

*Projekt FiSK*

# **Effekte adaptiver Feedbackbots im Simulierten Klassenraum auf prozedurales Professionswissen**

---

Teilvorhaben IPN: Naturwissenschaften

## **Schlussbericht**

<b>Zuwendungsempfänger:</b>	IPN – Leibniz-Institut für die Pädagogik der Naturwissenschaften und Mathematik
<b>Projektleitung:</b>	Prof. Dr. Ute Harms
<b>Förderkennzeichen:</b>	16DHB4004
<b>Vorhabenbezeichnung:</b>	FiSK - Effekte adaptiver Feedbackbots im Simulierten Klassenraum auf prozedurales Professionswissen Teilvorhaben: Naturwissenschaften (FiSK-IPN)
<b>Laufzeit des Vorhabens:</b>	01.03.2021 bis 29.02.2024 (verlängert bis 31.07.2024)

Kiel, 20.01.2025

---

Prof. Dr. Ute Harms

## Inhalt

1	Kurze Darstellung.....	3
1.1	Anknüpfung an den bisherigen wissenschaftlichen Stand.....	3
1.2	Aufgabenstellung .....	3
1.3	Durchführungsvoraussetzung .....	4
1.4	Planung und Ablauf des Vorhabens .....	4
1.5	Zusammenarbeit mit anderen Stellen .....	5
2	Eingehende Darstellung.....	5
2.1	Durchgeführte Arbeiten und erzielte Ergebnisse.....	5
2.1.1	Studie 1: Entwicklung der neuen Themenbereiche .....	6
2.1.2	Studie 2: Variation des Bot-Verhaltens .....	8
2.1.3	Studie 3: Themenabhängige Effekte .....	10
2.1.4	Studie 4: Nutzung des Chatbots .....	11
2.1.5	Implementation des SKR <sup>Bio</sup> in die Lehre .....	11
2.2	Wichtigste Positionen des zahlenmäßigen Nachweises .....	12
2.3	Notwendigkeit und Angemessenheit der geleisteten Arbeit.....	12
2.4	Nutzen sowie Verwertbarkeit der Ergebnisse und Erfahrungen .....	13
2.5	Erfolgte und geplante Veröffentlichungen .....	14
2.5.1	Präsentation der Ergebnisse.....	14
2.5.2	Erfolgte Veröffentlichungen .....	16
2.5.3	Geplante Veröffentlichungen .....	16
	Literaturverzeichnis .....	16

# 1 Kurze Darstellung

## 1.1 Anknüpfung an bisherigen wissenschaftlichen Stand

Diagnostische Fähigkeiten sind ein zentraler Bestandteil der professionellen Kompetenz von Lehrkräften. Sie beziehen sich auf die Fähigkeit, das Verständnis der Lernenden sowie Lernhindernisse zu beurteilen und diese Erkenntnisse zur Unterstützung des Unterrichts zu nutzen (Leuders et al., 2022). Diese Fähigkeiten umfassen deklaratives Wissen („Wissen, dass“; Bromme, 2001) zu Unterrichts- und Lernherausforderungen sowie prozedurales Wissen („Wissen, wie“; Bromme, 2001), das die passende Reaktion auf diese Herausforderungen beschreibt (Förtsch et al., 2018; Leuders et al., 2022). Oft bleibt deklaratives Wissen jedoch träge, was besonders zu Beginn der Lehrtätigkeit zu einem „Praxisschock“ führen kann (Klusmann et al., 2012; Dicke et al., 2016). Um den Übergang von deklarativem zu prozeduralem Wissen zu erleichtern, sollte die universitäre Lehrkräftebildung sowohl fachliche als auch handlungsorientierte Fähigkeiten frühzeitig fördern.

Simulationsbasierte Lernumgebungen bieten eine vielversprechende Möglichkeit, praktische Elemente in die Ausbildung zu integrieren („Approximations of Practice“; Grossman et al., 2009; Howell & Mikeska, 2021). Besonders lernförderlich sind Umgebungen, die durch Instruktionsmerkmale wie Feedback oder geführte Aktivitäten angereichert sind (Hillmayr et al., 2020). Lernumgebungen, die Bots und Künstliche Intelligenz (KI) einsetzen, gewinnen zunehmend an Bedeutung in der Bildungsforschung (Fleckenstein et al., 2024; Hwang & Chang, 2023).

Ein Bot, auch „Chatbot“ genannt, ist ein Software-Werkzeug, das mit Nutzern auf natürliche Weise interagiert (Molnár & Szüts, 2018). In Lernumgebungen werden Bots genutzt, um Lernende zu unterstützen, etwa bei administrativen Aufgaben oder beim Fremdsprachenerwerb (Barret et al., 2019; Yang & Evans, 2019). Moderne Bots verwenden KI und Natural Language Processing (NLP), um komplexe Interaktionen und kontextbezogene Dialoge zu ermöglichen (Hussain et al., 2019). Damit Bots Lehramtsstudierende optimal unterstützen können, müssen sie aber auch über spezifisches Domänenwissen verfügen. Erste Ansätze zur Strukturierung von Fachwissen, wie die Unterteilung in Lerninhalte und typische Fehler, wurden von Dutta (2017) beschrieben und bieten eine wertvolle Grundlage für gezieltes Feedback.

Für dieses Projekt wird die Klassenraumsimulation SKR<sup>Bio</sup> (Fischer et al., 2021; 2022) verwendet, die im vorherigen Projekt ProSim (gefördert durch das BMBF; FKZ 16DHL1015) für den Biologiebereich entwickelt wurde. Der SKR<sup>Bio</sup> ermöglicht die Untersuchung der Diagnosefähigkeiten von Biologie-Lehramtsstudierenden. Den Probanden werden SchülerInnen mit individuellen Fähigkeitsprofilen zugewiesen. Die Studierenden stellen im Rahmen einer virtuellen Schulstunde verschiedene Fragen (beispielsweise zum Thema Evolution) und beurteilen die gegebenen Antworten der simulierten SchülerInnen. Ein Feedback-Modul gibt am Ende eine summative Rückmeldung zur Genauigkeit der Beurteilungen.

## 1.2 Aufgabenstellung

Das Hauptziel des Verbundprojekts FiSK bestand darin, das prozedurale Professionswissen von Lehramtsstudierenden – insbesondere ihre Diagnosefähigkeit – zu fördern. Ein zentraler Ansatz war die Erweiterung der Klassenraumsimulation SKR<sup>Bio</sup> durch einen anfrageorientierten Chatbot, der automatisiertes Feedback zu fachdidaktischem und methodischem Wissen gibt, ohne dass ein direkter menschlicher Eingriff erforderlich ist. Diese Erweiterung sollte eine praxisnahe und individuelle Unterstützung bieten, um Lehramtsstudierende gezielt auf die komplexen Anforderungen ihres späteren Berufs vorzubereiten.

Ziel des Teilvorhabens IPN war es, die Lernumgebung SKR<sup>Bio</sup> inhaltlich und methodisch zu erweitern. Geplant war, die Themen Experimentieren und Argumentieren zu integrieren, um die Lernumgebung über das Thema Evolution und das Fach Biologie hinaus einsetzbar zu machen. Damit sollten sowohl die Funktionalität als auch die Anwendbarkeit des SKR<sup>Bio</sup> erweitert werden, um eine breitere Zielgruppe und vielfältigere Lernkontexte anzusprechen.

Die Wirksamkeit dieser Erweiterungen sollte in experimentellen Studien untersucht werden, um fundierte Erkenntnisse über die Integration und den Einfluss der neuen Ansätze zu gewinnen. Darüber hinaus war geplant, Merkmale zu identifizieren, die die Diagnosefähigkeit der Lehramtsstudierenden fördern oder hemmen. Diese Erkenntnisse sollten zur wissenschaftlichen Weiterentwicklung im Bereich der Lehrkräftebildung beitragen und die Grundlage für zukünftige Optimierungen der Lernumgebung schaffen.

### 1.3 Durchführungsvoraussetzung

Das Projekt FiSK gliederte sich in drei Teilvorhaben (Teilvorhaben A: Naturwissenschaften, IPN; Teilvorhaben B: Adaptive Feedbackbots durch Machine Learning Methoden, HFU; Teilvorhaben C: AI Coaching, AI Coaching GmbH). Zur Koordination des gesamten Projekts fanden regelmäßige digitale Arbeitstreffen aller Beteiligten statt.

Die interdisziplinäre Zusammenarbeit mit den Teilvorhaben B und C konzentrierte sich insbesondere auf die Erweiterung des SKR<sup>Bio</sup> durch adaptive Feedbackbots sowie die Analyse der erhobenen Daten. Die technische Umsetzung der Erweiterung des SKR<sup>Bio</sup> stellte sich jedoch als äußerst zeit- und ressourcenaufwändig heraus. Einige Anforderungen (beispielsweise direktes Feedback zu einzelnen Diagnosen im SKR<sup>Bio</sup>) konnten aufgrund fehlender technischer Voraussetzungen des SKR<sup>Bio</sup> nicht umgesetzt werden. Zudem war die Zusammenarbeit mit CAP3, dem Dienstleister für die Simulation, deutlich zeitintensiver als voraussehbar, insbesondere die notwendige Übertragung der Simulation auf die HFU-Server nahm erheblich mehr Zeit in Anspruch als ursprünglich geplant.

### 1.4 Planung und Ablauf des Vorhabens

Die Projektlaufzeit war von März 2021 bis Februar 2024 geplant. In diesem Zeitraum sollte der SKR<sup>Bio</sup> um die Themen Experimentieren und Argumentieren erweitert und mit einem adaptiven Feedbacksystem ergänzt werden, das durch Unterstützung der HFU sowie AI-Coaching entwickelt werden sollte. Dieses System sollte während der Bearbeitung der Simulation auf individuelle Antworten der Studierenden zugreifen und ihnen zeitnah Feedback zu ihrer Leistung geben, um das Lernen gezielt zu unterstützen.

Technische Verzögerungen, insbesondere bei der Übertragung der Klassenraumsimulation auf die HFU-Server – eine Voraussetzung für die Entwicklung des adaptiven Systems – sowie die fehlende Zugriffsmöglichkeit auf individuelle Diagnoseantworten, führten zu Verzögerungen im Projektablauf. Zudem erschwerte der unerwartete, nicht vorhersehbare Ausfall des ursprünglich für das Projekt vorgesehenen Feedbackbot-Systems, auf den das Projektteam keinen Einfluss hatte, den Fortschritt. Aus diesen Gründen wurde eine kostenneutrale Verlängerung des Projekts um fünf Monate (bis Ende Juli 2024) notwendig.

Im Verlauf des Projekts wurden insgesamt sechs Erhebungen durchgeführt (vgl. Kapitel [2.1 Durchgeführte Arbeiten und erzielte Ergebnisse](#)), die Biologielehramtsstudierende an sechs Universitätsstandorten in Deutschland einbezogen. Diese Erhebungen deckten alle drei Themenbereiche (Evolution, Experimentieren, Argumentieren) des SKR<sup>Bio</sup> ab und wurden unter

Einsatz unterschiedlicher entwickelter Feedbackbots durchgeführt. Die erhobenen Daten wurden im Rahmen des Projekts aufbereitet und ausgewertet. Die Ergebnisse wurden zur Veröffentlichung vorbereitet, publiziert bzw. bei internationalen Fachzeitschriften mit Review-System eingereicht (vgl. Kapitel [2.5 Erfolgte und geplante Veröffentlichungen](#)).

## 1.5 Zusammenarbeit mit anderen Stellen

Abgesehen von der projektinternen Zusammenarbeit (vgl. Kapitel [1.3 Durchführungsvoraussetzung](#)), gab es im ersten Projektjahr eine enge Kooperation zwischen HFU und dem Software-Dienstleister CAP3. Diese Zusammenarbeit war auf die technische Übertragung des SKR<sup>Bio</sup> auf die HFU-Server sowie die Fehlerbeseitigung in der Simulation ausgerichtet. Zur Koordination fanden mehrere digitale Arbeitstreffen statt.

## 2 Eingehende Darstellung

### 2.1 Durchgeführte Arbeiten und erzielte Ergebnisse

In diesem Kapitel werden die durchgeführten Arbeiten und die Ergebnisse der Erhebungen thematisch geordnet präsentiert. Die Erhebungen werden dabei als inhaltlich sinnvoll verknüpfte Studien dargestellt. Das bedeutet, dass zwei Erhebungen zu einer gemeinsamen Studie zusammengefasst oder Daten einzelner Erhebungen mehrfach in verschiedenen Studien berücksichtigt wurden. Ein Teil der Erhebungen wurde zudem verwendet, um die neu entwickelten Themenbereiche des SKR<sup>Bio</sup> (Experimentieren und Argumentieren) zu validieren.

Insgesamt werden vier Studien vorgestellt (**Studie 1:** Entwicklung der neuen Themen, **Studie 2:** Variation des Bot-Verhaltens, **Studie 3:** Themenabhängige Effekte, **Studie 4:** Nutzung des Chatbots) sowie Erkenntnisse zur Implementation des SKR<sup>Bio</sup> in die Lehre präsentiert. Tabelle 1 bietet eine Übersicht der Erhebungen und die Verwendung der Daten in den vier Studien.

Tabelle 1. Übersicht der Erhebungen und der Nutzung in den hier vorgestellten Studien

Erhebung	SKR <sup>Bio</sup> Thema	Bot-Nutzung	Probanden	Teil von Studie			
				1	2	3	4
EvoAug2022	Evolution	G1: kein Bot G2: Wissensbot	133	x	x		
ArgDec2022	Argumentieren	Wissensbot	129	x			
ExpFeb2023	Experimentieren (inkl. Einleitungsvideo)	Wissensbot	184	x			x
EvoMar2023	Evolution (inkl. Einleitungsvideo)	Wissensbot + Prozessbot	46		x		
EvoJan2024 / ArgJan2024	Evolution / Argumentieren (inkl. Einleitungsvideo)	Wissensbot	80 / 81			x	x
EvoApr2023	Evolution (inkl. Einleitungsvideo)	Wissensbot + Prozessbot	28			Implementation in die Lehre	

**Bemerkung.** Die Bezeichnung der Erhebungen entsprechen der Bezeichnung der Datensätzen, die im August 2026 bei GESIS veröffentlicht werden und dann für Forschung und Lehre zur Verfügung stehen.

Ein Teil der erzielten Ergebnisse wurde bereits veröffentlicht (vgl. Kapitel [2.5.2 Erfolgte Veröffentlichungen](#)), weitere Manuskripte sind bei Fachzeitschriften mit Review-System eingereicht bzw. in Vorbereitung zur Einreichung (vgl. Kapitel [2.5.3 Geplante Veröffentlichungen](#)). Um sicherzustellen, dass der Peer-Review-Prozess der geplanten Veröffentlichungen nicht beeinträchtigt wird, erfolgt in diesem Kapitel lediglich eine zusammenfassende Darstellung der Ergebnisse.

### 2.1.1 Studie 1: Entwicklung der neuen Themenbereiche

Zu Beginn des Projekts lag der SKR<sup>Bio</sup> nur zum Thema Evolution vor. Er wurde im vorausgehenden Projekt "Prozedurales Professionswissen im Simulierten Klassenraum entwickeln - ProSim" (BMBF, FKZ 16DHL 1014) am IPN entwickelt. Im Rahmen des FISK-Projekts wurde der SKR<sup>Bio</sup> zunächst um die Themenbereiche Experimentieren und Argumentieren erweitert. Das ursprüngliche SKR<sup>Bio</sup>-Design sah eine Differenzierung der Aussagen der virtuellen Schülerinnen und Schüler ausschließlich in korrekte und inkorrekte Kategorien vor. Um die Authentizität der Aussagen in allen drei Themenbereichen zu erhöhen, wurden die Aussagen so erweitert und umstrukturiert, dass sie nun in verschiedenen Niveaustufen vorliegen. Dazu wurden für alle drei Themenbereiche theoriegeleitete Niveaustufenmodelle entwickelt, die sich an validierten und publizierten Kompetenzniveaustufenmodellen orientieren. Während der simulierten Unterrichtsstunde identifizierten die Teilnehmenden zunächst inhaltspezifische Wissens Elemente in den Antworten der virtuellen Schüler:innen (vgl. Tabelle 2) und ordneten die Antwort im Anschluss einer Niveaustufe des Kompetenzmodells zu (vgl. Tabelle 3).

Tabelle 2. Übersicht der inhaltspezifischen Wissens Elemente der drei Themen im SKR<sup>Bio</sup>

Evolution	Experimentieren	Argumentieren
<i>Evolutionsbiologische Prinzipien:</i> <ul style="list-style-type: none"> <li>• Variabilität</li> <li>• Vererbung</li> <li>• Selektion</li> </ul> <i>Schwellenkonzepte:</i> <ul style="list-style-type: none"> <li>• Zufall</li> <li>• Wahrscheinlichkeit</li> <li>• zeitliche Maße</li> </ul> <i>Fehlvorstellungen (FV):</i> <ul style="list-style-type: none"> <li>• Anthropomorphe FV</li> <li>• Teleologische FV</li> <li>• FV von Nutzung &amp; Nichtnutzung</li> </ul>	<i>Aspekte des Experimentierens:</i> <ul style="list-style-type: none"> <li>• Nennung der Messgröße (AV)</li> <li>• Messung der Messgröße (AV)</li> <li>• Nennung der Einflussgröße (UV)</li> <li>• Variation der Einflussgröße (UV)</li> <li>• Nennung von Störgrößen</li> <li>• Konstanthaltung von Störgrößen</li> <li>• Identifizierung zusätzlicher Eigenschaften eines Experiments</li> </ul>	<i>Strukturelemente:</i> <ul style="list-style-type: none"> <li>• Fakt</li> <li>• Erläuterung / Schlussregel</li> <li>• Stützung</li> <li>• Ausnahmebedingung</li> </ul> <i>Argumentationsmuster:</i> <ul style="list-style-type: none"> <li>• Einordnungsmuster</li> <li>• Vergleichsmuster</li> <li>• Gegensatzmuster</li> <li>• Kausalmuster</li> <li>• Induktives Beispielmuster</li> <li>• Autoritätsmuster</li> </ul>

Tabelle 3. Beschreibung der theoriebasierten Niveaustufen der drei Themen im SKR<sup>Bio</sup>

Niveau- stufe	Evolution <i>Antwort im SKR<sup>Bio</sup> besteht aus...</i>	Experimentieren	Argumentieren
<b>Niveau 1</b>	... einer Fehlvorstellung.	... der Beschreibung einer Variable (Messgröße, Einflussgröße oder Störgrößen).	... einer Behauptung, einem oder mehreren Begründungen

			bestehend aus Fakten oder Erläuterungen.
<b>Niveau 2</b>	... einer Fehlvorstellung und maximal einem korrekten evolutionsbiologischem Prinzip.	... der Beschreibung zweier Variablen.	... einer Behauptung, einem oder mehreren Begründungen und einer Stützung.
<b>Niveau 3</b>	... einer Fehlvorstellung, mindestens zwei evolutionsbiologische Prinzipien und ggf. einem Schwellenkonzept.	... der Beschreibung einer abhängigen und einer unabhängigen Variable, die miteinander in Beziehung gesetzt werden.	... einer Behauptung, einem oder mehreren Begründungen und einer Ausnahmebedingung.
<b>Niveau 4</b>	... allen evolutionsbiologischen Prinzipien und mindestens einem Schwellenkonzept.	... der Beschreibung einer abhängigen und einer unabhängigen Variable, die unter Berücksichtigung einer Kontrollvariablen (Störgrößen) in Beziehung gesetzt werden.	... einer Behauptung, einem oder mehreren Begründungen, einer Stützung und einer Ausnahmebedingung.
<b>Niveau 5</b>	<i>[nicht vorhanden]</i>	... den in Niveaustufe 4 angegebenen Aspekten sowie der Identifizierung zusätzlicher Eigenschaften eines Experiments (z.B. Fehlerdiskussion)	<i>[nicht vorhanden]</i>

Auf dieser Grundlage entstanden erstmalig für die Themenbereiche Experimentieren und Argumentieren Frage-Antwort-Pfade, die auf eigenen sowie externen Datensätzen basieren. Für den bereits entwickelten Themenbereich Evolution wurden die Aussagen angepasst oder durch neue Aussagen aus bestehenden Datensätzen ergänzt, sodass nun alle Niveaustufen abgebildet werden können. Die konzeptionellen Grundlagen und Schülerantworten für die Themenbereiche Experimentieren und Argumentieren wurden vor dem ersten Einsatz der jeweiligen Klassenraumsimulation von Expert:innen der Fachdidaktik überprüft.

Die neuen Themenbereiche wurden anschließend erstmalig mit Biologie-Lehramtsstudierenden getestet (Erhebungen ArgDec2022 und ExpFeb2023). Der Themenbereich Evolution, der lediglich angepasst, aber nicht neu entwickelt wurde, wurde ebenfalls erneut getestet (Erhebung EvoAug2022), um die Wirksamkeit der Anpassungen zu überprüfen.

Im Themenbereich **Evolution** hatten die Studierenden häufig Schwierigkeiten bei der Identifikation von Schwellenkonzepten und Fehlvorstellungen. Etwa ein Drittel der Studierenden konnte alle Wissensselemente (d.h. vorhandene Prinzipien, Schwellenkonzepte oder Fehlvorstellungen in einer Aussage, vgl. Tabelle 2) korrekt erkennen, während rund zwei Drittel der Studierenden die den Aussagen zugrundeliegende Niveaustufe korrekt einschätzen konnten. Besonders bei gemischten Erklärungen (z. B. Niveau 2 und 3, vgl. Tabelle 3) gestaltete sich die korrekte Diagnose der Niveaustufen schwieriger als bei eindeutig wissenschaftlichen oder naiven Erklärungen (Niveau 4 und 1). Eine Erklärung für die fehlerhaften Diagnosen kann unzureichendes

Fachwissen der Studierenden sein, insbesondere in Bezug auf häufig vorkommende Fehlvorstellungen im Bereich der Evolution.

Zum Themenbereich **Experimentieren** wiederum konnten die Studierenden nur in etwa einem Drittel der Antworten alle vorhandenen Wissensselemente (z. B. Variation der Einflussgröße, vgl. Tabelle 2) korrekt identifizieren, während etwa die Hälfte der Aussagen der korrekten zugrundeliegenden Niveaustufe zugeordnet werden konnte. Allerdings zeigten viele Studierende Probleme bei der Identifikation der abhängigen (Messgröße) bzw. unabhängigen (Einflussgröße) Variable, was die Zuordnung der korrekten Niveaustufe oftmals erschwerte. Korrelationsanalysen zeigen Zusammenhänge der Diagnosefähigkeit mit dem Selbstkonzept im Bereich Experimentieren, den Lernstrategien und dem fachlichen Vorwissen der Studierenden.

Im Themenbereich **Argumentieren** konnten die Studierenden nur in etwa zwei Fünftel der Argumente alle Wissensselemente korrekt identifizieren, wobei die Diagnose von Aussagen auf Niveaustufe 3 besonders schwierig war. In mehr als der Hälfte der Argumente konnten die Studierenden mindestens ein hinterlegtes Argumentationsmuster (s. Tab. 2) korrekt diagnostizieren. Dabei erkannten sie Einordnungs-, Kausal- oder Autoritätsmuster weniger gut als Vergleichs-, Gegensatz- oder induktive Beispielmuster.

**Gewonnene Erkenntnisse:** Durch die thematische Erweiterung kann der SKR<sup>Bio</sup> nun gezielt als Instrument zur Messung und Untersuchung der Diagnosefähigkeit von Biologie-Lehramtsstudierenden in drei zentralen Themenbereichen (Evolution, Experimentieren und Argumentieren) eingesetzt werden. Die Ergebnisse der einzelnen Erhebungen deuten darauf hin, dass Biologie-Lehramtsstudierende in den drei Themenbereichen unterschiedliche Fähigkeitsprofile aufweisen, was wertvolle Einblicke in ihre Stärken und Entwicklungsbedarfe bietet und damit Ansatzpunkte für die Weiterentwicklung der Lehrkräftebildung zum Diagnostizieren erkennbar macht.

### 2.1.2 Studie 2: Variation des Bot-Verhaltens

Parallel zu Studie 1 wurden Feedback-Konzepte abstrahiert und systematisch klassifiziert, um sie in ein interaktives Feedback-System innerhalb des SKR<sup>Bio</sup> zu integrieren. Nach der unerwarteten Einstellung des Chatbot-Sub-Systems „Botmock“ durch das berechnete Unternehmen nach einer Übernahme (Dezember 2021) wurde nach einer Evaluierung potenzieller Ersatzsysteme auf „Dialogflow“ von Google umgestellt. Der Verbundpartner AI Coaching GmbH (zuständig für das Hosten der Systeme zur visuellen Darstellung des Chatverhaltens) informierte den Projektträger über die Abschaltung des alten Systems (Botmock) und der daraus resultierenden Notwendigkeit eines Alternativsystems (Dialogflow).

Ursprünglich war geplant, direktes Feedback zu einzelnen Diagnosen in den SKR<sup>Bio</sup> zu integrieren, um die Diagnosefähigkeit der Teilnehmenden zeitnah und kontrolliert zu beeinflussen. Nach intensiver Beurteilung der technischen Systemvoraussetzungen und der Prüfung alternativer Ansätze musste das Projektteam jedoch von dieser Idee absehen, da die Integration nicht realisierbar war oder nur mit erheblichem, im Rahmen des Projekts nicht umsetzbarem Umstrukturierungsaufwand möglich gewesen wäre. Stattdessen wurde entschieden, zwei Arten von Feedback und Botverhalten zu implementieren:

1. **Wissensbot:** Ein reaktiver Bot, der anfragenorientiert Wissensfragen beantwortet. Der Bot nutzt einen regelbasierten Ansatz und liefert statische Antworten zu (1) diagnostischen Konzepten, (2) einem Glossar von Begriffen und (3) allgemeinen Informationen zum Datenerhebungsprozess.

2. **Prozessbezogener Bot (kurz Prozessbot):** Ein proaktiver Bot, der Feedback zu diagnostischen Entscheidungen gibt (z. B. mehrmalige Auswahl desselben virtuellen Schülers), um auf unzureichende Strategien zur Erreichung des Gesamtziels (summative Bewertung aller virtuellen SchülerInnen) aufmerksam zu machen. Auch dieser Bot nutzt den regelbasierten Ansatz und liefert spezifische Rückmeldungen zu individuellen Diagnosestrategien im SKR<sup>Bio</sup>.

Die zugrundeliegenden technischen Modelle sowie das Darstellungsdesign des Chatbot mit dem SKR<sup>Bio</sup> wurden von den Projektpartnern in Furtwangen erarbeitet und bereitgestellt (vgl. Abbildung 1). Der Wissensbot wurde für alle drei Themenbereiche (Evolution, Experimentieren und Argumentieren) entwickelt, während der Prozessbot vorerst nur für den Themenbereich Evolution verfügbar war.

**Virtuelle Klasse**  
Alle virtuellen Schüler:innen werden mit einem Porträtfoto und Namen angezeigt. Schüler:innen, die eine Antwort geben wollen, sind orange markiert und mit einem kleinen Fingersymbol versehen.

**Chatbot**  
Der Chatbot wird über einen geteilten Bildschirm integriert.

Klassenraum Token: umvJD 44:51 Überspringen >

**Fragenpool**  
Zuerst der Themenbereich (z.B. Evolution), dann der biologische Bereich (hier zoologische oder botanische Beispiele). Durch „Anklicken“ des Fragenpools können die einzelnen Fragen aufgerufen werden.

**Einzelne Fragen**  
Zu jedem biologischen Bereich gibt es vier Einzelfragen. Die vollständige Frage, die der Klasse gestellt wurde und auf die die virtuellen Schüler:innen antworten, wird auf der rechten Seite (Vorschau) angezeigt.

**Fragen**  
Teufelszwirn  
Tabakpflanze  
Weißklee **Frage stellen**  
Rotes Straußgras

**Vorschau**  
Beim Weißklee gibt es Formen, welche Cyanidverbindungen synthetisieren können und diese extrazellulär einlagern. Einerseits ist der Weißklee dadurch ungenießbar für Pflanzenfresser, andererseits kann bei Frost Cyanid in das Gewebe gelangen und die Pflanze schädigen. Im Süden konnten beim Weißklee hohe Mengen an Cyanid gemessen werden. Dort gibt es viele Insekten. Im Norden hingegen enthalten die Pflanzen seltener Cyanid und es kommt häufiger zu Frösten. Wie ist die geringere Cyanidproduktion bei den nördlichen Populationen durch evolutionsbiologische Prinzipien zu erklären?

**Hallo, ich bin der EvolutionChat!**  
Ich unterstütze dich im SKR. Wenn du etwas wissen möchtest, kannst du mich gerne zu verschiedenen Themen fragen. Ich versuche dir dann zu antworten.  
Sollte etwas nicht so gut klappen, dann nutze das Wort **ABBRUCH** und wir starten einfach von vorne. Wie du bin auch ich noch am Lernen.

**Bitte gib zuerst dein achtstelliges Kennwort ein:**  
Zur Erinnerung, dass Kennwort besteht aus  
(1) der Zahl des heutigen Tages,  
(2) dem ersten und zweiten Buchstabe des Vornamens Ihrer Mutter,  
(3) der Zahl Ihres Geburtstags (nur der Tag) und  
(4) dem ersten und zweiten Buchstabe des Vornamens Ihres Vaters.

Frage eingeben

Abbildung 1. Aufbau des SKR<sup>Bio</sup>-Chatbot-Systems zum Themenbereich Evolution.

Ziel dieser Studie war es zu untersuchen, welchen Effekt verschiedene Chatbot-Systeme (drei Gruppen: ohne Bot-Zugriff, mit Wissensbot, mit Wissensbot und Prozessbot) auf die Diagnosefähigkeit von Biologie-Lehramtsstudierenden hat. Die Ergebnisse unserer Erhebungen (EvoAug2022 und EvoMar2023) zeigen jedoch, dass der Wissensbot nur begrenzt genutzt wurde, obwohl er potenziell als hilfreiches Lerninstrument dienen könnte. Der Wissensbot wurde häufiger in der Gruppe verwendet, die auch Zugriff auf den Prozessbot hatte, welcher Rückmeldungen zu diagnostischen Entscheidungen gab. Dennoch lagen die Anfragen im Zusammenhang mit diagnostischen Prozessen insgesamt auf einem niedrigen Niveau (durchschnittlich zwei bis drei Anfragen pro Probanden bei einer Durchführungszeit von 45 Minuten). Wenn der Chatbot genutzt wurde, dann vor allem, um Rückfragen zu themenspezifischen Konzepten zu stellen oder Fehlvorstellungen zu klären. Insgesamt konnten keine signifikanten Unterschiede in der Diagnosefähigkeit zwischen den Gruppen festgestellt werden. Diese Ergebnisse verdeutlichen die

Notwendigkeit, in zukünftigen Studien diagnosebezogene Rückmeldungen zu integrieren, um die Diagnosefähigkeit der Lehramtsstudierenden gezielt zu fördern.

**Gewonnene Erkenntnisse:** Während unsere Studie keine signifikanten Unterschiede der Diagnosefähigkeit basierend auf dem Bot-Verhalten aufzeigte, bleibt das Potenzial des SKR<sup>Bio</sup>-Chatbot-Systems zur Förderung der Diagnosefähigkeit von Lehramtsstudierenden dennoch weiterhin bestehen, insbesondere in frühen Ausbildungsphasen. Für Studierende bietet der SKR<sup>Bio</sup> die einzigartige Möglichkeit, erlerntes deklaratives Wissen in einer konkreten, simulierten Handlungssituation anzuwenden und durch den integrierten Chatbot Unterstützung bei fehlendem Diagnosewissen anzufragen.

### 2.1.3 Studie 3: Themenabhängige Effekte

Die vorliegende Studie baut auf den Ergebnissen der ersten Studie auf, in der bereits Hinweise darauf gefunden wurden, dass Teilnehmende unterschiedliche Fähigkeitsprofile in den verschiedenen Themenbereichen Evolution und Argumentieren aufweisen könnten. Da es sich bei der ersten Erhebung um unabhängige Stichproben handelte, war jedoch kein kausaler Vergleich möglich. Um diese Lücke zu schließen, wurde in Studie 3 ein neues Design gewählt, bei dem der Einfluss der Nutzung des SKR<sup>Bio</sup>-Chatbot-Systems in einem bestimmten Themenbereich (z.B. Evolution) auf die Nutzung des Systems im anderen Themenbereich (z.B. Argumentieren) untersucht wurde.

Ursprünglich war geplant, diese Erhebung unter den Bedingungen der Hochschule Furtwangen durchzuführen. Leider musste die Erhebung jedoch aufgrund eines Hacker-Angriffs auf die gesamte IT-Infrastruktur der Hochschule unterbrochen werden, wodurch auch die Server, auf denen alle Systeme des Projekts liefen, betroffen waren. In den darauffolgenden Wochen und Monaten wurde in Zusammenarbeit mit den Projektpartnern eine Übergangslösung erarbeitet, sodass die betroffenen Systeme auf Servern des IPN Kiel betrieben werden konnten. Diese Anpassung ermöglichte es, die Erhebung fortzuführen und die Studie schließlich wie geplant durchzuführen.

Für die Erhebung (EvoJan2024 und ArgJan2024) wurden Biologielehramtsstudierende gebeten, innerhalb von zwei Wochen an einer zweiteiligen Untersuchung teilzunehmen. Am ersten Tag wurden sie zufällig einem der beiden Themenbereiche – Evolution oder Argumentieren – zugewiesen. Am zweiten Tag wechselten sie in den jeweils anderen Themenbereich, um mögliche Reihenfolgeeffekte zu kontrollieren und die Vergleichbarkeit zu gewährleisten. Basierend auf den bisherigen Ergebnissen wurde angenommen, dass Diagnosen im Bereich Evolution den Teilnehmenden möglicherweise leichter fallen würden als im Bereich Argumentieren. Durch die verlängerte Erhebungszeit über zwei Wochen konnte zudem untersucht werden, inwiefern sich die Nutzung des SKR<sup>Bio</sup>-Chatbot-Systems auf die Diagnosefähigkeit auswirkt. Hierfür wurden in einem Vortest, Nachtest und Follow-up jeweils die Diagnosefähigkeiten zum Themenbereich des ersten Tages gemessen.

Die Untersuchung zur Themenspezifität ergaben, dass sich die Diagnosefähigkeiten der Studierenden im Bereich Evolution von denen im Bereich Argumentieren unterscheiden. Besonders auffällig waren die signifikanten Unterschiede bei den formativen Diagnosen: Im Bereich Evolution diagnostizierten die Studierenden häufiger korrekt als im Bereich Argumentieren. Zudem diagnostizierten sie im Bereich Evolution insgesamt mehr Antworten als im Bereich Argumentieren. Bei den summativen Diagnosen konnte jedoch kein Unterschied zwischen den beiden Themenbereichen festgestellt werden.

Die Ergebnisse zum Lerneffekt wiederum zeigten unterschiedliche Ergebnisse für die zwei Themenbereiche: Im Kontext Evolution konnte keine bedeutsame Veränderung der Diagnosefähigkeiten zwischen den verschiedenen Messzeitpunkten festgestellt werden. Die Werte für die Anzahl korrekt diagnostizierter Antworten blieben über alle Zeitpunkte hinweg nahezu konstant. Im Gegensatz dazu zeigte sich im Bereich Argumentieren eine deutliche Verbesserung der Diagnosefähigkeiten zwischen dem Pretest und dem Follow-up. Während die Studierenden im Pretest durchschnittlich weniger korrekt diagnostizierten, stieg die Zahl der richtigen Diagnosen im Follow-up signifikant an.

**Gewonnene Erkenntnisse:** Diagnosen im Bereich Evolution fallen den Studierenden offenbar leichter als im Bereich Argumentieren. Dies könnte auf unterschiedliche Vorkenntnisse oder die Komplexität der beiden Themenbereiche hinweisen. Zudem zeigte sich, dass die Nutzung des SKR<sup>Bio</sup>-Chatbot-Systems insbesondere im Bereich Argumentieren die Diagnosefähigkeit der Studierenden signifikant verbessert hat. Dies ist von Bedeutung für die zukünftige Anwendung und Weiterentwicklung von Diagnosetools wie dem SKR<sup>Bio</sup>. Insgesamt liefern die Ergebnisse wertvolle Hinweise darauf, wie digitale Lernsysteme gezielt die Diagnosefähigkeiten von angehenden Lehrkräften in unterschiedlichen fachlichen Kontexten fördern können.

#### 2.1.4 Studie 4: Nutzung des Chatbots

Für die Studie 4 wurde die Nutzung des Chatbots und die daraus entstandenen Dialoge der zu den drei Themenbereichen Evolution (Erhebung EvoJan2024), Experimentieren (Erhebung ExpFeb2023) und Argumentieren (Erhebung ArgJan2024) gesammelt und ausgewertet.

Im Themenbereich **Evolution** wurden insgesamt 115 Anfragen gestellt, wobei nur 36 % der Studierenden den Chatbot nutzten. Die meisten Anfragen konzentrierten sich auf diagnosebezogene Informationen, bei denen die Teilnehmenden nach Informationen zu (spezifischen) Schwellenkonzepten oder Fehlvorstellungen suchten. Anfragen zu den Niveaustufen fielen eher geringer aus. Insgesamt zeigt sich, dass die meisten Anfragen eine Beschreibung erbat, gefolgt von Beispielen und nur einer geringen Nachfrage nach spezifischen Schlüsselworten für die entsprechenden Konzepte oder Fehlvorstellungen.

Im Themenbereich **Experimentieren** nutzen 44 % der Studierende den Chatbot und stellten dabei insgesamt 133 Anfragen. Die Mehrheit der Anfragen konzentrierten sich auf die Konzepte der abhängigen oder unabhängigen Variable sowie die Niveaustufen. Dabei sind die Anfragen relativ ausgeglichen zwischen der Anfrage nach einer Beschreibung oder der Bitte um Beispiele.

Im Themenbereich **Argumentieren** konnte mit 63 % der Studierenden und insgesamt 358 Anfragen die höchste Nutzung des Chatbots dokumentiert werden. Dabei konzentrierten sich der Großteil der Anfragen auf die Argumentationsmuster. Die meisten Studierenden wollten bei ihren Anfragen entweder Beispiele oder eine Beschreibung.

**Gewonnene Erkenntnisse:** Die Auswertung der Chatbot-Nutzung zeigt eine unterschiedliche Häufigkeit der Anfragen in den verschiedenen Themenbereichen. Der größte Anteil der Anfragen entfiel auf den Bereich Argumentieren, gefolgt von Experimentieren und Evolution. Diese Ergebnisse geben Hinweise darauf, welche Themen besonders häufig nachgefragt wurden, was möglicherweise auf einen geringeren Wissensstand in diesen Bereichen hinweist oder darauf, dass diese Themen in der universitären Ausbildung möglicherweise weniger intensiv behandelt wurden.

#### 2.1.5 Implementation des SKR<sup>Bio</sup> in die Lehre

Im Rahmen einer einsemestrigen Lehrveranstaltung für Masterstudierende des Lehramts Biologie an der Universität Kiel wurde das SKR<sup>Bio</sup>-Chatbot-System zum Thema Evolution (welches in diesem Fall

aus Wissensbot und Prozessbot bestand) genutzt, um die Diagnosefähigkeit der Studierenden zu fördern. Diese Erhebung fand an zwei Seminartagen statt, an denen die Studierenden die Möglichkeit hatten, mit dem System zu arbeiten und nach der Durchführung jeweils ein individuelles Feedback zu ihrer erbrachten Leistung zu erhalten.

Die Ergebnisse zeigen, dass die Studierenden den Großteil der Antworten der korrekten Niveaustufe zuordnen konnten. In der Hälfte der Fälle wurden auch alle zugrundeliegenden Konzepte oder Fehlvorstellungen korrekt diagnostiziert. Besonders herausfordernd blieben jedoch die Diagnosen der Niveau-3-Erklärungen. Bei der summativen Bewertung konnte in zwei Drittel der Fälle die richtige Niveaustufe den virtuellen Schüler:innen zugeordnet werden, wobei sich keine signifikanten Unterschiede in den zugrundeliegenden Niveaustufen zeigten. Zusätzlich zeigte sich, dass die Diagnosen der Einzelaussagen sowie der Gesamtleistung der Schülerinnen am zweiten Tag leicht angestiegen sind. Es muss jedoch angemerkt werden, dass die Ergebnisse aufgrund der geringen Teilnehmerzahl am zweiten Tag (nur sechs Studierende, während am ersten Tag 28 Studierende teilnahmen) nicht repräsentativ sind.

Zusätzlich zu dieser offiziellen Untersuchung wurde das SKR<sup>Bio</sup>-Chatbot-System inklusive des individuellen Feedbacks (insbesondere zu den Themenbereichen Experimentieren und Argumentieren) auch an einzelnen Tagen im Rahmen von Lehrveranstaltungen für Bachelorstudierende verwendet. Dabei gab es von den Studierenden generell positive Rückmeldungen (mündlich an die Seminarleitung).

**Gewonnene Erkenntnisse:** Das SKR<sup>Bio</sup>-Chatbot-System lässt sich effektiv in die Lehre integrieren und kann den Studierenden eine wertvolle Möglichkeit bieten, ihr Wissen in einem geschützten Raum zu testen und zu vertiefen. Besonders eine wiederholte Nutzung des Systems könnte dazu beitragen, Routinen für diagnostische Prozesse zu entwickeln, die später in realen Unterrichtskontexten angewendet werden können. Darüber hinaus bietet die Integration des Chatbots eine sichere Möglichkeit, fachliche oder fachdidaktische Fragen zu stellen, ohne diese Unklarheiten und Unsicherheiten der Seminarleitung gegenüber äußern zu müssen.

## 2.2 Wichtigste Positionen des zahlenmäßigen Nachweises

Die Positionen des zahlenmäßigen Nachweises entsprechen den im Antrag formulierten Planungen und werden im Schlussverwendungsnachweis dargelegt. In erster Linie wurden die Mittel für Personalausgaben verwendet. Die übrigen Ausgaben ergeben sich durch sachliche Verwaltungsmittel in Form von Reisemitteln für nationale und internationale Tagungen sowie der Entschädigung von Versuchspersonen.

## 2.3 Notwendigkeit und Angemessenheit der geleisteten Arbeit

Die durchgeführten Arbeiten sowie die dafür angewandten Ressourcen waren notwendig und angemessen, da sie (1) der im Projektantrag dargelegten Planung entsprachen und (2) alle im Arbeitsplan formulierten Aufgaben - insofern die technischen Voraussetzungen gegeben waren - bearbeitet wurden. Es gibt zwar noch ausstehende Positionen, wie beispielsweise die Veröffentlichung einzelner Studienergebnisse, für diese wurden aber alle technischen und inhaltlichen Voraussetzungen geschaffen. Im Teilvorhaben IPN konnten die Ergebnisse des Projekts zudem in drei Masterarbeiten (nicht publiziert) erfolgreich verwertet werden.

## 2.4 Nutzen sowie Verwertbarkeit der Ergebnisse und Erfahrungen

Der SKR<sup>Bio</sup> hat sich durch die thematische Erweiterung um die Bereiche „Experimentieren“ und „Argumentieren“ als wertvolles Instrument zur Untersuchung der Diagnosefähigkeit von Biologie-Lehramtsstudierenden etabliert. Dabei konnten erstmals differenzierte Erkenntnisse über die Diagnosefähigkeit in den Kompetenzbereichen Sachkompetenz, Erkenntnisgewinnung und Kommunikation/Bewertung gewonnen werden. Besonders hervorzuheben ist, dass durch die Implementierung anfragenorientierter Wissensbots spezifische Diagnoseschwierigkeiten, wie etwa die Unterscheidung von abhängiger und unabhängiger Variable im Thema „Experimentieren“, identifiziert werden konnten. Diese Einsichten tragen maßgeblich dazu bei, das Verständnis für die Stärken und Schwächen der Studierenden in diesen Schlüsselbereichen zu vertiefen.

Nach Abschluss des Projekts soll der SKR<sup>Bio</sup> mit den drei fachdidaktisch abgesicherten Themen (Evolution, Experimentieren, Argumentieren) sowie umfangreicher Funktionalität und breiter Anwendbarkeit auch über das Fach Biologie hinaus interessierten Personen der universitären Lehrkräftebildung zur Verfügung stehen. Die Erweiterung um die Chatbots ist dabei optional, wird jedoch von den Projektbeteiligten empfohlen, wenn keine inhaltliche Einführung in die Themen oder zugrunde liegenden Kompetenzstufen erfolgt ist. Der SKR<sup>Bio</sup> wird bereits in Lehrveranstaltungen für Biologie-Lehramtsstudierende am Standort Kiel eingesetzt und auch Hochschulstandorte wie München und Linz haben Interesse an seiner Nutzung für Lehrzwecke bekundet. Dies verdeutlicht das Potenzial und die Vielseitigkeit des SKR<sup>Bio</sup> als innovative Lernumgebung.

Im Rahmen des Verbundprojekts „Digitalisierungsbezogene und digital gestützte Professionalisierung von MIN-Lehrkräften“ (DigiProMIN) im Kompetenzverbund lernen:digital (gefördert durch die Europäische Union und das BMBF, FKZ 01JA23M05B) wird der SKR<sup>Bio</sup> im Teilprojekt 2c "Simulationen zur Förderung von Diagnosekompetenz in der Fort- und Weiterbildung von Biologielehrkräften" (Leitung Prof. Harms) auch als digitale Lernumgebung in Veranstaltungen für Biologie-Lehrkräfte der zweiten und dritten Bildungsphase in Schleswig-Holstein, Hamburg, Brandenburg und Berlin genutzt, um die Diagnosefähigkeit der (angehenden) Lehrkräfte in den zuvor genannten Kompetenzbereichen weiter zu stärken. Darüber hinaus wurde innerhalb des Teilprojekts in Zusammenarbeit mit der Humboldt-Universität zu Berlin eine thematische Erweiterung entwickelt, die das Thema „Modellieren“ als Teil der Erkenntnisgewinnung in den SKR<sup>Bio</sup> integriert. Diese Erweiterung zeigt das stetige Bestreben, die digitale Lernumgebung weiter zu optimieren und den Bedürfnissen der Lehrkräfte und Studierenden gerecht zu werden.

Die oben beschriebenen Studienergebnisse wurden in der Projektlaufzeit auf verschiedenen nationalen und internationalen Tagungen präsentiert. Dies führte zu regem Interesse am SKR<sup>Bio</sup> an verschiedenen Standorten und in Folge zu einer weiteren Nutzung des SKR<sup>Bio</sup> über die oben genannten Standorte hinaus. So wird die Klassenraumsimulation nun u.a. auch an der Ludwig-Maximilians-Universität in München sowie an der PH Oberösterreich in Linz in der Lehrkräftebildung eingesetzt.

Ein weiterer wichtiger Aspekt der Verwertbarkeit der Projektergebnisse liegt in der Veröffentlichung der gesammelten Daten und Erkenntnisse, die bereits in Fachpublikationen (wie dem Buch „Lehrkräftebildung von morgen“, vgl. [2.5.2 Erfolgte Veröffentlichungen](#)) präsentiert wurden. Die digitalen Forschungsdaten wurden gemäß guter wissenschaftlicher Praxis und Datenschutzvorgaben einerseits im IPN-internen Forschungsdatenarchiv hinterlegt und andererseits GESIS zur Archivierung übergeben. Bei GESIS werden nach Ablauf des Embargos im August 2026 die Daten mit einer DOI veröffentlicht und für Forschung und Lehre frei verfügbar sein ([www.gesis.org](http://www.gesis.org)).

Ein Nachweis der Forschungsdaten wird bereits vorab auf der Website des Verbunds Forschungsdaten Bildung ([www.forschungsdaten-bildung.de](http://www.forschungsdaten-bildung.de)) einsehbar sein. Diese Transparenz ermöglicht es, die Erkenntnisse der Bildungsforschung zur Verfügung zu stellen und die Diskussion über die Weiterentwicklung der Lehrerbildung anzuregen.

## 2.5 Erfolgte und geplante Veröffentlichungen

### 2.5.1 Präsentation der Ergebnisse

Im Folgenden sind alle Präsentationen der Ergebnisse auf nationalen und internationalen Fachtagungen sowie bei anderen relevanten Veranstaltungen aufgeführt, auf denen das Projekt vorgestellt wurde.

- März 2021 **Vortrag beim digitalen Symposium „COSIMA meets ProSim et al.“**  
Harms, U., & Fischer, J. (2021, 22. März). FiSK - Effekte adaptiver Feedbackbots im Simulierten Klassenraum auf prozedurales Professionswissen [Vortrag]. COSIMA meets ProSim et al. – Simulationsbasierte Lernumgebungen in der Hochschule (digitales Symposium), Kiel.
- Juni 2021 **Vortrag beim digitalen IPN-Forschungskolloquium**  
Fischer, J., & Harms, U. (2021, 17. Juni). FiSK - Effekte adaptiver Feedbackbots im Simulierten Klassenraum auf prozedurales Professionswissen [Vortrag]. IPN-Forschungskolloquium (digital), Kiel.
- Juni 2022 **Projektvorstellung bei der virtuellen BMBF-Veranstaltung „Vernetzungsveranstaltung KI in der Hochschulbildung am 22./23. Juni 2022“**
- Aug. 2022 **Vortrag bei der Fachtagung “13th Conference of ERIDOB”**  
Fiedler, D., Schönle, D., Reich, C. & Harms, U. (2022, 31. August). Diagnosing evolutionary explanations: Effects of a chatbot support system on teacher students' diagnostic competence [Vortrag]. *13th Conference of European Researchers in Didactics of Biology (ERIDOB)*, Nicosia, Cyprus.
- Sep. 2022 **Vorstellung und Einsatz der Klassenräume zu Evolution und Experimentieren beim Science Day** im Rahmen des Festivals der Wissenschaft am 30. September 2022 am IPN in Kiel
- Nov. 2022 **Projektvorstellung beim Retreat Didaktik der Biologie**  
Fiedler, D. (2022, 03. November). Förderung der Diagnosekompetenz angehender Biologielehrkräfte im Bereich Evolution, Experimentieren und Argumentieren – Das FiSK-Projekt [Vortrag]. *Retreat Didaktik der Biologie*, Akademiezentrum Sankelmark in Oeversee (Schleswig-Holstein).
- April 2023 **Vortrag bei der Fachtagung “96th Annual NARST Conference”**  
Fiedler, D., Schönle, D., Reich, C. & Harms, U. (2023, 20. April). Applying a classroom simulation with chatbot to support pre-service biology teachers' diagnostic competence in evolution [Vortrag]. *96th Annual Conference of National Association of Research in Science Teaching (NARST)*, Chicago, USA.
- Mai 2023 **Vorstellung des SKR<sup>Bio</sup> beim digitalen Kick-off Meeting des Cluster 2c (DigiProMIN)**  
Fiedler, D. & Harms, U. (2023, 17. Mai). Die Klassenraumsimulation SKR<sup>Bio</sup> [Vortrag]. *Kick-Off Meeting des Cluster 2c* des vom BMBF geförderten Verbundprojekts

*Digitalisierungsbezogene und digital gestützte Professionalisierung von MIN-Lehrkräften (DigiProMIN) im Programm lernen:digital, Online.*

- Juni 2023 **Projektvorstellung beim Retreat Didaktik der Biologie**  
Fiedler, D. (2023, 29. Juni). When practical situations are rare: Improving diagnostic ability of pre-service biology teachers using a classroom simulation with chatbot [Vortrag]. *Retreat Didaktik der Biologie*, Akademiezentrum Sankelmark in Oeversee (Schleswig-Holstein).
- Juli 2023 **Projektvorstellung bei der 2. IPN-Sommertagung**  
Fiedler, D., Harms, U., Schönle, D. & Reich, C. (2023, 17. Juli). Effekte adaptiver Feedbackbots im Simulierten Klassenraum auf prozedurales Professionswissen [Poster]. *Projektvorstellung während der IPN-Sommertagung*, Kiel.
- Juli 2023 **Vorstellung des SKR<sup>Bio</sup> mit Chatbot beim IPN-Diagnosekompetenz-Netzwerk**  
Fiedler, D. (2023, 27. Juli). Effekte adaptiver Feedbackbots im Simulierten Klassenraum auf prozedurales Professionswissen. *Projektvorstellung während des IPN-Diagnosenetzwerk-Treffens*, Kiel.
- Sep. 2023 **Vortrag bei der Fachtagung "24. Internationale Tagung der FDdB im VBIO"**  
Fiedler, D., Schönle, D., Reich, C., & Harms, U. (2023, 17. September). When practical situations are rare: Improving pre-service biology teachers' diagnostic competency in a classroom simulation with chatbot. *24. Internationale Tagung der Fachsektion Didaktik der Biologie (FDdB im VBio)*, Ludwigsburg.
- Nov. 2023 **Vorstellung des Projekts FiSK und des SKRBio mit Chatbot bei der Evaluation des IPN durch die Leibniz-Gemeinschaft am 6. November 2023 in Kiel.**
- März 2024 **Vortrag bei der Fachtagung "97th Annual NARST Conference"**  
Fiedler, D., Schönle, D., Reich, C. & Harms, U. (2024, 18. März). Boosting diagnostic competence in evolution using chatbots in classroom simulations: Insights into an explorative study [Vortrag]. *97<sup>th</sup> Annual Conference of National Association of Research in Science Teaching (NARST)*, Denver, USA.
- März 2024 **Vortrag bei der Fachtagung "12. Tagung der GEBF"**  
Poser, J., Fiedler, D., Schönle, D., Reich, C. & Harms, U. (2024, 18. März). Messung der Diagnosekompetenz zum Experimentieren mit der Klassenraumsimulation SKRBio [Vortrag]. *12. Tagung der Gesellschaft für Empirische Bildungsforschung (GEBF)*, Potsdam.
- Mai 2024 **Vortrag beim IPN-Forschungskolloquium**  
Poser, J., Fiedler, D., Schönle, D., Reich, C. & Harms, U. (2024, 30. Mai). Diagnosekompetenz von Lehramtsstudierenden im Bereich Experimentieren - Eine Untersuchung mit der Klassenraumsimulation SKR<sup>EXP</sup> mit Chatbot [Vortrag]. *IPN-Forschungskolloquium*, Kiel.
- Juli 2024 **Vorstellung des SKR<sup>Bio</sup> beim Kongress "Digitales und digital gestütztes Unterrichten"** am 24. Juli 2024 in Kiel.

### 2.5.2 Erfolgte Veröffentlichungen

Im Folgenden sind alle veröffentlichten Ergebnisse und Erkenntnisse aufgeführt, die während des Förderzeitraums von FiSK entstanden sind:

1. Fiedler, D., Baer, D., & Harms, U. (2024). *Die Klassenraumsimulation SKR<sup>Bio</sup> im Biologie-Lehramtsstudium: Fähigkeiten zur Diagnose von Argumenten messen und fördern*. In N. Graulich, J. Arnold, S. Sorge, & M. Kubsch (Eds.), *Lehrkräftebildung von morgen: Beiträge der Naturwissenschaftsdidaktiken zur Förderung überfachlicher Kompetenzen*. (p. 291-300). Waxmann Verlag. <https://doi.org/10.31244/9783830997962.31>
2. Fiedler, D., & Harms, U. (2024). Diagnosefähigkeit im Biologie-Lehramtsstudium messen und fördern: Einsatz der Klassenraumsimulation SKR<sup>Bio</sup> mit und ohne Chatbot-Unterstützung. *IPN-Journal 11*, <https://static.leibniz-ipn.de/journal11/#0>
3. Fiedler, D., Poser, J., & Harms, U. (2024). Teacher questions and student answers for the SCR<sup>Bio</sup> in the context of evolution [Data set]. *PsychArchives*. <https://doi.org/10.23668/psycharchives.14404>
4. Schönle, D., Fiedler, D., Harms, U., Reich, C., & Ould Abdeslam, D. (2024). Transformative educational technologies: Chatbot integration and teacher trainee reception in simulated classrooms. In *The 16th annual International Conference on Education and New Learning Technologies, 2024*. (S. 7406-7416)
5. Schönle, D., Reich, C., Ould Abdeslam, D., Fiedler, D., Harms, U., & Poser, J. (2024). Automatic Assessment of Student Answers using Large Language Models: Decoding Didactic Concepts. In *The 2024 IARIA Annual Congress on Frontiers in Science, Technology, Services, and Applications*. (S. 158-167)

### 2.5.3 Geplante Veröffentlichungen

Die digitalen Forschungsdaten sind bei GESIS archiviert und werden nach Ablauf des Embargos im August 2026 mit einer DOI veröffentlicht und frei für Forschung und Lehre zugänglich gemacht (vgl. [Kapitel 2.4 Nutzen sowie Verwertbarkeit der Ergebnisse und Erfahrungen](#)). Die Daten werden zudem in der Dissertation von Johannes Poser weiterverwendet. Nachfolgend sind alle geplanten Veröffentlichungen aufgeführt, die voraussichtlich in den nächsten zwei Jahren erscheinen werden.

1. Fiedler, D., Schönle, D., Reich, C., & Harms, U. (08/2024). Impact of simulation-chatbot system on teachers' diagnostic skills. *Manuskript unter Begutachtung*.
2. Poser, J., Harms, U., Schönle, D., Reich, C., & Fiedler, D. (12/2024). Simulation-based analysis of the content-specificity of pre-service biology teachers' judgment accuracy. *Manuskript in Vorbereitung*.
3. Harms, U., Fiedler, D., Czincel, B. (12/2024). *Explaining evolutionary change in the digital age: Challenges and perspectives for digital technology in evolution education*. In K. Kampourakis & R. H. Nehm (Eds.), *Explanation in Biology Education: Theory and Practice*. *Manuskript in Vorbereitung*.

## Literaturverzeichnis

Barrett, M., Branson, L., Carter, S., DeLeon, F., Ellis, J., Gundlach, C., & Lee, D. (2019). Using Artificial Intelligence to Enhance Educational Opportunities and Student Services in Higher Education. *The Journal of the Virginia Community Colleges*, 22(1), 11.

- Bromme, R. (2001). Teacher Expertise. In N. J. Smelser, P. B. Baltes, & F. E. Weinert (Hrsg.), *International Encyclopedia of the Behavioral Sciences: Education*. (S. 15459–15465).
- Dicke, T., Holzberger, D., Kunina-Habenicht, O., Linninger, C., Schulze-Stocker, F., Seidel, T., Terhart, E., Leutner, D., & Kunter, M. (2016). „Doppelter Praxisschock“ auf dem Weg ins Lehramt? Verlauf und potenzielle Einflussfaktoren emotionaler Erschöpfung während des Vorbereitungsdienstes und nach dem Berufseintritt. *Psychologie in Erziehung und Unterricht*, 63(4), 244-257.
- Dutta, D. (2017). Developing an Intelligent Chat-bot Tool to assist high school students for learning general knowledge subjects. *Georgia Institute of Technology*.
- Fischer, J., Machts, N., Bruckermann, T., Möller, J., & Harms, U. (2022). The Simulated Classroom Biology—A simulated classroom environment for capturing the action-oriented professional knowledge of pre-service teachers about evolution. *Journal of Computer Assisted Learning*, 38(6), 1765-1778.
- Fischer, J., Machts, N., Möller, J., & Harms, U. (2021). Der Simulierte Klassenraum Biologie – Erfassung deklarativen und prozeduralen Wissens bei Lehramtsstudierenden der Biologie. *Zeitschrift für Didaktik der Naturwissenschaften*, 27(1), 215–229.
- Fleckenstein, J., Meyer, J., Jansen, T., Keller, S. D., Köller, O., & Möller, J. (2024). Do teachers spot AI? Evaluating the detectability of AI-generated texts among student essays. *Computers and Education: Artificial Intelligence*, 6, Article 100209.
- Förtsch, C., Sommerhoff, D., Fischer, F., Fischer, M. R., Girwidz, R., Obersteiner, A., Reiss, K., Stürmer, K., Siebeck, M., Schmidmaier, R., Seidel, T., Ufer, S., Wecker, C., & Neuhaus, B. J. (2018). Systematizing professional knowledge of medical doctors and teachers: development of an interdisciplinary framework in the context of diagnostic competences. *Education Sciences*, 8(4), 207–225.
- Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. (2009). Teaching Practice: A Cross-Professional Perspective. *Teachers College Record*, 111(9).
- Hillmayr, D., Ziernwald, L., Reinhold, F., Hofer, S. I., & Reiss, K. M. (2020). The potential of digital tools to enhance mathematics and science learning in secondary schools: A context-specific meta-analysis. *Computers & Education*, 153, article 103897.
- Howell, H., & Mikeska, J. N. (2021). Approximations of practice as a framework for understanding authenticity in simulations of teaching. *Journal of Research on Technology in Education*, 53(1), 8–20.
- Hussain, S., Sianaki, O. A., & Ababneh, N. (2019). A survey on conversational agents/chatbots classification and design techniques. In *Workshops of the International Conference on Advanced Information Networking and Applications* (pp. 946-956). Springer.
- Hwang, G.-J. & Chang, C.-Y. (2023). A review of opportunities and challenges of chatbots in education. *Interactive Learning Environments*, 31(7), 4099-4112.
- Klusmann, U., Kunter, M., Voss, T., & Baumert, J. (2012). Berufliche Beanspruchung angehender Lehrkräfte: Die Effekte von Persönlichkeit, pädagogischer Vorerfahrung und professioneller Kompetenz. *Zeitschrift für Pädagogische Psychologie*, 26, 275-290.
- Leuders, T., Loibl, K., Sommerhoff, D., Herppich, S., & Praetorius, A.-K. (2022). Toward an overarching framework for systematizing research perspectives on diagnostic thinking and practice. *Journal für Mathematik-Didaktik*, 43(1), 13-38.
- Molnár, G., & Szüts, Z. (2018). The Role of Chatbots in Formal Education. In *International Symposium on Intelligent Systems and Informatics* (pp. 197–202).

Yang, S., & Evans, C. (2019). Opportunities and Challenges in Using AI Chatbots in Higher Education. In Proceedings of the 2019 3rd International Conference on Education and E-Learning (pp. 79–83).

## Kurzbericht

---

### **Effekte adaptiver Feedbackbots im Simulierten Klassenraum auf prozedurales Professionswissen (FiSK) - Teilvorhaben IPN: Naturwissenschaften**

Diagnostische Fähigkeiten sind entscheidend für die professionelle Kompetenz von Lehrkräften. Sie bilden die Grundlage dafür, den Wissensstand von Schüler:innen zu beurteilen und in die Unterrichtsgestaltung einzubeziehen. Besonders bei Berufseinsteigern zeigt sich oft eine Diskrepanz zwischen theoretischem Wissen (bspw. zum Diagnostizieren) und dessen praktischer Anwendung. Simulationsbasierte Lernumgebungen und der Einsatz von Bots (basierend auf Künstlicher Intelligenz) bieten einen vielversprechenden Ansatz, den Übergang von der Theorie in die Praxis zu erleichtern und dabei zugleich individuelle Unterstützung zu geben.

Das Projekt FiSK hatte zum Ziel, die Diagnosefähigkeit von Biologielehramtsstudierenden mithilfe der Klassenraumsimulation SKR<sup>Bio</sup> zu fördern. Die Klassenraumsimulation SKR<sup>Bio</sup> ermöglichte es den Studierenden, Erklärungen von virtuellen Schüler:innen zur biologischen Evolution fachlich zu beurteilen. Um den SKR<sup>Bio</sup> für den Einsatz in der Lehrkräftebildung weiterzuentwickeln, wurde ein Chatbot integriert, der automatisiertes Feedback zu fachlichem und fachdidaktischem Wissen gibt. Gleichzeitig sollte der SKR<sup>Bio</sup> um weitere relevante Themenbereiche wie das Experimentieren und Argumentieren erweitert werden. Die Wirksamkeit dieser Entwicklungen wurde durch experimentelle Studien untersucht, um Faktoren zu identifizieren, die die Diagnosefähigkeit angehender Biologielehrkräfte fördern oder behindern.

Die Ergebnisse der experimentellen Studien zeigen, dass der SKR<sup>Bio</sup> durch seine thematische Erweiterung nun gezielt als Instrument zur Messung der Diagnosefähigkeit von Biologie-Lehramtsstudierenden in den Bereichen Evolution, Experimentieren und Argumentieren eingesetzt werden kann. Die Studie verdeutlicht unterschiedliche Fähigkeitsprofile der Studierenden in diesen Bereichen, wobei Diagnosen im Bereich Evolution den Studierenden tendenziell leichter fallen als im Bereich Argumentieren. Besonders im Bereich Argumentieren konnte durch die Nutzung des SKR<sup>Bio</sup>-Chatbot-Systems eine signifikante Verbesserung der Diagnosefähigkeit erzielt werden. Die Nutzung des Chatbots variierte je nach Themenbereich, wobei der größte Bedarf an Unterstützung im Bereich Argumentieren festgestellt wurde.

Das SKR<sup>Bio</sup>-Chatbot-System lässt sich lernförderlich in die Lehre integrieren und bietet den Studierenden eine innovative digitale Lerngelegenheit, ihr diagnostisches fachliches Wissen in einem geschützten Raum unter hoher Reduzierung der Komplexität einer realen Klassenraumsituation gezielt zu testen und zu vertiefen. Besonders eine wiederholte Nutzung könnte dabei helfen, Routinen für diagnostische Prozesse zu entwickeln, die später im Unterricht angewendet werden können. Darüber hinaus ermöglicht der Chatbot autarkes Lernen im SKR<sup>Bio</sup>, da fachliche und fachdidaktische Fragen unabhängig von einem Dozenten oder einer Dozentin geklärt werden können.

Die Ergebnisse des Projekts bieten wertvolle Einblicke, wie digitale Lernsysteme gezielt die Diagnosefähigkeiten angehender Lehrkräfte fördern können, insbesondere in den frühen Phasen der Lehrkräftebildung. Die digitalen Forschungsdaten wurden in Übereinstimmung mit den Anforderungen guter wissenschaftlicher Praxis und den Datenschutzvorgaben an GESIS zur Archivierung übermittelt. Nach Ablauf des Embargos im August 2026 werden die Daten mit einer DOI versehen und für Forschung und Lehre öffentlich zugänglich gemacht werden.