Evaluationsprojekte pro Bildungsetappe zu zentralen thematischen Schwerpunkten realisiert werden. Die Evaluationen werden voraussichtlich ab dem 3. Quartal 2015 mit einer dreijährigen Laufzeit durchgeführt. Es sollen Konzepte einbezogen werden, bei denen zu erwarten ist, dass bis zum vorgesehenen Beginn der Evaluation eine umfangreiche Weiterentwicklung des Sprachförderansatzes erfolgt und ein fortgeschrittener Entwicklungsstand erreicht wird (mind. Entwicklungsstufe II). Dies schließt die konzeptuelle Ausarbeitung, die gemeinsame und verbindliche Abstimmung einer systematischen Umsetzung im Verbund, die Erprobung von Bausteinen des Ansatzes und die Über-

prüfung ausgewählter Ziele des Konzeptes im Rahmen einer Selbstevaluation ein. Die Konzepte müssen überdies eine ausreichende thematische Passung zu den Schwerpunkten der Evaluation aufweisen. Welche Schwerpunkte gegenwertig mit einer ausreichenden Anzahl an Konzepten länderübergreifend untersucht werden können, sofern die dafür erforderlichen finanziellen Mittel bereitgestellt werden, ist in Tabelle 1 dargestellt. Soweit wie möglich sollen in diesen Evaluationsstudien neben Faktoren der Implementationsqualität sowohl die unmittelbare Wirksamkeit der Konzepte als auch mittelbarer Effekte untersucht werden.

Tabelle 1. Empfehlung für thematische Schwerpunkte der zweiten Evaluationsrunde

Elementarbereich	Primarstufe	Sekundarstufe
Gezielte alltagsintegrierte sprachliche Bildung in Schlüsselsituationen für Kinder zwischen 2 und 6 Jahren (Modul E1)	Sprachförderung in alltäglichen und fachlichen Kontexten (P1)	Sprachförderung im Sach- und Fachunterricht (S4)
Alltagsintegrierte Sprachbildung beim Übergang vom Elementar- zum Primarbereich (Module E6, E1)	Fachübergreifende Sprachförderung (P3, P4)	Leseförderung im Sach- und Fachunterricht (S2)
Alltagsintegrierte sprachliche Bildung für Kinder unter 3 Jahren (Modul E2)		

5. Funktion des Trägerkonsortiums während der Durchführung der Evaluationsvorhaben

Im Rahmen seiner Zuständigkeit für die Gesamtkoordination des Programms ist das Trägerkonsortium für die Abstimmungsprozesse mit dem Lenkungsausschuss verantwortlich, die im Zusammenhang mit der externen Evaluation erforderlich sind. Die Zuständigkeit für die administrative Koordination und Abstimmung der vom Bund geförderten Evaluationsprojekte liegt bei Vorhaben, die vom BMFSFJ gefördert werden, bei der Regiestelle und bei den vom BMBF geförderten Vorhaben beim Projektträger im DLR.

Die vom Bund geförderten Evaluationsvorhaben werden durch das Trägerkonsortium koordiniert. Unter anderem werden pro Jahr bis zu zwei Netzwerktreffen zwischen dem Trägerkonsortium und den Evaluationsprojekten stattfinden. Dabei wird es einerseits um den Austausch des aktuellen Arbeitsstandes gehen, andererseits soll ein gemeinsames Rückmeldekonzept entwickelt werden, auf dessen Basis die Projektergebnisse auch für diejenigen Verbände nutzbar gemacht werden können, die nicht an den externen Evaluationen teilnehmen. Eine Erprobung von Rückmeldestrategien durch die Evaluationsteams und das Trägerkonsortium erfolgt im Themencluster „Qualitätsmonitoring“.

Das Trägerkonsortium wird programmbegleitende Beratungen für diejenigen Verbände anbieten, die Evaluationsvorhaben aus eigenen bzw. Landesmitteln umsetzen.

Um die fortlaufende Qualitätsentwicklung der Sprachförderansätze in den Verbänden zu unterstützen, begleitet das Trägerkonsortium die Konzeptweiterentwicklung im Themencluster „Qualitätsmonitoring“, das Bestandteil des Fortbildungsprogramms ist. In diesem etappenübergreifenden Cluster sind zunächst vier zweitägige Veranstaltungen bis zum Sommer 2016 vorgesehen mit denen drei zentrale Ziele verfolgt werden. Erstens erfolgen ein kontinuierlicher Austausch und eine Reflektion über den Entwicklungsverlauf der

Sprachförderkonzepte in den Verbänden. Um eine intensive Zusammenarbeit zwischen den Verbänden zu ermöglichen, sollen dazu länderübergreifende Netzwerkstrukturen in Schwerpunktgruppen aufgebaut werden. Dies betrifft im Schulbereich z.B. Verbände, deren Konzepte primär die Sprach- oder die Leseförderung vorsehen. Das zweite Ziel besteht darin, Multiplikationspotenzial in Workshops zur Selbstevaluation zu entwickeln, indem die Teilnehmenden punktuell bei der Planung, Durchführung und Auswertung ihres Selbstevaluationsvorhabens unterstützt und darauf vorbereitet werden, die dabei erworbenen Kompetenzen später selbstständig im Verbund anzuwenden und weiterzugeben. Beim dritten Ziel geht es schließlich darum, die Ergebnisse aus den externen Evaluationsprojekten im Rahmen des Clusters vorzustellen, gemeinsam zu reflektieren sowie Strategien zu erarbeiten, wie diese Erkenntnisse für die Weiterentwicklung für alle Konzepte in BiSS sowie für die Sprachförderung in den Ländern allgemein nutzbar gemacht werden können.

Anfang 2016 ist zudem eine onlinegestützte Befragung in den Verbänden vorgesehen, bei der es um die Akzeptanz, Zufriedenheit und Nutzung der durch das BiSS-Trägerkonsortium bereitgestellten Angebote (z.B. Beratungen, Tools, Fortbildungen) gehen wird. Das Ziel dieser Befragung besteht darin, diese Angebote für den weiteren Programmverlauf bedarfsgerecht zu optimieren.

6. Kostenkalkulation für die Evaluationsvorhaben

Für die Evaluationsvorhaben wurden in der BiSS-Expertise jährlich 1,8 Mio. € kalkuliert. Bei einer Programmlaufzeit von fünf Jahren entspricht dies einem Gesamtetat von 9 Mio. €. Für jedes Evaluationsprojekt wurden in der Expertise jährlich bis zu 300.000 € veranschlagt. Es wird angestrebt, mindestens drei Evaluationsprojekte pro Bildungsetappe umzusetzen. Die tatsächlich benötigten finanziellen Ressourcen werden zwischen den einzelnen Projekten variieren und beispielsweise von der Größe der untersuchten Stichproben und der Komplexität des Untersuchungsdesigns abhängen.

7. Darstellung und Verwertung der Ergebnisse aus den Evaluationsvorhaben

Da die Durchführung der Evaluationsprojekte den üblichen Standards der Evaluation entsprechen sollte (vgl. DeGEval, 2002), müssen sich alle Evaluationsprojekte dazu verpflichten, die Ergebnisse ihrer Studien vollständig, fair und in geeigneten Formaten an die beteiligten Akteure in den Verbänden zurückzumelden. Grundsätzlich sollten die Ergebnisse aus den Evaluationen dazu genutzt werden, konkrete Handlungskonsequenzen für die Optimierung der Konzepte abzuleiten und umzusetzen. Darüber hinaus sollten die Ergebnisse in wissenschaftlichen Fachzeitschriften sowie gemeinsam mit dem Trägerkonsortium im Rahmen eines Herausgeberbandes publiziert werden.

Literatur

- Benit, N. (2013). Externe Evaluation. In M. A. Wirtz (Hrsg.), *Dorsch – Lexikon der Psychologie* (16. Auflage, S. 505). Bern: Verlag Hans Huber.
- Bortz, J. & Döring, N. (2006). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler* (4. Auflage). Berlin: Springer.
- DeGEval – Deutsche Gesellschaft für Evaluation e. V. (2002). *Standards für Evaluation*. Köln: Geschäftsstelle DeGEval.
- Gold, A., Trenk-Hinterberger, I. & Souvignier, E. (2009). "Die Textdetektive" - Ein strategieorientiertes Programm zur Förderung des Leseverständnisses. In W. Lenhard & W. Schneider (Hrsg.), *Diagnostik und Förderung des Leseverständnisses* (S. 207-226). Göttingen: Hogrefe.
- Hager, W. & Hasselhorn, M. (1995). Konzeption und Evaluation von Programmen zur kognitiven Förderung: theoretische Überlegungen. In W. Hager (Hrsg.), *Programme zur Förderung des Denkens bei Kindern. Konstruktion, Evaluation und Metaevaluation* (S. 41-85). Göttingen: Hogrefe.
- Mittag, W. & Hager, W. (2000). Ein Rahmenkonzept zur Evaluation psychologischer Interventionsmaßnahmen. In W. Hager, J.-L. Patry & H. Brezing (Hrsg.), *Evaluation psychologischer Interventionsmaßnahmen. Standards und Kriterien: Ein Handbuch* (S. 102-128). Bern: Huber.
- Redder, A., Schwippert, K., Hasselhorn, M., Forschner, S., Fickermann, D. & Ehlich, K. (2011). *Bilanz und Konzeptualisierung von strukturierter Forschung zu „Sprachdiagnostik und Sprachförderung“*. ZUSE-Berichte Band 2. Hamburg: Hamburger Zentrum für Unterstützung der wissenschaftlichen Begleitung und Erforschung schulischer Entwicklungsprozesse (ZUSE).
- Sanders, J. R. (Hrsg.). (2000). *Handbuch der Evaluationsstandards. Die Standards des "Joint Committee on Standards for Educational Evaluation"* (2. Auflage). Opladen: Leske + Budrich.
- Schneider, W., Baumert, J., Becker-Mrotzek, M., Hasselhorn, M., Kammermeyer, G., Rauschenbach, T., Roßbach, H.-G., Roth, H.-J., Rothweiler, M. & Stanat, P. (2012). Expertise „Bildung durch Sprache und Schrift (BISS)“ - Bund-Länder-Initiative zur Sprachförderung, Sprachdiagnostik und Leseförderung, abgerufen am 31.10.2013, http://www.bmbf.de/pubRD/BISS_Expertise.pdf
- Steuerungsgruppe „Feststellung der Leistungsfähigkeit des Bildungswesens im internationalen Vergleich“. (2012). Programmskizze "Bildung durch Sprache und Schrift", abgerufen am 1.11.2013, http://www.kmk.org/fileadmin/veroeffentlichungen_beschluesse/2012/2012_10_18-Initiative_Sprachfoerderung_Programmskizze.pdf
- Stufflebeam, D. L. (1983). The CIPP Model for Program Evaluation. In G. F. Madaus, M. Scriven & D. L. Stufflebeam (Eds.), *Evaluation Models: Viewpoints on Educational and Human Services Evaluation*. Boston: Kluwer Nijhof.
- Wottawa, H. & Thierau, H. (1998). *Lehrbuch Evaluation* (2. überarbeitete Auflage). Bern: Verlag Hans Huber.