



Schriftliche Anfrage

der Abgeordneten **Gabriele Triebel, Julia Post BÜNDNIS 90/DIE GRÜNEN**
vom 08.12.2025

Transparenz und wissenschaftliche Absicherung des Sprachstandserhebungs- verfahrens BaSiS

Die frühzeitige Sprachförderung ist ein wesentlicher Faktor für Bildungs- und Teilhabegerechtigkeit. Die Sprachstandserhebungen der Viereinhalbjährigen erfüllen dabei nicht nur eine diagnostische, sondern auch eine bildungspolitische Steuerungsfunktion, da ihre Ergebnisse Grundlage für verpflichtende Fördermaßnahmen sowie für organisatorische und rechtlich relevante Entscheidungen gegenüber Kindern und Eltern sind.

Vor diesem Hintergrund kommt der fachlichen Qualität, der wissenschaftlichen Fundierung und der Transparenz eingesetzter Testverfahren eine besondere Bedeutung zu. Derzeit wird im Freistaat Bayern das Verfahren BaSiS zur Sprachstandsfeststellung eingesetzt.

Die Staatsregierung wird gefragt:

1. Transparenz und Nachvollziehbarkeit 3
 - 1.1 In welchem Umfang ist eine Einsichtnahme in das vollständige Testmanual sowie in die fachlichen und wissenschaftlichen Begleitunterlagen des Verfahrens BaSiS möglich? 3
 - 1.2 Falls eine vollständige Einsichtnahme nicht vorgesehen ist, aus welchen rechtlichen oder vertraglichen Gründen wurde diese Einschränkung beschlossen? 3
2. Wissenschaftliche Fundierung und Qualitätssicherung 3
 - 2.1 Welche überprüften Gütekriterien (Objektivität, Reliabilität, Validität) liegen dem Verfahren BaSiS zugrunde? 3
 - 2.2 Welche konkreten Kennwerte werden jeweils erreicht und auf welche empirischen Studien, Evaluationen oder wissenschaftlichen Publikationen beziehen sich diese Angaben? 3
 - 2.3 Wie stellt die Staatsregierung sicher, dass bildungspolitisch relevante Entscheidungen nicht auf Verfahren beruhen, deren wissenschaftliche Qualität nicht hinreichend transparent belegt ist? 3
3. Normierung, Vergleichbarkeit und Chancengerechtigkeit 4

3.1	Ist BaSiS ein normiertes Verfahren und, wenn ja, auf welcher Stichprobe (Größe, Altersstruktur, sprachliche und soziale Diversität der getesteten Kinder, Erhebungszeitraum) beruht diese Normierung?	4
3.2	Nach welchen anerkannten wissenschaftlichen Standards wurde die Normierungsstichprobe erhoben und veröffentlicht?	4
3.3	Welche Maßnahmen ergreift die Staatsregierung, um sicherzustellen, dass durch den Einsatz von BaSiS keine systematischen Benachteiligungen einzelner Gruppen entstehen?	5
4.	Evaluation	6
4.1	Welche Alternativen wurden im Vorfeld geprüft und nach welchen fachlichen und bildungspolitischen Kriterien wurde die Entscheidung für BaSiS getroffen?	6
4.2	Sind regelmäßige unabhängige Evaluationen des Verfahrens vorgesehen und, wenn ja, in welchem Turnus und mit welcher Veröffentlichungspraxis?	6
	Hinweise des Landtagsamts	7

Antwort

des Staatsministeriums für Unterricht und Kultus

vom 12.01.2026

1. Transparenz und Nachvollziehbarkeit

1.1 In welchem Umfang ist eine Einsichtnahme in das vollständige Testmanual sowie in die fachlichen und wissenschaftlichen Begleitunterlagen des Verfahrens BaSiS möglich?

1.2 Falls eine vollständige Einsichtnahme nicht vorgesehen ist, aus welchen rechtlichen oder vertraglichen Gründen wurde diese Einschränkung beschlossen?

Die Fragen 1.1 und 1.2 werden aufgrund des Sachzusammenhangs gemeinsam beantwortet.

Wie auch bei anderen wissenschaftlich fundierten Testverfahren üblich, ist das Bayerische Screening des individuellen Sprachstands (BaSiS) zzgl. der in Frage 1.1 benannten Unterlagen nicht öffentlich einsehbar, um eine missbräuchliche oder unsachgemäße Anwendung zu verhindern.

In Bezug auf das Informationsrecht der Abgeordneten wird vorsorglich mitgeteilt, dass dieses Recht mit dem Interesse an der Durchführung der Sprachstandserhebung als staatlicher Aufgabe, die ebenfalls Verfassungsrang genießt, abzuwägen wäre unter Einstellung der Informationsmöglichkeiten, die eine Weitergabe an Dritte und damit ein allgemeines Bekanntwerden ausschließen.

2. Wissenschaftliche Fundierung und Qualitätssicherung

2.1 Welche überprüften Gütekriterien (Objektivität, Reliabilität, Validität) liegen dem Verfahren BaSiS zugrunde?

2.2 Welche konkreten Kennwerte werden jeweils erreicht und auf welche empirischen Studien, Evaluationen oder wissenschaftlichen Publikationen beziehen sich diese Angaben?

2.3 Wie stellt die Staatsregierung sicher, dass bildungspolitisch relevante Entscheidungen nicht auf Verfahren beruhen, deren wissenschaftliche Qualität nicht hinreichend transparent belegt ist?

Die Fragen 2.1 bis 2.3 werden aufgrund des Sachzusammenhangs gemeinsam beantwortet.

Die vorliegenden Kennwerte stützen sich auf die Pilotierungsstudie vom Dezember 2024 und auf die Reanalyse der Daten der ersten Erhebungswelle im Jahr 2025.

Zu den einzelnen Bereichen:

- **Objektivität:** Die Objektivität ist durch die standardisierte, webbasierte Testdurchführung mit aufgabenweiser Präsentation sowie eindeutigen Darbietungs- und Auswertungsregeln gewährleistet. Bei jeder Aufgabe wird neben dem Stimulusmaterial präsentiert, welche Instruktion zu geben ist, wie von der Testleitung nachgefragt werden muss und welche Lösungen von den Kindern als korrekt bewertet werden, sofern die Webanwendung die Bewertung nicht ohnehin automatisch selbst vornimmt. Die Durchführungs- und Auswertungsobjektivität wird somit als gegeben betrachtet. Die Interpretation erfolgt anhand der Normwerte, sodass auch dieses Kriterium als gegeben angesehen wird.
- **Reliabilität:** Angaben zur Messgüte basieren auf der Analyse der Pilotierungsdaten und der Daten der ersten Erhebungswelle. Die interne Konsistenz der Subtests lag in der Pilotierung zwischen $\omega = .884$ und $\omega = .957$ in der populationsweiten Erhebung, für die Gesamtskala bei $\omega = .978$. Die Split-Half-Reliabilität (nach Spearman-Brown) bewegt sich zwischen .863 und .980.
- **Validität:** Die Konstruktvalidität stützt sich auf die theoriegeleitete Auswahl der Subtests (Wortschatz rezeptiv/produktiv, Satzverständnis, Satzspanne/phonologische Schleife), die in der Fachliteratur als zentrale Prädiktoren für Schulerfolg etabliert sind und auf konfirmatorische Faktorenanalysen der Pilotierungsdaten. Weitere Testgütekriterien, die der Entwicklung zugrunde gelegt wurden, sind Testfairness, Ökonomie und Normierung.

Die vorläufige Testgüte mit einer Aufstellung der erreichten Kennwerte und der Beschreibung des Gesamtablaufs der Testentwicklung sind Teil eines Berichts.

Die Pilotstudie wurde im Dezember 2024 mit $n = 361$ Kindern im Alter von 4,0 bis 5,6 Jahren ($m = 4,8$; $sd = 0,31$) durchgeführt. Sie diente sowohl der Itemrevision und Testkonstruktion als auch der Normierung des Instruments. Auf Basis der Pilotdaten wurden Items mit unzureichender Trennschärfe ($rit < .2$), extremen Schwierigkeiten oder signifikanten Geschlechter-DIFs ausgeschlossen (siehe Frage 3.2). Zudem wurden IRT-Analysen (1PL-Modell) durchgeführt und Items mit $Infitt > 1.2$ eliminiert.

Die vorläufige Testgüte mit einer Aufstellung der erreichten Kennwerte und der Beschreibung des Gesamtablaufs der Testentwicklung ist Teil eines Berichts. Eine über die Pilotierung und die erste Erhebungswelle hinausgehende Evaluation soll erfolgen.

3. Normierung, Vergleichbarkeit und Chancengerechtigkeit

3.1 Ist BaSiS ein normiertes Verfahren und, wenn ja, auf welcher Stichprobe (Größe, Altersstruktur, sprachliche und soziale Diversität der getesteten Kinder, Erhebungszeitraum) beruht diese Normierung?

3.2 Nach welchen anerkannten wissenschaftlichen Standards wurde die Normierungsstichprobe erhoben und veröffentlicht?

Die Fragen 3.1 und 3.2 werden aufgrund des Sachzusammenhangs gemeinsam beantwortet.

BaSiS ist ein normiertes Verfahren. Die Normierung basiert auf der in der Antwort zu Fragen 2.1 bis 2.3 beschriebenen Pilotstudie. Hinsichtlich der Heterogenität der Normstichprobe wurden die Variablen Familiensprache, Geschlecht, Regierungsbezirk sowie städtischer bzw. ländlicher Raum erfasst.

Die Normierung von BaSiS folgt modernen psychometrischen Standards, weist jedoch aufgrund der Rahmenbedingungen methodische Besonderheiten und Einschränkungen auf. Für die Normierung wurde auf Continuous Norming (Lenhard, A., Lenhard, W., Suggate, S., Segerer, R. [2018] A Continuous Solution to the Norming Problem. *Assessment*, 25(1), 112–125) zurückgegriffen.

Um die Repräsentativität der Normstichprobe zu erhöhen, wurde eine Post-Stratifizierung mittels *Iterative Proportional Fitting* nach Familiensprache und Geschlecht durchgeführt. Die Gewichtung erfolgte auf Basis der bekannten Verteilungen in der bayerischen Population, wobei für die Familiensprache auf Daten des Bildungsberichts 2024 zurückgegriffen wurde.

3.3 Welche Maßnahmen ergreift die Staatsregierung, um sicherzustellen, dass durch den Einsatz von BaSiS keine systematischen Benachteiligungen einzelner Gruppen entstehen?

Um systematische Benachteiligungen einzelner Gruppen durch den Einsatz von BaSiS zu vermeiden, wurden bereits in der Testentwicklung verschiedene psychometrische Verfahren zur Fairness-Prüfung eingesetzt.

Im Rahmen der Itemanalyse wurden umfassende *Differential Item Functioning* (DIF)-Analysen durchgeführt. Dabei wurde für jedes Item untersucht, ob es bei gleichem Fähigkeitsniveau systematische Unterschiede zwischen Jungen und Mädchen gibt. Items, die signifikante geschlechtsspezifische DIF-Effekte aufwiesen, wurden aus dem finalen Itempool ausgeschlossen.

Zusätzlich wurde zur Überprüfung der Geschlechterfairness auf Skalenebene der Wald-Test im Rahmen von IRT-Analysen eingesetzt. Die finalen Subtests zeigten keine signifikanten Verletzungen der Messinvarianz zwischen Jungen und Mädchen.

Zur Prüfung der grundlegenden Modellannahmen wurde Andersens Likelihood-Ratio-Test durchgeführt. Darüber hinaus wurden Item-Fit-Statistiken berechnet, wobei Items mit Infit-Werten größer als 1.2 ausgeschlossen wurden, da diese auf eine unzureichende Passung zum latenten Konstrukt hinweisen.

Bei der Normierung wurde durch die Post-Stratifizierung nach Familiensprache und Geschlecht sichergestellt, dass alle Gruppen entsprechend ihrem Anteil in der Population berücksichtigt werden. Dies verhindert, dass die Normen durch eine einseitige Stichprobenzusammensetzung verzerrt werden. Die Gewichtung erfolgte auf Basis der bekannten Populationsverteilungen in Bayern.

In der Testdurchführung wurden bewusst Maßnahmen ergriffen, um kulturelle oder sprachproduktionsbedingte Benachteiligungen zu minimieren. So werden bei den produktiven Aufgaben (produktiver Wortschatz, Satzspanne) Aussprachefehler, artikulatorische Probleme, Akzente und dialektale Besonderheiten nicht als falsch gewertet. Die Bewertung fokussiert auf den lexikalischen Gehalt der Antworten, nicht auf die phonetische Realisierung. Bei der Satzspanne wird eine Antwort als korrekt gewertet, wenn alle Wörter in korrekter Reihenfolge wiedergegeben werden, unabhängig von Wortendungen, um morphologische Besonderheiten von Zweitsprachlernenden nicht zu sanktionieren.

Die standardisierte, webbasierte Testdurchführung mit einheitlichen Instruktionen und automatisierter Auswertung gewährleistet zudem, dass alle Kinder unter vergleichbaren Bedingungen getestet werden und Bewertungsunterschiede zwischen verschiedenen Testleitungen minimiert werden.

4. Evaluation

4.1 Welche Alternativen wurden im Vorfeld geprüft und nach welchen fachlichen und bildungspolitischen Kriterien wurde die Entscheidung für BaSiS getroffen?

Die Entscheidung des Staatsministeriums für Unterricht und Kultus, für die neue Aufgabe der flächendeckenden Sprachstandserhebungen in Bayern eineinhalb Jahre vor der Einschulung mit BaSiS ein Verfahren aus Bayern und für Bayern zu konzipieren, erfolgte auf Grundlage einer systematischen Abwägung der an ein entsprechendes Verfahren zu stellenden Anforderungen – insbesondere in Bezug auf Wissenschaftsbasierung, Digitalisierung, Adaptivität, Rechtssicherheit, Marktunabhängigkeit sowie Ressourcenschonung in Entwicklung, Implementierung und Auswertung – und der Prüfung, ob ein bereits am Markt verfügbares Verfahren diesen Anforderungen in vergleichbarer Weise entsprechen kann.

4.2 Sind regelmäßige unabhängige Evaluationen des Verfahrens vorgesehen und, wenn ja, in welchem Turnus und mit welcher Veröffentlichungspraxis?

Unabhängig von der Pilotierung und der ersten Erhebungswelle durchgeführte Evaluationen sind geplant; die Details hierzu werden im Zuge der Gesamtplanung festgelegt.

Hinweise des Landtagsamts

Zitate werden weder inhaltlich noch formal überprüft. Die korrekte Zitierweise liegt in der Verantwortung der Fragestellerin bzw. des Fragestellers sowie der Staatsregierung.

Zur Vereinfachung der Lesbarkeit können Internetadressen verkürzt dargestellt sein. Die vollständige Internetadresse ist als Hyperlink hinterlegt und in der digitalen Version des Dokuments direkt aufrufbar. Zusätzlich ist diese als Fußnote vollständig dargestellt.

Drucksachen, Plenarprotokolle sowie die Tagesordnungen der Vollversammlung und der Ausschüsse sind im Internet unter www.bayern.landtag.de/parlament/dokumente abrufbar.

Die aktuelle Sitzungsübersicht steht unter www.bayern.landtag.de/aktuelles/sitzungen zur Verfügung.