

**Using an Intelligent Tutoring System within a Task-Based Learning Approach in
English as a Foreign Language Classes to Foster Motivation and Learning Outcome
(Interact4School):**

Pre-registration of the Study Design

Parrisius, C.¹, Pieronczyk, I.¹, Blume, C.², Wendebourg, K.¹, Pili-Moss, D.³, Assmann, M.³,
Beilharz, S.³, Bodnar, S.⁴, Colling, L.⁴, Holz, H.⁴, Middelanis, L.², Nuxoll, F.⁴, Schmidt-
Peterson, J.¹, Meurers, D.⁴, Nagengast, B.¹, Schmidt, T.³, & Trautwein, U.¹

¹ Hector Research Institute of Education Sciences and Psychology, University of Tübingen,
Germany

² Competence Centre for Teacher Education and Research (DoKoLL), Technical University
Dortmund, Germany

³ Institute of English Studies, Leuphana University Lüneburg, Germany

⁴ Department of Linguistics, University of Tübingen, Germany

Date of pre-registration: 31 January 2022

Note

The study is ongoing. In some classes, data has already been collected at three of the five timepoints. The questionnaire data, logfile data, and English proficiency test data have not yet been examined or edited, nor have any analyses for study purposes been carried out. However, some logging information and proficiency test information has been looked at from the system developers to ensure the system behaves as intended.

The current pre-registration exclusively serves the purpose to comprehensively describe the *study design* that was realized for the evaluation of the FeedBook. Precise research questions and the corresponding hypotheses will be described in more detail in subsequent amendments to this pre-registration and will be focused on in separate projects. Amendments will be accessible via PsychArchives (<https://doi.org/10.23668/psycharchives.5366>) at “Related PsychArchives Objects.”

Introduction

In second language learning, frequent individual feedback is essential to help learners to develop an accurate and fluent use of the language (Nassaji & Kartchava, 2017). However, in real-life settings, teachers are often confronted with challenges in providing adequate feedback on their students' individual progress and performance. Among other potential advantages, intelligent tutoring systems (ITS) can address these high demands on second language school teachers (Würffel, 2017). ITS offer new possibilities to provide different forms of feedback to the students, and their effectiveness has been shown in several meta-analyses (Kulik & Fletcher, 2016, Ma et al., 2014, Steenbergen-Hu & Cooper, 2014). For instance, Meurers et al. (2019) demonstrated that for specific, form-focused exercises with which students practiced certain grammatical constructions during individual learning phases, individualized feedback provided by an ITS led to significant learning progress regarding these constructions. Taken together, digital environments such as ITS enable new ways of learning and teaching, and have the potential to foster learning outcome, but also motivation and teaching quality (Ahmadi, 2018).

The Interact4School project investigates individualized learning as a component of institutionalized schooling through the use of an ITS (the online workbook *FeedBook*) for second language learning in a multisite, cluster-randomized controlled field trial. The project systematically addresses the effectiveness of various forms of adaptive individualized student support provided by the FeedBook on learning progress and motivation. The study is being conducted in the school year 2021/22.

The FeedBook and its Features

The FeedBook is a web-based English workbook for second-language learners. It is embedded in a task-based language learning (TBLL; Ellis, 2003; Hallet & Legutke, 2013) approach and comprises English exercises that address grammar topics, lexis, pragmatics, and

listening comprehension aligned with the seventh-grade English curricula of the various German federal states in academic-track schools. It is an ITS insofar as that it is enriched with automatic, interactive, scaffolded feedback (Rudzewitz et al., 2018) on forms and meaning. That is, not only does the FeedBook indicate to students whether their answers in the exercises were correct or incorrect; it also anticipates misconceptions that underlie incorrect answers, and provides the students with scaffolded feedback so that they can correct their answers step by step. The FeedBook and its feedback algorithms were developed in a year-long process and were piloted in the school year 2017/18 (Rudzewitz et al., 2017). In a subsequent first efficacy trial, Meurers et al. (2019) provided evidence that students learned substantially more regarding those language means for which they received scaffolded feedback in the FeedBook.

For the current trial, the FeedBook exercises were newly developed and are didactically embedded in lesson plans that have a TBLL orientation. Moreover, the existing feedback that the system provides was expanded to also include criterial feedback and motivational elements. That is, a student dashboard was implemented that provides the students with information about their learning progress and performance level (criterial feedback; see Figure 1). Additionally, motivational elements were added, which include, for example, the display of incentive phrases ("Well done! Next time, you'll do even better.") and guidance from a digital pedagogical agent (see Figure 2). It is technically possible to provide students with only a subset of features.

Figure 1

Example for Criterial Feedback: Student Dashboard with Overview of Submitted Exercises and Achieved Mastery (Trophy)

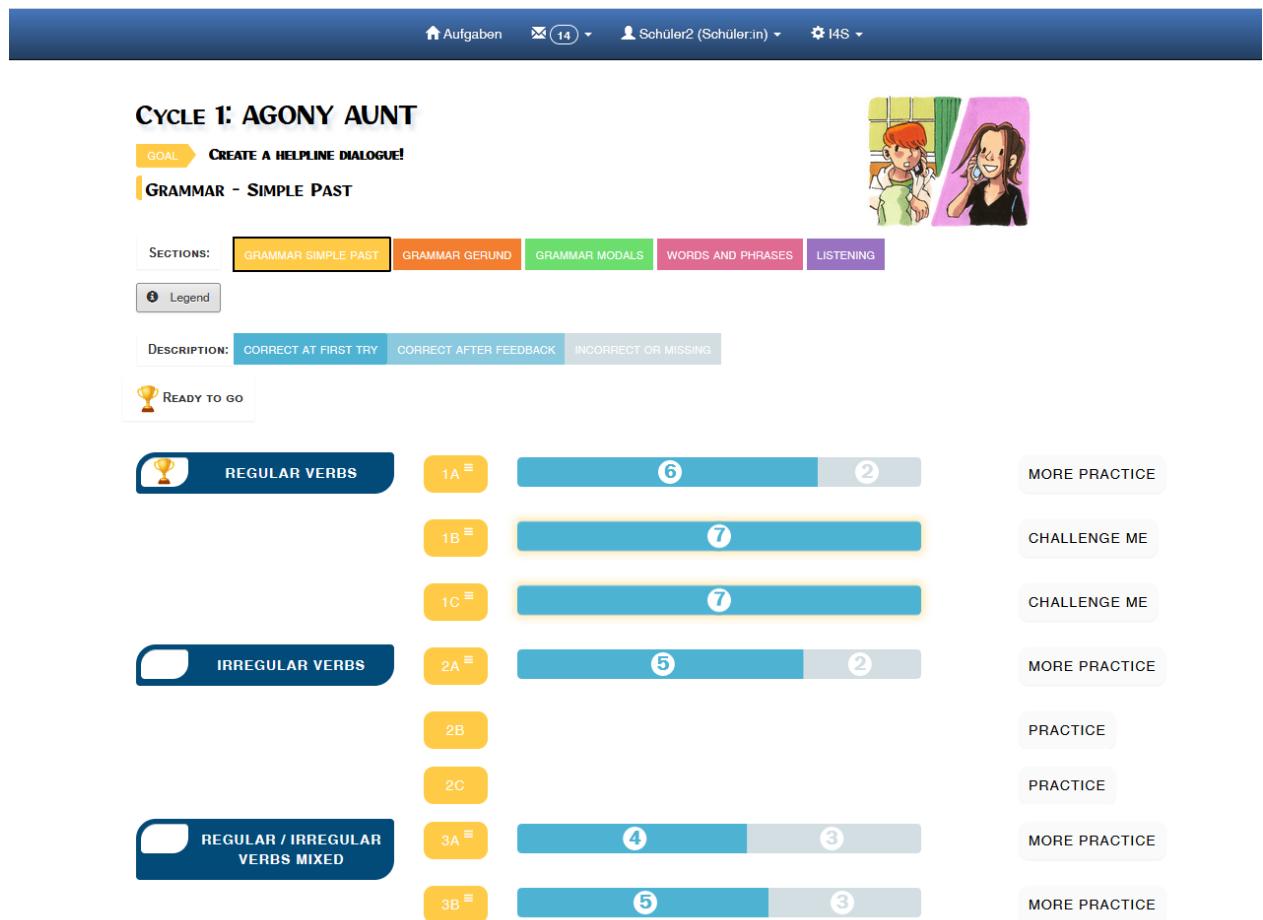
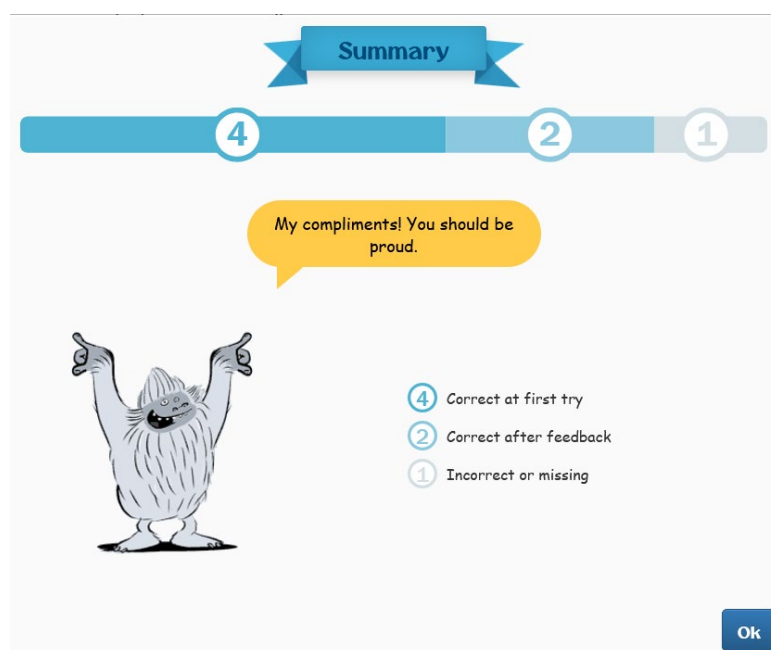


Figure 2

Example for Motivational Elements: Incentive Phrase and Pedagogical Agent



Research Questions

The effects of the FeedBook in terms of learning objectives in seventh-grade English will be examined by addressing the following main questions in a one-year-long study of regular English classes using the FeedBook. By making use of either data from the full sample (Projects 1 to 4) or from a subset of students from one participating school in the federal state of Hamburg (Projects 5 and 6), for which additional qualitative data was and will be collected, we will investigate the following questions:

- Do students show higher English proficiency when they are assigned to work with the FeedBook compared to business-as-usual control classes? (*Project 1*)
- Does criterial feedback linking the language means being practiced to the functional language task for which they are useful foster students' English proficiency? (*Project 2*)

- Does criterial and/or motivational feedback embedded in the FeedBook enhance students' motivation? (*Project 3*)
- Do students show higher English proficiency for those language means for which they receive individualized grammatical feedback? (*Project 4*)
- How does the use of the FeedBook relate to linguistic outcomes (English proficiency), transfer of linguistic structures in communicative activities, and classroom-based interaction? How do students' cognitive individual differences (long-term and working memory) relate to the same set of outcomes? (*Project 5*)
- How do teachers' attitudes and competence regarding the use of intelligent tutoring systems, task-based language learning, and pedagogical issues such as diagnosis and feedback develop with use of the FeedBook? (*Project 6*)

Each research question and the corresponding hypotheses will be described in more detail in subsequent amendments to this pre-registration and will be focused on in separate projects (amendments will be accessible via PsychArchives

(<https://doi.org/10.23668/psycharchives.5366>) at “Related PsychArchives Objects”). An additional focus of the study lies on the extensive implementation of explorative learning analytics. The current pre-registration exclusively serves the purpose to comprehensively describe the *study design* that was realized for the evaluation of the use of the FeedBook within a TBLL approach, including the study conditions and measures.

Method

Sample and Procedure

Data for this large-scale test of the FeedBook are currently being collected in seventh-grade classrooms in academic track schools in three German federal states (namely, Baden-Württemberg, North Rhine-Westphalia, and Hamburg) from September 2021 to approximately June 2022. For Baden-Württemberg, the Ministry of Education and Cultural

Affairs in Baden-Württemberg approved the study and the collection of the data (date of approval: July 19, 2021, file number: 31-6499.21/629/1). For Hamburg, the Authority for School and Vocational Training approved the study and the collection of the data (date of approval: July 13, 2021, file number e514.101.5000-002/221,025). In line with school law in North Rhine-Westphalia, according to which no superior ministry is responsible for approving school studies, the study and the collection of the data in this federal state were approved by the individual headmasters from the respective participating schools. The Ethics Committee for Psychological Research at the University of Tübingen confirmed that the study procedures were in line with ethical standards of research with human subjects (date of approval: August 4, 2021, file number: A2.5.4-184_ns).

Treatment and Conditions

The FeedBook contains exercises that are aligned with, and prepare for, a total of four complex Target Tasks that require the integration of several skills and competences, such as the use of certain grammatical structures, the knowledge of certain vocabulary fields, and listening, speaking or writing skills. As these Target Tasks are generally communicative, they are carried out in class. The process of preparation for a Target Task is called a Task Cycle. The FeedBook contains digital practice material for each of the four Task Cycles. Each Task Cycle is planned to last approximately 3 weeks, respectively. The FeedBook also provides teachers with detailed lesson plans, which give them ideas of how to optimally integrate the digital exercises (see Method section for more detail).

In order to investigate the aforementioned research questions, students were randomly assigned to different conditions (see Table 1). To address the research question for Project 1, schools were randomly assigned to either work with the FeedBook and supplemental material in the school year 2021/22 (*FeedBook condition*) or to work with the FeedBook in the school year 2022/23 (*waitlist control condition*). To address the research questions for Projects 2 and

3, classes were randomly assigned to work with the original version of the FeedBook (*original FeedBook condition*; see Meurers et al., 2019), to work with the FeedBook in which, additionally, criterial feedback is provided (*criterial feedback condition*; see Figure 1), or to work with the FeedBook with the criterial feedback and in which, additionally, motivational elements are implemented (*criterial feedback & motivational elements condition*; see Figure 2). And finally, to address the research question for Project 4, students were randomly assigned to *Group A*, which receives individual, corrective feedback for all language constructs in Task Cycles 1 and 3, or to *Group B*, which receives individual, corrective feedback for all language constructs in Task Cycles 2 and 4. While working on a Task Cycle, one half of the students always receive scaffolded feedback regarding specific language constructs or grammatical structures that are primarily focused on by the respective Task Cycle (e.g., Conditional Clauses Type 2) (see Figure 3). The other students do not receive this feedback for the same specific grammatical forms and constitute the respective control group with respect to specific grammatical forms. The allocation of feedback/no feedback alternates with each Task Cycle, so that by the end of the school year, all students in a class have had an equal opportunity to benefit from the various forms of grammatical feedback in the FeedBook. There were no specific conditions realized to address the other projects.

The teachers working with the FeedBook have access to a wide range of teaching material, including lesson plans, handouts and other additional paper-based exercise sheets, solutions, PowerPoint-Slides, and additional video material for viewing and listening comprehension. Teachers were asked to use the teaching material while working on the Task Cycles and to conduct the Target Tasks during class.

The intervention materials (i.e., the English exercises used in the FeedBook as well as the additional material provided to the teachers to prepare their students for the Target Tasks in class, including paper-based handouts, slides for presentations, and videos) were developed

Table 1

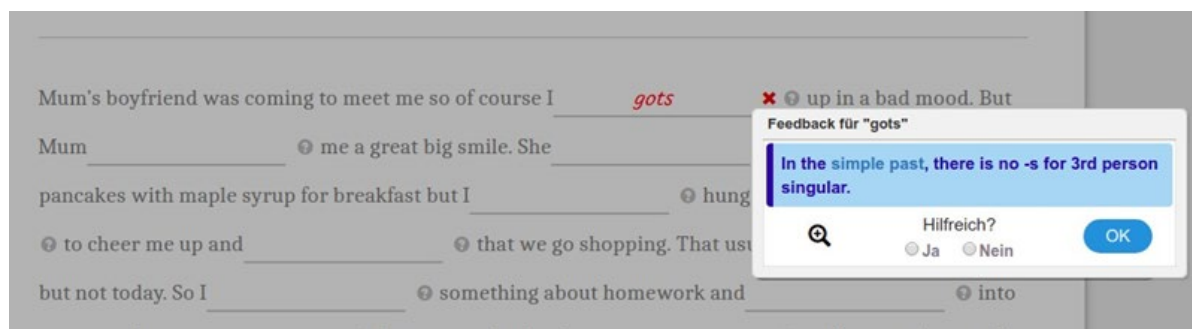
Intervention Conditions

Assignment						
Conditions at the school level: Use of FeedBook	In school year 2021/22 (FeedBook condition)					In school year 2022/23 (waitlist control condition)
Conditions at the class level: Criterial feedback and motivational elements	Original FeedBook condition		Criterial feedback condition		Criterial feedback & motivational elements condition	
Conditions at the individual level: Grammatical feedback	Group A	Group B	Group A	Group B	Group A	Group B

Note: Students in Group A receive extensive grammatical feedback on language constructs targeted in Task Cycles 1 and 3, whereas students in Group B receive extensive grammatical feedback on language constructs targeted in Task Cycles 2 and 4. For the language constructs targeted in the respective other Task Cycles, the groups receive only limited grammatical feedback.

Figure 3

Example for Scaffolded Feedback



by a group of experts, consisting of teachers in Hamburg and Baden-Württemberg as well as researchers of English didactics. The entire teaching material is based on the contents of the textbook Camden Town 3 by the Westermann publishing house, and is thus aligned with the year seven English curricula of the various German federal states in academic-track schools.

Study Design

We developed a design with which we were able to realize a number of intervention conditions and to test for all our research questions. Random assignment was only implemented for the schools in the federal states of Baden-Württemberg and North Rhine-Westphalia (not Hamburg). In Hamburg, the quantitative data obtained from this study is supplemented by qualitative data from three seventh-grade classes in one school, including interviews with the teachers and video recordings from multiple lessons. The focus with this one participating school in Hamburg is on integrating the qualitative and quantitative studies in a mixed-method design (Riazi & Candlin, 2014). Therefore, this school was contacted ahead of time and was exempted from the random assignment to experimental conditions. Instead, the school and its three participating teachers with three classes was manually assigned to the *criteria feedback & motivational elements condition*. Despite the deliberate assignment of the school as a whole to this condition, the random assignment of students within these classes to Group A or Group B was kept intact.

In the remaining two federal states, we realized different designs to answer our research questions (see also Table 2). That is, ...

1. to address the first research question (effectiveness of FeedBook use), schools (within federal states) were initially planned to be randomly assigned to the *FeedBook condition* (vs. the *waitlist control condition*) with a probability of 0.67. We updated this probability after recruitment, and realized a probability of 0.75 to be randomly assigned to the FeedBook condition (vs. the waitlist control condition; see Recruitment for more detail).
2. to address the second and third research questions (effectiveness of criteria feedback and motivational elements to foster English proficiency and student motivation), teachers (within schools) in the FeedBook condition along with

their classes were randomly assigned to one of three conditions (i.e., *original FeedBook condition*, *criterial feedback condition*, and *criterial feedback & motivation elements condition*) with a probability of 0.33, respectively.

3. to address the fourth research question (effect of individualized grammatical feedback on English proficiency), individual students (within classes) in the FeedBook condition were randomly assigned to *Group A* (vs. *Group B*) with a probability of 0.50.

Table 2

Overview of the Research Questions, Subsequent Information for the Power Analyses, and Results of the Power Analyses (MDES)

Projects & research questions	Conditions	Primary outcome	Design for power analyses	MDES from initial power analyses	MDES from updated power analyses
<i>Project 1: Do students show higher English proficiency when they are assigned to work with the FeedBook compared to business-as-usual control classes?</i>	<i>FeedBook condition vs. waitlist control condition (and FeedBook condition vs. extended waitlist control condition)</i>	English proficiency (i.e., student-level variable)	Three-level blocked cluster-level random assignment design (i.e., students nested in teachers/classes, nested in schools, blocked in federal states) with treatment at Level 3 (schools)	Between $\delta = 0.38$ 95% CI [0.11,0.65] and $\delta = 0.23$ 95% CI [0.07,0.40]	Between $\delta = 0.79$ 95% CI [0.19,1.40] and $\delta = 0.48$ 95% CI [0.11,0.85]
<i>Project 2: Does criterial feedback linking the language means being practiced to the functional language task for which they are useful foster students' English proficiency?</i>	<i>Original FeedBook condition vs. criterial feedback condition vs. criterial feedback & motivational elements condition</i>	English proficiency (i.e., student-level variable)	Three-level blocked cluster-level random assignment design (i.e., students nested in teachers/classes, nested in schools, blocked in federal states) with treatment at Level 2 (classes)	Between $\delta = 0.30$ 95% CI [0.08,0.51] and $\delta = 0.24$ 95% CI [0.07,0.41] (comparing each two conditions)	Between $\delta = 0.50$ 95% CI [0.13,0.88] and $\delta = 0.40$ 95% CI [0.10,0.70] (comparing each two conditions)
<i>Project 3: Does criterial and/or motivational feedback embedded in the FeedBook enhance students' motivation?</i>		Student motivation (i.e., student-level variable)			

<i>Project 4:</i> Do students show higher English proficiency for those language means for which they receive individualized grammatical feedback?	<i>Group A vs. Group B</i>	English proficiency (i.e., student-level variable)	Three-level blocked individual-level random assignment design (i.e., students blocked in teachers/classes and in schools) with treatment at Level 1 (individuals)	Between $\delta = 0.102$ 95% CI [0.03,0.17] and $\delta = 0.091$ 95% CI [0.03,0.16]	Between $\delta = 0.120$ 95% CI [0.04,0.20] and $\delta = 0.106$ 95% CI [0.03,0.18]
<i>Project 5:</i> How does the use of the FeedBook relate to linguistic outcomes (English proficiency), transfer of linguistic structures in communicative activities, and classroom-based interaction? How do students' cognitive individual differences (long-term and working memory) relate to the same set of outcomes?			This research question will be investigated following a mixed-method approach.		

Project 6: How do teachers' attitudes and competence regarding the use of intelligent tutoring systems, task-based language learning, and pedagogical issues such as diagnosis and feedback develop with use of the FeedBook?

This research question will be investigated following a mixed-methods approach.

Note. MDES = minimal detectable effect size.

Initial Power Analyses. To determine the sample size in the two federal states in which random assignment was planned (i.e., Baden-Württemberg and North Rhine-Westphalia), we conducted multiple power analyses in R 4.1.2 (R Core Team, 2021) and the PowerUpR package (v1.1.0; Bulus et al., 2021). We ran all power analyses under the constraint to achieve a power of 80% and we set the alpha level to 5% (for two-tailed testing). We aimed at achieving acceptable minimal detectable effect sizes (MDES) per design under a given number of schools. To derive realistic estimates for the necessary parameters, we drew on data from the previous FeedBook study (Meurers et al., 2019), from educational RCTs that the project members have previously designed and implemented (e.g., Gaspard et al., 2015; Gaspard et al., 2021), and from published estimates (e.g., Stallasch et al., 2021). An overview of these parameters are presented in Table 3. Based on the observed parameters, we defined two scenarios (conservative scenario and optimistic scenario) with sets of parameters that denote two extremes around the observed parameters from previous studies. For each design (i.e., at the school level, at the class level, and at the individual level), we ran one power analysis per scenario (conservative vs. optimistic) to mark the range of effect sizes in which the MDES can be expected to fall. The respective assumptions and the results of the power analyses can be obtained at the following link: <https://doi.org/10.23668/psycharchives.5366>.

Given our resources, a multisite cluster-randomized trial with 16 schools (with three classes per school and 24 students per class) from two federal states seemed to represent an ideal balance between MDES and feasibility. Therefore, we set out to recruit $N_S = 16$ schools from the two federal states (i.e., eight schools per federal state) with a total of $N_C = 48$ classes and $n = 1,152$ students. Under the assumption of this total sample size, the results of the power analyses indicated that...

1. at the school level, we would be able to detect intervention effects of at least $\delta = 0.38$, 95% CI [0.11,0.65] (conservative) to $\delta = 0.23$, 95% CI [0.07,0.40]

(optimistic) comparing the FeedBook condition with the waitlist control condition,

2. at the class level, we would be able to detect intervention effects of at least $\delta = 0.30$, 95% CI [0.08,0.51] (conservative) to $\delta = 0.24$, 95% CI [0.07,0.41] (optimistic) comparing each two conditions at that level, and
3. at the individual level, we would be able to detect intervention effects of at least $\delta = 0.102$, 95% CI [0.03,0.17] (conservative) to $\delta = 0.091$, 95% CI [0.03,0.16] (optimistic) comparing Group A with Group B per Task Cycle.

Table 3

Overview of Parameters Included in the Respective Power Analyses

Design	Considered parameters	Conservative scenario	Optimistic scenario
Three-level blocked cluster-level random assignment design with treatment at Level 3 (school)	<ul style="list-style-type: none"> Intraclass correlation at the class level (ρ_2) Intraclass correlation at the school level (ρ_3) Amount of variance at individual level explained by a pretest measure used as a covariate (R_1^2) Amount of variance at class level explained by a pretest measure used as a covariate (R_2^2) Amount of variance at the school level explained by a pretest measure used as a covariate (R_3^2) Probability of assignment to FeedBook condition (vs. waitlist control condition) 	<ul style="list-style-type: none"> $\rho_2 = 0.15$ $\rho_3 = 0.25$ $R_1^2 = 0.50$ $R_2^2 = 0.50$ $R_3^2 = 0.90$ 0.67 	<ul style="list-style-type: none"> $\rho_2 = 0.07$ $\rho_3 = 0.02$ $R_1^2 = 0.50$ $R_2^2 = 0.50$ $R_3^2 = 0.90$ 0.67
Three-level blocked cluster-level random assignment design with treatment at Level 2 (class)	<ul style="list-style-type: none"> Intraclass correlation at the class level (ρ_2) Intraclass correlation at the school level (ρ_3) Treatment effect heterogeneity (ω) Amount of variance at individual level explained by a pretest measure used as a covariate (R_1^2) 	<ul style="list-style-type: none"> $\rho_2 = 0.05$ $\rho_3 = 0.005$ $\omega = 0.002$ $R_1^2 = 0.50$ 	<ul style="list-style-type: none"> $\rho_2 = 0.005$ $\rho_3 = 0.005$ $\omega = 0.002$ $R_1^2 = 0.50$

	<ul style="list-style-type: none"> • Amount of variance at class level explained by a pretest measure used as a covariate (R_2^2) • Amount of variance at the school level explained by a pretest measure used as a covariate (R_3^2) • Probability of assignment to one out of three conditions (i.e., original FeedBook condition vs. criterial feedback condition vs. criterial feedback & motivational elements condition) 	<ul style="list-style-type: none"> • $R_2^2 = 0.30$ • $R_3^2 = 0.90$ • 0.33 	<ul style="list-style-type: none"> • $R_2^2 = 0.70$ • $R_3^2 = 0.90$ • 0.33
Three-level blocked individual-level random assignment design	<ul style="list-style-type: none"> • Intraclass correlation at the class level (ρ_2) • Intraclass correlation at the school level (ρ_3) • Treatment effect heterogeneity at the class level (ω_2) • Treatment effect heterogeneity at the school level (ω_3) • Amount of variance at individual level explained by a pretest measure used as a covariate (R_1^2) • Amount of treatment effect variance among classes explained by a pretest measure used as a covariate e (R_2^2) • Amount of treatment effect variance among schools explained by a pretest measure used as a covariate (R_3^2) • Probability of assignment to Group A (vs. Group B) 	<ul style="list-style-type: none"> • $\rho_2 = 0.15$ • $\rho_3 = 0.25$ • $\omega_2 = 0.05$ • $\omega_3 = 0.05$ • $R_1^2 = 0.50$ • $R_2^2 = 0.25$ • $R_3^2 = 0.25$ • 0.50 	<ul style="list-style-type: none"> • $\rho_2 = 0.07$ • $\rho_3 = 0.02$ • $\omega_2 = 0.05$ • $\omega_3 = 0.02$ • $R_1^2 = 0.50$ • $R_2^2 = 0.25$ • $R_3^2 = 0.25$ • 0.50

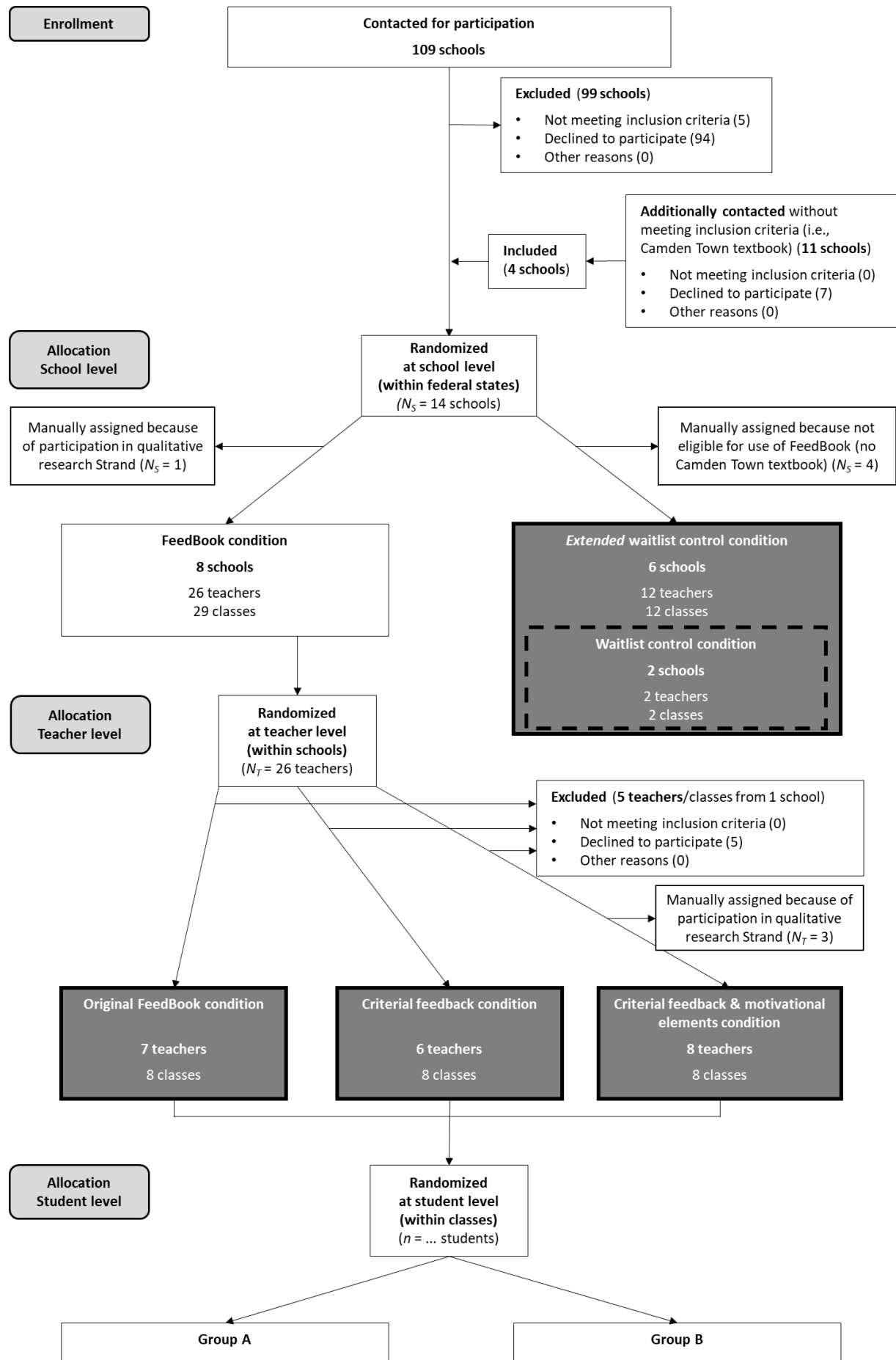
Recruitment. As an inclusion criterion, schools had to be using the English textbook Camden Town 3 from the Westermann publishing house to be able to participate in the study because the exercises included in the FeedBook and the teaching materials provided within the scope of the study were aligned with this textbook. To recruit the targeted number of schools and classes, we contacted 18 academic-track schools in Baden-Württemberg and 90 academic-track schools in North Rhine-Westphalia that, to our knowledge, used the Camden Town 3 textbook. We contacted both the headmasters or headmistresses as well as the English

teachers at these schools, as long as the necessary information about individual teachers were provided as public information. The headmasters/headmistresses and teachers received detailed information about the project, including information about the FeedBook itself and the concomitant study. If teachers decided to participate with their classes, information material were handed out to the individual students and their parents. There were no incentives for participation for either the schools, the teachers, or the students.

Given the inclusion criterion (i.e., the Camden Town textbook, which is only used in a subset of schools) and the COVID pandemic, which tremendously affected and still affects the schools and ultimately their willingness to participate, the participation rate was lower than anticipated. From the potential pool of schools working with the required textbook, a remainder of ten schools (four in Baden-Württemberg, five in North Rhine-Westphalia, and the one school from Hamburg which had been recruited beforehand) agreed to participate (see also Figure 4). However, one school in North Rhine-Westphalia with five classes declined to participate before the study started but after condition assignment (without awareness of their condition). Given the difference of seven schools with 22 classes as compared to our targeted number of schools and classes based on our power analyses, we additionally recruited schools that do not use the Camden Town textbook, but who were willing to participate in the study as a control group (*extended* waitlist control group). We gained an additional four schools and eleven classes in Baden-Württemberg to enlarge our control condition. Including these schools, **the final study sample comprises a total of $N_S = 13$ schools with $N_T = 33$ teachers and their $N_C = 36$ classes** (see also Figure 4).

Figure 4

Flow Diagram of the Study Design



Realized Assignment to Conditions. Because of the recruitment outcome, we adjusted our design slightly: Within Baden-Württemberg and North Rhine-Westphalia, schools using the Camden Town textbook were randomly assigned to the FeedBook condition (vs. the waitlist control condition) with a probability of 0.75 (as compared to 0.67 as initially intended) so that the absolute number of schools working with FeedBook was closer to the intended absolute number despite the low participation rates at the school level. Schools not working with the Camden Town textbook were manually assigned to an *extended* waitlist control condition.

The other probabilities for condition assignments were kept as initially intended. The participating school in Hamburg was manually assigned to the condition with the richest set of features (i.e., the *criterial feedback & motivational elements condition*) because the research focus in this school relies on a mixed methods analysis of the given elements. The final allocation to the intervention conditions is as follows:

- Seven schools with a total of 21 teachers together with their 24 classes were assigned to the FeedBook condition (including the Hamburg school), two schools with a total of two teachers together with their two classes to the waitlist control condition, and six schools with a total of twelve teachers together with their twelve classes to the *extended* waitlist control condition (i.e., supplemented with the four schools not working with the Camden Town textbook).
- Within the seven schools in the FeedBook condition, seven teachers together with their eight classes were assigned to the *original FeedBook condition*, six teachers together with their eight classes to the *criterial feedback condition*, and eight teachers together with their eight classes to the *criterial feedback & motivational elements condition* (including the three classes from the Hamburg

school). The unequal numbers of teachers that were assigned to the respective conditions resulted from the five classes in North Rhine-Westphalia that declined participation after condition assignment (without awareness of their condition).

- Within classes in the FeedBook condition, roughly 50% of the students were automatically assigned to Group A and to Group B, respectively, by the FeedBook system.

Updated Power Analyses for Final Study Design. We updated our power analyses based on the actual sample size but under identical assumptions as for the original power analyses. The results of the updated power analyses can be obtained at the following link: <https://doi.org/10.23668/psycharchives.5366>. They indicated that...

1. at the school level, we would be able to detect intervention effects of at least $\delta = 0.79$, 95% CI [0.19,1.40] (conservative) to $\delta = 0.48$, 95% CI [0.11,0.85] (optimistic) comparing the schools randomly assigned to the FeedBook condition or the waitlist control condition. When considering the *extended* control condition (i.e., supplemented with schools that are not working with the Camden Town textbook), the power analyses indicated that, at the school level, we would be able to detect intervention effects of at least $\delta = 0.44$, 95% CI [0.12,0.76] (conservative) to $\delta = 0.27$, 95% CI [0.07,0.46] (optimistic) comparing the FeedBook condition with the extended waitlist control condition,
2. at the class level, we would be able to detect intervention effects of at least $\delta = 0.50$, 95% CI [0.13,0.88] (conservative) to $\delta = 0.40$, 95% CI [0.10,0.70] (optimistic) comparing each two conditions at that level, and

3. at the individual level, we would be able to detect intervention effects of at least $\delta = 0.120$, 95% CI [0.04,0.20] (conservative) to $\delta = 0.106$, 95% CI [0.03,0.18] (optimistic) comparing Group A with Group B per Task Cycle.

Data Collection

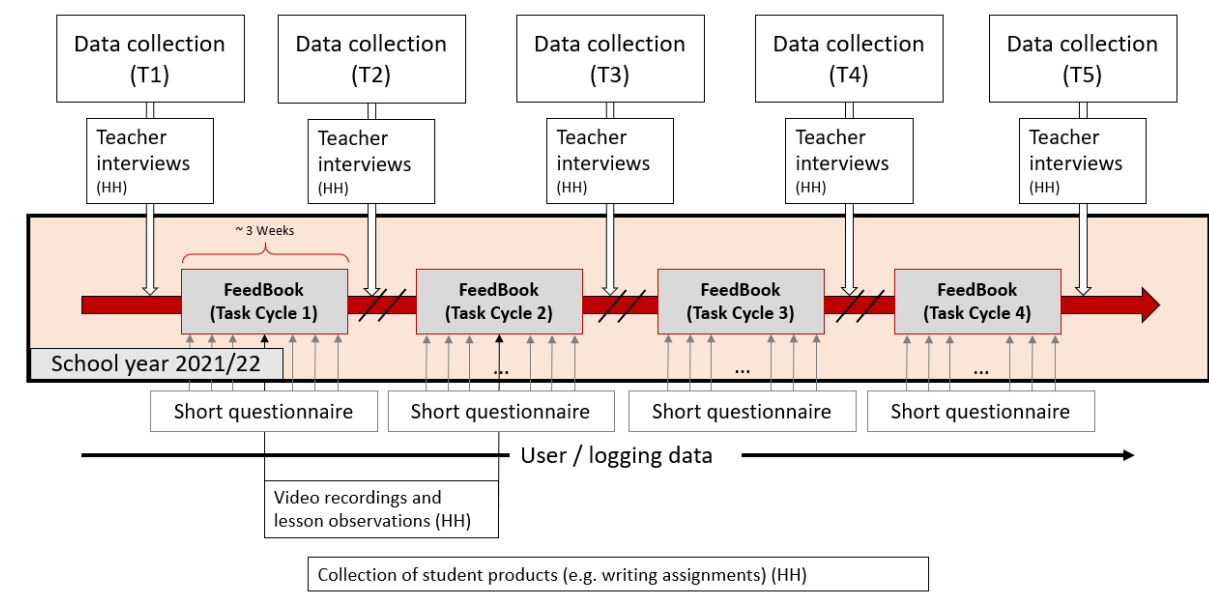
During the current study, the FeedBook will be used throughout an entire school year including four Task Cycles, each lasting approximately 3 weeks. To evaluate the effects of the intervention (e.g., the use of the FeedBook), data will be collected in four strands at multiple time points (see Figure 5 and Table 4).

- **Strand 1: Survey and test data.** Survey and test data were collected at the beginning of the school year (i.e., before the FeedBook was introduced; T1) and will be collected after every Task Cycle (i.e., after Task Cycle 1: T2; after Task Cycle 2: T3, after Task Cycle 3: T4; after Task Cycle 4: T5). Schools in the (extended) waitlist control condition only participate in T1 and T5, although the tests and questionnaires are identical to those in the FeedBook condition. The survey and test data comprise:
 - language proficiency tests (general language skills test, GCT; and language construct test tailored to the respective Task Cycle, LCT)
 - student questionnaires (SQs)
 - teacher questionnaires (TQs)
 - a working memory capacity test (WMCT) (only at T4)
- **Strand 2: Ambulatory assessment data.** Student short questionnaires will continuously be collected: Every time the students log out of the system, they are asked to complete a short questionnaire
- **Strand 3: Logfile data.** Student and teacher user data will continuously be collected, including user-system interactions such as click patterns within the FeedBook throughout the study period; this data enables comprehensive learning analytics

- **Strand 4: Qualitative data.** At the participating school in Hamburg, qualitative data was and will be collected in three classrooms. This data includes
 - 30-minute structured interviews with the three participating teachers at five time points (simultaneously to T1 to T5)
 - Lesson observations and audio/video recordings in the three classes in Hamburg for two 90-min lessons per class, one at the end of Task Cycle 1 and one at the end of Task Cycle 2
 - Evaluations of student products of all four Task Cycles (e.g., oral tasks and writing assignments)

Figure 5

Study Design



Note: T = timepoint, HH = Hamburg (one of the three federal states in which data collection takes place).

Measures

In the following, we will provide an overview of the primary outcomes of the study for Projects 1 to 4. The full set of items as well as further measures, including further survey

measures as well as measures from the qualitative data strand, can be derived from the Appendix.

English Proficiency. At T1 and T5, students will complete a C-test (Grotjahn, 2006), which is used to measure their general language skills (General Competence Test, GCT). The C-test consists of five authentic cloze texts that must be completed within 25 minutes in total. The GCT is supplemented by regular Language Construct Tests (LCT) before and after each Task Cycle (i.e., five times in total). The LCT focuses on those grammatical structures that are the target constructs of the respective Task Cycle. The LCTs were developed for the first FeedBook study (Meurers et al., 2019) and adapted for the current study. Additionally, information on students' English grades will be collected from school records.

Student-Reported Trait-Like Motivation. Students will be asked to answer parallel items on their trait-like motivation for English as a subject at T1 through T5. All items will be answered on a 4-point Likert-type scale from 1 = *not true at all* to 4 = *completely true*. Students will be asked to report their trait-like motivation with the introductory question "To what degree do these statements apply to you?" Based on Eccles et al.'s situated expectancy-value theory (1983, 2020), students' competence beliefs will be measured with a scale adapted from previous studies (e.g., Baumert et al., 1997; Ramm et al., 2006); and students' value beliefs (i.e., their intrinsic value, attainment value, utility value, and cost) will be measured with a scale developed by Gaspard et al. (2017), which was slightly adapted and shortened for the present study.

Student-Reported State Motivation. Students will additionally be asked to answer parallel items on their state motivation for English homework exercises every time they log out of the FeedBook. All items will be answered on a 4-point Likert-type scale from 1 = *not true at all* to 4 = *completely true*. Students will be asked to report their state motivation with the introductory question "To what degree do these statements apply to you and your exercise

completion in FeedBook today?” Items to assess students’ state competence beliefs and state value beliefs were adapted from the trait-like measures and shortened to fewer items.

Working Memory and Long-Term Memory. To assess the working memory and long-term memory capacity, we will select a set of measures from the following tasks: Visual Digit Span Backwards (adapted from Wechsler, 2008), OSPAN task (Unsworth et al., 2005), and N-Back task (adapted from Pelegrina et al., 2015, Schleepen & Jonkmann, 2009), for working memory; MLAT-V (Carroll & Sapon, 1959) and Explicit Visual Sequence Learning (adapted from Siegelman et al., 2017), for explicit long-term memory; Implicit Visual Sequence Learning (adapted from Siegelman et al., 2017) and Serial Reaction Time task (Lum et al., 2010), for implicit long-term memory.

The final set of measures has not been decided on at the time of the pre-registration and will be specified in an amendment to this pre-registration (amendments will be accessible via PsychArchives (<https://doi.org/10.23668/psycharchives.5366>) at “Related PsychArchives Objects”).

Table 4

Measures and Time Points

	T1	T2	T3	T4	T5
Measures	GCT LCT pre TC 1	LCT post TC 1 LCT pre TC 2	LCT post TC 2 LCT pre TC 3	LCT post TC 3 LCT pre TC 4	GCT LCT post TC 4
	SQ	SQ	SQ	SQ WMCT	SQ
	TQ	TQ	TQ	TQ	TQ
	Continuously: Logfile data from students and teachers, student short questionnaires				

Note. T = timepoint, GCT = General Construct Test, LCT = Language Construct Test, TC = Task Cycle, SQ = student questionnaire, TQ = teacher questionnaire, WMCT = working memory capacity test.

References

- Ahmadi, M. R. (2018). The use of technology in English language learning: A literature review. *International Journal of Research in English Education (IJREE)*, 3(2), 115–125.
- Baumert, J., Gruehn, S., Heyn, S., Köller, O., & Schnabel, K. U. (1997). *Bildungsverläufe und psychosoziale Entwicklung im Jugendalter (BIJU). Dokumentation – Band 1: Skalen Längsschnitt Welle 1–4 [Learning processes, educational careers, and psychosocial development in adolescence and young adulthood (BIJU). Documentation–Volume 1: Scales of Data collections 1–4]*. Max-Planck-Institute for Human Development.
- Baumert, J., Blum, W., Brunner, M., Dubberke, T., Jordan, A., Klusmann, U., Krauss, S., Kunter, M., Löwen, K., Neubrand, M., & Tsai, Y.-M. (2009). *Professionswissen von Lehrkräften, kognitiv aktivierender Mathematikunterricht und die Entwicklung von mathematischer Kompetenz (COACTIV): Dokumentation der Erhebungsinstrumente Materialien aus der Bildungsforschung Nr. 83 [Professional competence of teachers, cognitively activating instruction, and the development of mathematical competences (COACTIV): Scale documentation. Materials in educational research No. 83]*. Max-Planck-Institute for Human Development.
- Bulus, M., Dong, N., Kelcey, B., & Spybrook, J. (2021). *PowerUpR: Power Analysis Tools for Multilevel Randomized Experiments. R package version 1.1.0*. <https://CRAN.R-project.org/package=PowerUpR>
- Carroll, J. B., & Sapon, S. M. (1959). *Modern language aptitude test*. Psychological Corporation.
- Cloninger, C. R. (1987). A systematic method for clinical description and classification of personality variants: A proposal. *Archives of General Psychiatry*, 44(6), 573–588.

- Eccles, J. S., Adler, T. F., Futterman, R., Goff, S. B., Kaczala, C. M., Meece, J. L., & Midgley, C. (1983). Expectancies, values, and academic behaviors. In J. T. Spence (Ed.), *Achievement and achievement motives* (pp. 75–146). W. H. Freeman and Company.
- Eccles, J. S., & Wigfield, A. (2020). From expectancy-value theory to situated expectancy-value theory: A development, social cognitive, and sociocultural perspective on motivation. *Contemporary Educational Psychology*, 61, Article 101859.
<https://doi.org/10.1016/j.cedpsych.2020.101859>
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.
- Gaspard, H., Dicke, A.-L., Flunger, B., Brisson, B. M., Häfner, I., Nagengast, B., & Trautwein, U. (2015). Fostering adolescents' value beliefs for mathematics with a relevance intervention in the classroom. *Developmental Psychology*, 51(9), 1226–1240.
<https://doi.org/10.1037/dev0000028>
- Gaspard, H., Häfner, I., Parrisius, C., Trautwein, U., & Nagengast, B. (2017). Assessing task values in five subjects during secondary school: Measurement structure and mean level differences across grade level, gender, and academic subject. *Contemporary Educational Psychology*, 48, 67–84. <https://doi.org/10.1016/j.cedpsych.2016.09.003>
- Gaspard, H., Parrisius, C., Piesch, H., Kleinhansl, M., Wille, E., Nagengast, B., Trautwein, U., & Hulleman, C. S. (2021). The potential of relevance interventions for scaling up: A cluster-randomized trial testing the effectiveness of a relevance intervention in math classrooms. *Journal of Educational Psychology*, 113(8), 1507–1528.
<https://doi.org/http://dx.doi.org/10.1037/edu0000663>
- Grotjahn, R. (2006). Der C-Test: Theorie, Empirie, Anwendungen [The C-Test: Theory, empirical research, applications]. Peter Lang GmbH, Internationaler Verlag der Wissenschaften.

- Hallet, W., & Legutke, M. K. (2013). Task-approaches revisited. New orientations, new perspectives. *The European Journal of Applied Linguistics and TEFL*, 2(2), 139–160.
- Hertel, S., Hochweber, J., Mildner, D., Steinert, B., & Jude, N. (2014). *PISA 2009 Skalenhandbuch [PISA 2009 Scale documentation]*. Waxmann.
<https://doi.org/10.25656/01:9554>
- Jackel, A.-K., Wagner, W., Trautwein, U., & Göllner, R. (2021). “The teacher motivates us - or me?” - The role of the addressee in student ratings of teacher support. [Unpublished manuscript]. Hector Research Institute of Education Sciences and Psychology, University of Tübingen.
- Klimczak, M., Kampa, M., Bürgermeister, A., Harks, B., Rakoczy, K., Besser, M., Klieme, E., Blum, W., & Leiss, D. (2012). *Dokumentation der Befragungsinstrumente der Interventionsstudie im Projekt “Conditions and Consequences of Classroom Assessment” (Co2CA) [Documentation of the measures used in the intervention study “Conditions and Consequences of Classroom Assessment” (CO²CA)]*. DIPF.
<https://doi.org/10.25656/01:12119>
- Kulik, J. A., & Fletcher, J. D. (2016). Effectiveness of intelligent tutoring systems. *Review of Educational Research*, 86(1), 42–78. <https://doi.org/10.3102/0034654315581420>
- Kunter, M., Schümer, G., Artelt, C., Baumert, J., Klieme, E., Neubrand, M., Prenzel, M., Schiefele, U., Schneider, W., Stanat, P., Tillmann, K.-J., & Weiß, M. (2002). PISA 2000: Dokumentation der Erhebungsinstrumente. Materialien aus der Bildungsforschung Nr. 72 [*PISA 2000: Scale documentation. Materials in educational research No. 72*]. Max-Planck-Institute for Human Development.
- Kunter, M., Frenzel, A., Nagy, G., Baumert, J., & Pekrun, R. (2011). Teacher enthusiasm: Dimensionality and context specificity. *Contemporary Educational Psychology*, 36, 289–301. <https://doi.org/10.1016/j.cedpsych.2011.07.001>

- Kunter, M., Tsai, Y.-M., Klusmann, U., Brunner, M., Krauss, S., & Baumert, J. (2008). Students' and mathematics teachers' perceptions of teacher enthusiasm and instruction. *Learning and Instruction, 18*, 468–482.
<https://doi.org/10.1016/j.learninstruc.2008.06.008>
- Liebendörfer, M., Göller, R., Biehler, R., Hochmuth, R., Kortemeyer, J., Ostsieker, L., Rode, J., & Schaper, N. (2021). LimSt – Ein Fragebogen zur Erhebung von Lernstrategien im mathematikhaltigen Studium. *Journal für Mathematik-Didakt, 42*, 25–59.
<https://doi.org/10.1007/s13138-020-00167-y>
- Lum, J. A. G., Gelgic, C., & Conti-Ramsden, G. (2010). Procedural and declarative memory in children with and without specific language impairment. *International Journal of Language & Communication Disorders, 45*(1), 96-107.
<https://doi.org/10.3109/13682820902752285>
- Ma, W., Adesope, O. O., Nesbit, J. C., & Liu, Q. (2014). Intelligent tutoring systems and learning outcomes: A meta-analysis. *Journal of Educational Psychology, 106*(4), 901–918. <https://doi.org/10.1037/a0037123>
- McAuley, E., Duncan, T., & Tammen, V. V. (1987). Psychometric properties of the Intrinsic Motivation Inventory in a competitive sport setting: A confirmatory factor analysis. *Research Quarterly for Exercise and Sport, 60*, 48-58.
- Meurers, D., De Kuthy, K., Nuxoll, F., Rudzewitz, B., & Ziai, R. (2019). Scaling up intervention studies to investigate real-life foreign language learning in school. *Annual Review of Applied Linguistics, 39*, 161–188.
<https://doi.org/10.1017/S0267190519000126>
- Nassaji, H., & Kartchava, E. (2017). *Corrective feedback in second language teaching and learning: Research, theory, applications, implications*. Routledge.
<https://doi.org/10.29344/0717621x.19.1587>

- Pelegrina, S., Lechuga, M. T., García-Madruga, J. A., Elosúa, M. R., Macizo, P., Carreiras, M., Fuentes, L. J., & Bajo, M. T. (2015). Normative data on the n-back task for children and young adolescents. *Frontiers in Psychology*, 6, Article 1544. <https://doi.org/10.3389/fpsyg.2015.01544>
- R Core Team (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. <https://www.r-project.org>
- Ramm, G., Prenzel, M., Baumert, J., Blum, W., Lehmann, R., Leutner, R., Neubrand, M., Pekrun, R., Rolff, H.-G., Rost, J., & Schiefele, U. (2006). *Pisa 2003: Dokumentation der Erhebungsinstrumente [Pisa 2003: Scale documentation]*. Waxmann.
- Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K): Entwicklung und Validierung eines ökonomischen Inventars zur Erfassung der fünf Faktoren der Persönlichkeit [Short version of the Big Five Inventory (BFI-K): Development and validation of an economical inventory for the five personality factors]. *Diagnostica*, 51(4), 195–206. <https://doi.org/10.1026/0012-1924.51.4.195>
- Riazi, A. M. & Candlin, C. N. (2014). Mixed-methods research in language teaching and learning: Opportunities, issues and challenges. *Language Teaching*, 47(2), 135–173. <https://doi.org/10.1017/S0261444813000505>
- Rieger, S., Göllner, R., Spengler, M., Trautwein, U., Nagengast, B., & Roberts, B. W. (2017). Social cognitive constructs are just as stable as the Big Five between grades 5 and 8. *AERA Open*, 3(3), 1–9. <https://doi.org/10.1177/2332858417717691>
- Rudzewitz, B., Ziai, R., De Kuthy, K., & Meurers, D. (2017). Developing a web-based workbook for English supporting the interaction of students and teachers. *Proceedings of the Joint 6th Workshop on NLP for Computer Assisted Language Learning and 2nd Workshop on NLP for Research on Language Acquisition at NoDaLiDa 2017*, 134, 36–46.

- Rudzewitz, B., Ziai, R., De Kuthy, K., Möller, V., Nuxoll, F., & Meurers, D. (2018). Generating feedback for English foreign language exercises. *Proceedings of the Thirteenth Workshop on Innovative Use of NLP for Building Educational Applications*, 127–136. <https://doi.org/10.18653/v1/w18-0513>
- Schleepen, T. M. J., & Jonkman, L. M. (2009). The development of non-spatial working memory capacity during childhood and adolescence and the role of interference control: An n-back task study. *Developmental Neuropsychology*, 35(1), 37–56. <https://doi.org/10.1080/87565640903325733>
- Schmidt, D. A., Baran, E., Thompson, A. D., Mishra, P., Koehler, M. J., & Shin, T. S. (2009). Technological pedagogical content knowledge (TPACK): The development and validation of an assessment instrument for preservice teachers. *Journal of Research on Technology in Education*, 42(2), 123–149. <https://doi.org/10.1080/15391523.2009.10782544>
- Siegelman, N., Bogaerts, L., & Frost, R. (2017). Measuring individual differences in statistical learning: current pitfalls and possible solutions. *Behavioral Research Methods*, 49, 418–432.
- Stallasch, S. E., Lüdtke, O., Artelt, C., & Brunner, M. (2021). Multilevel design parameters to plan cluster-randomized intervention studies on student achievement in elementary and secondary school. *Journal of Research on Educational Effectiveness*, 14(1), 172–206. <https://doi.org/10.1080/19345747.2020.1823539>
- Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology*, 106(2), 331–347. <https://doi.org/10.1037/A0034752>
- Steinberg, M. L., Krejci, J. A., Collett, K., Brandon, T. H., Ziedonis, D. M., & Chen, K. (2007). Relationship between self-reported task persistence and history of quitting

- smoking, plans for quitting smoking, and current smoking status in adolescents. *Addictive Behaviors*, 32(7), 1451-1460. <https://doi.org/10.1016/j.addbeh.2006.10.008>
- Trautwein, U., Lüdtke, O., Roberts, B. W., Schnyder, I., & Niggli, A. (2009). Different forces, same consequence: Conscientiousness and competence beliefs are independent predictors of academic effort and achievement. *Journal of Personality and Social Psychology*, 97(6), 1115–1128. <https://doi.org/10.1037/a0017048>
- Tseng, W. T., Cheng, H. F., & Gao, X. (2020). Validating a motivational self-guide scale for language learners. *Sustainability*, 12, 6468–6468. <https://doi.org/10.3390/su12166468>
- Unsworth, N., Heitz, R. P., Schrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, 37(3), 498-505. <https://doi.org/10.3758/BF03192720>
- Voss, T., Kunter, M., Seiz, J., Hoehne, V., & Baumert, J. (2014). Die Bedeutung des pädagogisch-psychologischen Wissens von angehenden Lehrkräften für die Unterrichtsqualität [The importance of pedagogical-psychological knowledge of preservice teachers for the quality of teaching]. *Zeitschrift für Pädagogik*, 60(2), 184-201. <https://doi.org/25656/01:14653>
- Wechsler, D. (2009). *Wechsler Adult Intelligence Scale*. Fourth Edition. Pearson.
- Williams, G. C., & Deci, E. L. (1996). Internalization of biopsychosocial values by medical students: A test of self-determined theory. *Journal of Personality and Social Psychology*, 70, 767–779.
- Würffel, N. (2017). Individualisiertes, differenziertes und barrierefreies Fremdsprachenlernen - Digitalisierung als Chance eines inklusiven Fremdsprachenunterrichts? [Individualized, differentiated, and barrier-free foreign language learning – Digitalization as a chance for inclusive foreign language teaching?] In E. Burwitz-Melzer, F. G. Königs, C. Riemer, & L. Schmelter (Eds.), *Inklusion, Diversität und das Lehren und Lernen fremder Sprachen*.

*Arbeitspapiere der 37. Frühjahrskonferenz zur Erforschung des
Fremdsprachenunterrichts* (pp. 351–362). Narr Francke Attempto Verlag.

Appendix

The Appendix contains the following information:

Appendix A: Overview of measures included for Strand 1 (survey and test data)

Appendix B: Overview of measures included for Strand 2 (ambulatory assessment data)

Appendix C: Interview guides included for Strand 4 (qualitative data)

Measures are currently provided in German only. English translations (or, if applicable, the English originals) will be published in an amendment to this pre-registration. Amendments will be accessible via PsychArchives (<https://doi.org/10.23668/psycharchives.5366>) at “Related PsychArchives Objects.”

Appendix A

Overview of Measures Included for Strand 1 (Survey and Test Data)

Table A.1

Student Questionnaires (T1 to T5)

Scale/construct	Item (German)	Source
Trait-like motivation – Self-concept	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	
	Für Englisch habe ich einfach keine Begabung.	Gaspard et al. (2017)
	Englisch liegt mir nicht besonders.	Gaspard et al. (2017)
	Ich bin gut in Englisch.	Gaspard et al. (2017)
	Englisch fällt mir leicht.	Gaspard et al. (2017)
	Ich bin mit meiner Leistung in Englisch zufrieden.	Adapted from the Intrinsic Motivation Inventory (IMI), McAuley et al. (1989)
	Ich fühle mich im Umgang mit Englisch meist recht sicher.	Adapted from the Intrinsic Motivation Inventory (IMI), McAuley et al. (1989)
Trait-like motivation – Intrinsic value	Ich habe mich durch die Aufgaben im FeedBook gut auf die letzte Target Task vorbereitet gefühlt.	Self-developed
	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	
	Englisch macht mir Spaß.	Gaspard et al. (2017)

	Ich mache Englisch gerne.	Gaspard et al. (2017)
	Englisch mag ich einfach.	Gaspard et al. (2017)
Trait-like motivation – Attainment value	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	
	Gut in Englisch zu sein, bedeutet mir viel.	Gaspard et al. (2017)
	Mir ist eine gute Englischnote sehr wichtig.	Gaspard et al. (2017)
	Mir liegt viel daran den Stoff in Englisch zu behalten.	Gaspard et al. (2017)
	Das Fach Englisch hat für mich eine große Bedeutung.	Adapted from Gaspard et al. (2017)
Trait-like motivation – Utility value	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	
	Das Wissen aus dem Fach Englisch bringt mir im täglichen Leben viele Vorteile.	Gaspard et al. (2017)
	Englischkenntnisse kann man im Alltag und in der Freizeit gut gebrauchen	Gaspard et al. (2017)
	Für meine berufliche Zukunft wird es sich auszahlen, gut in Englisch zu sein.	Gaspard et al. (2017)
	In Englisch gut zu sein lohnt sich, weil man es in der Schule einfach braucht.	Adapted from Gaspard et al. (2017)
	Die Aufgaben im FeedBook waren nützlich als Vorbereitung auf die Target Task.	Self-developed
Trait-like motivation – Cost	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	
	Mich mit Englisch zu beschäftigen ist anstrengend.	Gaspard et al. (2017)
	Die Beschäftigung mit Englisch kostet mich eine Menge Energie.	Gaspard et al. (2017)
	Wenn ich mich mit Englisch beschäftige, fühle ich mich oft hoffnungslos.	Self-developed

	Ich bin enttäuscht, wenn ich in Englisch eine Aufgabe nicht richtig lösen kann.	Self-developed
	Mich mit Englisch auseinanderzusetzen ist langweilig.	Self-developed
	Die Beschäftigung mit Englisch frustriert mich oftmals.	Self-developed
	Um in Englisch gut zu sein, müsste ich viel Freizeit opfern.	Gaspard et al. (2017)
Ideal L2Self	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	
	Ich wäre gerne irgendwann in der Lage, mich mit internationalen Freunden auf Englisch zu unterhalten.	Tseng et al. (2020)
	Ich wäre gerne irgendwann in der Lage, meine Gedanken in sozialen Medien auch auf Englisch zu teilen.	Adapted from Tseng et al. (2020)
	Ich wäre gerne irgendwann in der Lage, englische Bücher und Nachrichten zu lesen	Tseng et al. (2020)
	Ich wäre gerne irgendwann in der Lage, englische Filme und Videos ohne Untertitel zu schauen.	Adapted from Tseng et al. (2020)
Effort	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	
	Ich strenge mich in Englisch an.	Adapted from Trautwein et al. (2009)
	Ich tue mein Bestes bei den Englischaufgaben.	Adapted from Trautwein et al. (2009)
	Ich bearbeite alle Aufgaben in Englisch ernsthaft.	Adapted from Trautwein et al. (2009)
Resistance to frustration	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	
	Auch wenn ich frustriert bin, lerne ich trotzdem weiter.	Liebendörfer et al. (2021)
	Auch wenn ich beim Lernen überhaupt nicht vorankomme, versuche ich es immer wieder, bis es klappt.	Liebendörfer et al. (2021)

	Ich denke über Aufgaben immer wieder nach, auch wenn ich gerade nicht weiterkomme.	Liebendörfer et al. (2021)
	Wenn eine Übungsaufgabe schwer ist, dann versuche ich nicht lange sie zu lösen.	Liebendörfer et al. (2021)
	Bei Übungsaufgaben versuche ich, alles hinzubekommen.	Liebendörfer et al. (2021)
	Wenn ich eine Aufgabe an einem Tag nicht lösen kann, versuche ich es am nächsten Tag weiter.	Liebendörfer et al. (2021)
Engagement	<i>Wie oft kommt es in einer typischen Englischstunde vor, dass du ...</i>	
	...dich meldest, um aktiv am Unterricht teilzunehmen?	Gaspard et al. (2017)
	...dich für die Aufgaben begeistern kannst?	Gaspard et al. (2017)
	...mit deinen Mitschüler:innen über die Unterrichtsaufgaben diskutierst?	Gaspard et al. (2017)
	...Interesse an den Unterrichtsinhalten zeigst?	Gaspard et al. (2017)
	...sofort mit der Bearbeitung von Aufgaben beginnst?	Gaspard et al. (2017)
Task Persistence	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	
	Ich versuche etwas immer und immer wieder, auch wenn ich beim ersten Mal keinen Erfolg hatte.	Steinberg et al. (2007), originally adapted from TPQ by Cloninger (1987)
	Ich arbeite häufig an Dingen weiter, wenn Andere schon aufgegeben haben.	Steinberg et al., (2007), originally adapted from TPQ by Cloninger (1987)
Personality – Conscientiousness	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	
	Ich bin tüchtig und arbeite flott.	BFI-K, Rammstedt & John (2005)
	Ich mache Pläne und führe sie auch durch.	BFI-K, Rammstedt & John (2005)

Personality – Extraversion	Ich erledige meine Aufgaben gründlich.	BFI-K, Rammstedt & John (2005)
	Ich bin bequem, neige zu Faulheit.	BFI-K, Rammstedt & John (2005)
	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	
	Ich bin eher zurückhaltend, reserviert.	BFI-K, Rammstedt & John (2005)
	Ich bin begeisterungsfähig und kann andere leicht mitreißen.	BFI-K, Rammstedt & John (2005)
	Ich bin eher der/die "Stille", wortkarg.	BFI-K, Rammstedt & John (2005)
Personality – Agreeableness	Ich gehe aus mir heraus, bin gesellig.	BFI-K, Rammstedt & John (2005)
	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	
	Ich neige dazu, andere zu kritisieren.	BFI-K, Rammstedt & John (2005)
	Ich schenke anderen leicht Vertrauen, glaube an das Gute im Menschen.	BFI-K, Rammstedt & John (2005)
	Ich kann mich kalt und distanziert verhalten.	BFI-K, Rammstedt & John (2005)
	Ich kann mich schroff und abweisend anderen gegenüber verhalten.	BFI-K, Rammstedt & John (2005)
Personality – Neuroticism	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	

	Ich werde leicht deprimiert, niedergeschlagen.	BFI-K, Rammstedt & John (2005)
	Ich bin entspannt, lasse mich durch Stress nicht aus der Ruhe bringen.	BFI-K, Rammstedt & John (2005)
	Ich mache mir viele Sorgen.	BFI-K, Rammstedt & John (2005)
	Ich werde leicht nervös und unsicher.	BFI-K, Rammstedt & John (2005)
Personality - Openness	<i>Wie sehr treffen diese Aussagen auf dich zu?</i>	
	Ich bin vielseitig interessiert.	BFI-K, Rammstedt & John (2005)
	Ich bin tiefsinnig, denke gerne über Sachen nach.	BFI-K, Rammstedt & John (2005)
	Ich habe eine aktive Vorstellungskraft, bin phantasievoll.	BFI-K, Rammstedt & John (2005)
	Ich schätze künstlerische und ästhetische Eindrücke.	BFI-K, Rammstedt & John (2005)
	Ich habe nur wenig künstlerisches Interesse.	BFI-K, Rammstedt & John (2005)
Instructional quality – Cognitive activation	<i>Wie sehr treffen diese Aussagen auf den Englisch-Unterricht bzw. deine Englischlehrkraft zu?</i>	
	Unserer Englischlehrkraft ist es wichtig, dass wir regelmäßig üben.	Jaekel et al. (2021), originally adapted from Baumert et al. (1997) and Ramm et al. (2006)

Instructional quality – Classroom management – Monitoring	Die Übungsphasen im Unterricht sind für uns abwechslungsreich.	Jaekel et al. (2021), originally adapted from Baumert et al. (1997) and Ramm et al. (2006)
	Unsere Lehrkraft legt Wert darauf, dass wir bestimmte Dinge auswendig können.	Jaekel et al. (2021)
	Wir üben regelmäßig, um Aufgaben immer schneller und sicherer bearbeiten zu können.	Jaekel et al. (2021), originally adapted from Baumert et al. (1997) and Ramm et al. (2006)
	<i>Wie sehr treffen diese Aussagen auf den Englisch-Unterricht bzw. deine Englischlehrkraft zu?</i>	
	Unsere Englischlehrkraft weiß im Unterricht genau, was in der Klasse vor sich geht.	Jaekel et al. (2021), originally adapted from Baumert et al. (1997)
	Unsere Englischlehrkraft greift ein, bevor Unruhe und Störungen überhaupt erst entstehen.	Jaekel et al. (2021), originally adapted from Baumert et al. (1997)
	Unsere Englischlehrkraft achtet sehr darauf, dass wir aufpassen.	Jaekel et al. (2021), originally adapted from Baumert et al. (1997)
	Unsere Englischlehrkraft merkt sofort, wenn Schüler:innen beginnen, etwas anderes zu treiben.	Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

**Instructional quality –
Classroom management
– Disturbances**

Wie sehr treffen diese Aussagen auf den Englisch-Unterricht bzw. deine Englischlehrkraft zu?

Im Unterricht ist es manchmal laut und alles geht durcheinander.

Jaekel et al. (2021), originally adapted from Kunter et al. (2002)

Der Unterricht wird oft gestört.

Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

Manchmal muss unsere Englischlehrkraft zu Beginn der Stunde lange warten, bis Ruhe eintritt.

Jaekel et al. (2021), originally adapted from Kunter et al. (2002)

Im Unterricht wird fortwährend laut gequatscht.

Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

Im Unterricht wird manchmal auch Zeit vertrödelte.

Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

Im Unterricht dauert es meistens sehr lange, bis alle zur Arbeit bereit sind.

Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

**Instructional quality –
Support – Handling
mistakes**

Wie sehr treffen diese Aussagen auf den Englisch-Unterricht bzw. deine Englischlehrkraft zu?

Unsere Englischlehrkraft ist geduldig, wenn jemand im Unterricht einen Fehler macht

Ramm et al. (2006)

Bei unserer Englischlehrkraft ist Fehlermachen nichts Schlimmes.

Ramm et al. (2006)

Instructional quality – Support – Internal differentiation	Unsere Englischlehrkraft achtet darauf, dass in unserer Klasse niemand ausgelacht wird, der einen Fehler macht.	Ramm et al. (2006)
	<i>Wie sehr treffen diese Aussagen auf den Englisch-Unterricht bzw. deine Englischlehrkraft zu?</i>	
	Unsere Englischlehrkraft gibt den einzelnen Schüler:innen oft verschiedene Aufgaben.	Adapted from Baumert et al. (1997)
	Bei unserer Englischlehrkraft können schnellere Schüler:innen schon zum Nächsten übergehen.	Baumert et al. (1997)
Instructional quality – Support – Enthusiasm	Unsere Englischlehrkraft stellt unterschiedlich schwere Fragen, je nachdem, wie gut eine Schülerin oder ein Schüler ist.	Baumert et al. (1997)
	Unsere Englischlehrkraft verlangt von den guten Schüler:innen deutlich mehr.	Baumert et al. (1997)
	<i>Wie sehr treffen diese Aussagen auf den Englisch-Unterricht bzw. deine Englischlehrkraft zu?</i>	
	Unsere Englischlehrkraft scheint großen Spaß am Unterrichten zu haben.	Kunter et al. (2008)
	Unsere Englischlehrkraft unterrichtet mit Begeisterung.	Kunter et al. (2008)
Instructional quality – Support – Feedback	Unsere Englischlehrkraft ist selbst vom Fach Englisch begeistert.	Adapted from Kunter et al. (2008)
	<i>Wie sehr treffen diese Aussagen auf den Englisch-Unterricht bzw. deine Englischlehrkraft zu?</i>	
	Von unserer Englischlehrkraft erfahren wir immer wieder, wo wir mit unseren Leistungen stehen.	Jaekel et al (2021), originally adapted from Klimczak et al. (2012)
	Unsere Englischlehrkraft gibt uns regelmäßig Rückmeldung, was wir schon können.	Jaekel et al (2021), originally adapted from Klimczak et al. (2012)

	Unsere Englischlehrkraft gibt uns regelmäßig Rückmeldung, was wir noch nicht so gut können.	Jaekel et al (2021), originally adapted from Klimczak et al. (2012)
	Unsere Englischlehrkraft zeigt uns immer wieder, wie wir uns verbessern können.	Jaekel et al (2021), originally adapted from Klimczak et al. (2012)
Instructional quality – Support – Personal support	<i>Wie sehr treffen diese Aussagen auf den Englisch-Unterricht bzw. deine Englischlehrkraft zu?</i>	
	Ich habe das Gefühl, dass meine Englischlehrkraft mir Wahlmöglichkeiten und Optionen bietet.	Williams & Deci (1996)
	Ich fühle mich von meiner Englischlehrkraft verstanden.	Williams & Deci (1996)
	Meine Englischlehrkraft vermittelt mir Vertrauen in meine Englisch-Fähigkeiten.	Adapted from Williams & Deci (1996)
	Meine Englischlehrkraft ermutigt mich, Fragen zu stellen.	Adapted from Williams & Deci (1996)
	Meine Englischlehrkraft hört sich an, wie ich Sachen gerne machen würde.	Adapted from Williams & Deci (1996)
Teacher familiarity	Wie gut kennst du deine Englischlehrkraft bereits?	Self-developed
English homework	Gibt es an deiner Schule Hausaufgaben, die ihr auch zuhause bearbeiten sollt oder arbeitet ihr stattdessen ausschließlich in Lernbüros/Lernwerkstätten bzw. während individuellen Selbstlernzeiten in der Schule?	Self-developed
	Wie oft bekommt ihr Hausaufgaben in Englisch?	Self-developed
	Wie häufig habt ihr Aufgaben im FeedBook in Task Cycle [1/2/3/4] während des Englischunterrichts bearbeitet?	Self-developed

	Wie oft habt ihr in Task Cycle [1/2/3/4] Aufgaben im FeedBook bekommen (z.B. als Hausaufgaben oder in individuellen Lernzeiten)?	Self-developed
	Wenn ihr Hausaufgaben in Englisch aufbekommt, wie viele Minuten brauchst du durchschnittlich dafür, diese Englischhausaufgaben zu erledigen?	Meurers et al. (2019)
	Wenn ihr Aufgaben im FeedBook bearbeiten solltet (z.B. als Hausaufgaben oder in individuellen Lernzeiten), wie viele Minuten hast du in der Regel dann daran gearbeitet?	Meurers et al. (2019)
Homework engagement	<i>Wie verhältst du dich bei der Erledigung deiner Englischhausaufgaben auf Papier?</i>	
	Ich tue mein Bestes bei den Hausaufgaben in Englisch.	Meurers et al. (2019)
	Ich bearbeite in letzter Zeit die Englischhausaufgaben so gut ich kann.	Meurers et al. (2019)
	Ich versuche immer meine Englischhausaufgaben vollständig zu erledigen.	Meurers et al. (2019)
	Ich schreibe die Hausaufgaben in Englisch oft von anderen in der Schule ab.	Meurers et al. (2019)
	Bei den Hausaufgaben ist mir oft nicht klar, was ich machen soll.	Meurers et al. (2019)
	Die Hausaufgaben sind klar formuliert.	Meurers et al. (2019)
	Ich weiß oft nicht, was ich für Hausaufgaben aufhabe.	Meurers et al. (2019)
Homework engagement 2	<i>Wie kommst Du mit den Englischhausaufgaben zurecht?</i>	
	sind oft zu schwer	Meurers et al. (2019)
	helfen mir, Englisch zu lernen.	Meurers et al. (2019)
	sind wichtig, um im Unterricht mitmachen zu können.	Meurers et al. (2019)
	sind oft sehr leicht	Meurers et al. (2019)
	mache ich oft zusammen mit meinen Eltern.	Meurers et al. (2019)
	kann ich meist nur mit meinen Eltern lösen	Meurers et al. (2019)
FeedBook engagement	<i>Wie verhältst du dich bei der Erledigung der Aufgaben im FeedBook?</i>	

	Ich tue mein Bestes bei den Aufgaben im FeedBook.	Meurers et al. (2019)
	Ich bearbeite in letzter Zeit meine Englischaufgaben im Feedbook so gut ich kann.	Meurers et al. (2019)
	Ich versuche immer meine Englischaufgaben im Feedbook vollständig zu erledigen.	Meurers et al. (2019)
	Ich schreibe die Lösungen für die Aufgaben im FeedBook oft von anderen in der Schule ab.	Meurers et al. (2019)
	Bei den Aufgaben im FeedBook ist mir oft nicht klar, was ich machen soll.	Meurers et al. (2019)
	Die Aufgaben im FeedBook sind klar formuliert.	Meurers et al. (2019)
	Ich weiß oft nicht, was ich für Aufgaben im FeedBook als Hausaufgaben aufhabe.	Meurers et al. (2019)
FeedBook engagement 2	<i>Wie kommst Du mit den Englischaufgaben im FeedBook zurecht?</i>	
	sind oft zu schwer.	Meurers et al. (2019)
	helfen mir, Englisch zu lernen.	Meurers et al. (2019)
	sind wichtig, um im Unterricht mitmachen zu können.	Meurers et al. (2019)
	sind oft sehr leicht.	Meurers et al. (2019)
	mache ich oft zusammen mit meinen Eltern.	Meurers et al. (2019)
	kann ich meist nur mit meinen Eltern lösen.	Meurers et al. (2019)
Perception of FeedBook – General	<i>Wie sind Deine Erfahrungen mit dem FeedBook im Allgemeinen?</i>	
	Die Rückmeldungen im FeedBook helfen mir die Aufgaben richtig zu machen.	Meurers et al. (2019)
	FeedBook Aufgaben machen mehr Spaß als im Papierarbeitsheft	Meurers et al. (2019)
	FeedBook Aufgaben machen mehr Spaß als Aufgaben in der Klasse.	Meurers et al. (2019)
	Die Aufgaben im FeedBook funktionieren gut.	Meurers et al. (2019)
	Bei den Aufgaben im FeedBook ist klar, was ich machen muss.	Meurers et al. (2019)

Perception of FeedBook – Dashboard <i>[only in respective intervention condition]</i>	Bei den Aufgaben im Papierarbeitsheft ist klar, was ich machen muss.	Meurers et al. (2019)
	Wenn ich mit dem FeedBook arbeite, lerne ich was.	Meurers et al. (2019)
	Das FeedBook ist leicht zu bedienen.	Meurers et al. (2019)
	Mit den Aufgaben im FeedBook lerne ich mehr in Englisch.	Meurers et al. (2019)
	Das Schüler:innen-Dashboard mit der Aufgabenübersicht im FeedBook finde ich hilfreich.	Self-developed
	Die Aufgabenübersicht hilft mir, meine Fähigkeiten einzuschätzen.	Self-developed
	Die Aufgabenübersicht hilft mir, meinen Lernfortschritt einzuschätzen.	Self-developed
	Ich arbeite gerne in der Aufgabenübersicht.	Self-developed
	Die Aufgabenübersicht ist leicht zu verstehen.	Self-developed
	Die Aufgabenübersicht im FeedBook finde ich hilfreich.	Self-developed
	Die Aufgabenübersicht ist einfach zu bedienen.	Self-developed
	Ich verstehe, was die Balken, die nach jeder Aufgabe und in der Aufgabenübersicht angezeigt werden, bedeuten.	Self-developed
	Das Ausklappen der Balken bzw. Aufgaben in der Aufgabenansicht finde ich praktisch.	Self-developed
	Mir ist es wichtig, Pokale zu bekommen.	Self-developed
	Die "More Practice"-Schaltfläche bringt mich zu...	Self-developed
Perception of FeedBook – Agent	Die "Challenge Me"-Schaltfläche bringt mich zu...	Self-developed
	Den Pokal erhalte ich, wenn...	Self-developed
	<i>Nun geht es um den Lernbuddy, der nach jeder Aufgaben-Abgabe auftaucht (also der Big Ben, die Maus oder der Yeti). Wie findest du diesen Lernbuddy?</i>	

*[only in respective
intervention condition]*

Der Lernbuddy nervt mich.	Self-developed
Ich bin mit meinem Lernbuddy zufrieden.	Self-developed
Ich behalte immer denselben Lernbuddy und stelle keinen anderen ein.	Self-developed
Ich hätte lieber gar keinen Lernbuddy.	Self-developed
Seit ich mit dem Lernbuddy arbeite, strenge ich mich mehr an.	Self-developed
Ich fand den Lernbuddy cool.	Self-developed
Durch den Lernbuddy macht mir das Arbeiten mit dem FeedBook mehr Spaß.	Self-developed
<i>Die Sätze des Lernbuddys...</i>	Self-developed

**Perception of FeedBook
– Agent Feedback**
*[only in respective
intervention condition]*

...fand ich überflüssig.	Self-developed
...waren zu negativ.	Self-developed
... haben mich frustriert.	Self-developed
...waren langweilig.	Self-developed
...habe ich gelesen.	Self-developed
...waren ermutigend.	Self-developed
...haben mich motiviert, weiterzuarbeiten.	Self-developed
...waren zu positiv.	Self-developed
...waren hilfreich, um meine Leistung einzuschätzen.	Self-developed
...habe ich verstanden.	Self-developed

	...haben dazu geführt, dass ich es beim nächsten Mal noch besser machen wollte.	Self-developed
	...haben mich stolz gemacht.	Self-developed
Parental interest	Meine Eltern interessiert es, ob ich die Hausaufgaben gemacht habe.	Self-developed
Internet/media access	Hast Du zu Hause Zugang zum Internet? Pro Tag...	Adapted from Kunter et al. (2002)
	Hast Du in der Schule Zugang zum Internet? Pro Tag...	Adapted from Kunter et al. (2002)
	Wie kommst Du normalerweise ins Internet?	Adapted from Kunter et al. (2002)
	Nutzt Du Lernprogramme auf dem Handy, Computer oder Tablet?	Adapted from Kunter et al. (2002)
Learning apps computer	Welche Lernprogramme nutzt Du auf dem Computer/Tablet?	Adapted from Meurers et al. (2019)
	Wie häufig verwendest du diese? (im Durchschnitt)	Self-developed
	Wie sinnvoll findest du diese? (im Durchschnitt)	Self-developed
Learning apps smartphone	Welche Lernprogramme nutzt Du auf dem Handy?	Meurers et al. (2019)
	Wie häufig verwendest du diese? (im Durchschnitt)	Self-developed
	Wie sinnvoll findest du diese? (im Durchschnitt)	Self-developed
Vocab learning	<i>Wie lernst Du Vokabeln?</i>	
	mit dem Computer, Tablet oder Handy	Adapted from Meurers et al. (2019)
	mit den Eltern, Geschwistern	Meurers et al. (2019)
	mit Karteikarten	Meurers et al. (2019)

	mit dem Buch	Meurers et al. (2019)
	mit dem Vokabelheft	Meurers et al. (2019)
Computer proficiency	<i>Wie gerne benutzt du Computer (oder Tablet-Computer)?</i>	
	Mit dem Computer kenne ich mich aus.	Meurers et al. (2019)
	Ich arbeite gerne mit Computern.	Meurers et al. (2019)
	Computer sind frustrierend.	Meurers et al. (2019)
	Es fällt mir schwer, am Computer längere Wörter und Sätze einzugeben.	Meurers et al. (2019)
	Ich tippe schnell und gut.	Meurers et al. (2019)
Gender	Bist du weiblich, männlich oder divers?	—
Age	In welchem Jahr bist du geboren?	—
	In welchem Monat bist du geboren?	—
Repetition of school year(s)	Hast du schon mal eine Klasse wiederholt?	—
	Wenn ja, kreuze alle Klassen an, die du schon wiederholt hast.	—
Second language at school	Welche Fremdsprachen lernst Du in der Schule?	—
Private tutoring lessons	Nimmst du zurzeit Nachhilfe in Englisch?	—
Migration background	<i>In diesem Teil des Fragebogens geht es um Angaben zu deiner Familie. Wähle bei jeder Frage die Antwort aus, die am besten zutrifft. Fragen wir nach deiner Mutter oder deinem Vater, sind damit diejenigen Personen gemeint, mit denen du die meiste Zeit in einem Haushalt wohnst. Dies kann auch die Partnerin oder der Partner eines Elternteils sein.</i>	
	In welchem Land wurdest du geboren?	—
	In welchem Land wurde deine Mutter geboren?	—

Language at home	In welchem Land wurde dein Vater geboren?	—
	<i>Wenn du zu Hause mit deinen Eltern sprichst, welche Sprache benutzt ihr dabei meistens?</i>	
	Du	—
	Mutter	—
Parents' educational background	Vater	—
	<i>Was ist der höchste Schulabschluss deiner Eltern?</i>	
	Mutter	—
Parents' job	Vater	—
	Was macht deine Mutter zurzeit?	Hertel et al. (2014)
	Welchen Beruf übt deine Mutter aus? (z.B. Lehrerin, Küchenhilfe, Verkaufsleiterin)	Hertel et al. (2014)
	Was macht deine Mutter in diesem Beruf? (z. B. in einer Grundschule unterrichten, in einem Restaurant der Köchin/dem Koch helfen, die Gerichte zuzubereiten, ein Verkaufsteam leiten)	Hertel et al. (2014)
	Was macht dein Vater zurzeit?	Hertel et al. (2014)
	Welchen Beruf übt dein Vater aus? (z.B. Lehrer, Küchenhilfe, Verkaufsleiter)	Hertel et al. (2014)
	Was macht dein Vater in diesem Beruf? (z. B. in einer Grundschule unterrichten, in einem Restaurant der Köchin/dem Koch helfen, die Gerichte zuzubereiten, ein Verkaufsteam leiten)	Hertel et al. (2014)
Family English proficiency	<i>Wie gut kann deine Familie Englisch?</i>	
	Mutter	Meurers et al. (2019)
	Vater	Meurers et al. (2019)

	Geschwister	Meurers et al. (2019)
Parental help	<i>Wie helfen Dir Deine Eltern zu Hause?</i>	
	finden es wichtig, dass ich in der Schule gut bin.	Meurers et al. (2019)
	finden es wichtig, dass ich Englisch lerne.	Meurers et al. (2019)
	wollen, dass ich mich in der Schule anstrenge	Meurers et al. (2019)
	helfen mir beim Üben auf Klassenarbeiten in Englisch.	Meurers et al. (2019)
	wissen, wann die nächste Klassenarbeit ist.	Meurers et al. (2019)
	fragen mich Englischvokabeln ab.	Meurers et al. (2019)
Experience abroad	Warst Du schon mal länger in einem englischsprachigen Land (England, USA, Australien, ...)?	Meurers et al. (2019)
	Wie lange?	Meurers et al. (2019)
Use of English in everyday life	<i>Wann verwendest Du Englisch?</i>	
	... um englische Filme oder Serien in Originalfassung zu sehen?	Meurers et al. (2019)
	... englische Bücher zu lesen?	Meurers et al. (2019)
	... englische Videospiele zu spielen?	Meurers et al. (2019)
	... um (Lern-)Apps zu nutzen.	Meurers et al. (2019)
Comment	Falls du mit einer Frage oder Aufgabe Probleme hattest (z.B. falls etwas nicht richtig funktioniert hat oder du etwas nicht verstanden hast) oder du uns sonst noch etwas Wichtiges mitteilen möchtest, nutze das Feld unten.	—
	Möchtest Du uns noch etwas [zur Aufgabenübersicht/zum Lernbuddy] oder zum FeedBook allgemein mitteilen? <i>[only used in the respective intervention condition]</i>	—

Note. Not all items were used at all time points.

Table A.2*Teacher Questionnaires (T1 to T5)*

Scale/construct	Item (German)	Source
Motivation for teaching – Self-concept	<i>In den [ersten Englischstunden dieses Schuljahres/letzten Englischstunden] in dieser Klasse...</i>	
	... habe ich meine Stundenziele gut erreicht.	Self-developed
	... gelang es mir, auch in unvorhergesehenen Situationen angemessen zu reagieren.	Self-developed
	... fiel es mir sehr leicht, guten Unterricht zu gestalten.	Self-developed
	... hatte ich das Gefühl, meinen Schüler:innen wirklich etwas beigebracht zu haben.	Self-developed
	... war ich zufrieden damit, wie ich den Unterricht gestaltet habe.	Self-developed
	... hatte ich das Gefühl, eine gute Lehrkraft gewesen zu sein.	Self-developed
	... fühlte ich mich beim Unterrichten kompetent.	Self-developed
	... hatte ich das Gefühl, die Schüler:innen richtig erreicht zu haben.	Self-developed
Motivation for teaching – Intrinsic value	<i>Die [ersten Englischstunden dieses Schuljahres/letzten Englischstunden] in dieser Klasse...</i>	
	... haben mir sehr viel Spaß gemacht.	Self-developed
	... gehörten zu meinen Lieblingsstunden.	Self-developed
	... habe ich sehr genossen.	Self-developed
	... haben mir sehr viel Freude bereitet, meine Schüler:innen zu unterrichten.	Self-developed
	... hatte ich das Gefühl, den richtigen Beruf gewählt zu haben.	Self-developed
	... unterrichtete ich mit Begeisterung.	Self-developed

Motivation for teaching – Attainment value	<i>In den [ersten Englischstunden dieses Schuljahres/den letzten Englischstunden] in dieser Klasse...</i>	
	... war es mir persönlich besonders wichtig, meinen Schüler:innen etwas beizubringen.	Self-developed
	... war es mir wichtig, von den Schüler:innen als gute Lehrer:in wahrgenommen zu werden.	Self-developed
	... lag mir sehr viel daran, guten Unterricht zu machen.	Self-developed
Motivation for teaching – Utility value	... hatte ich persönlich den Anspruch, sehr guten Unterricht zu gestalten.	Self-developed
	<i>Die [ersten Englischstunden dieses Schuljahres/letzten Englischstunden] in dieser Klasse haben mir wieder einmal die Relevanz des Lehrberufs vor Augen geführt, weil...</i>	
	... ich durch das Unterrichten viel für das eigene Leben lernen konnte.	Self-developed
	... das Unterrichten für mich abwechslungsreich war.	Self-developed
Motivation for teaching – Cost	... ich durch den Beruf als Lehrer:in viel Wertschätzung erfahren habe.	Self-developed
	... ich die Entwicklung von Kindern und Jugendlichen mitgestalten konnte.	Self-developed
	<i>Die [ersten Englischstunden dieses Schuljahres/letzten Englischstunden] in dieser Klasse...</i>	
	... haben mich eine Menge Energie gekostet.	Self-developed
	... haben mich unter Stress gesetzt.	Self-developed
	... haben viele Selbstzweifel in mir ausgelöst.	Self-developed
	... haben mich hoffnungslos gestimmt.	Self-developed
	... haben mich enttäuscht.	Self-developed
	... waren langweilig.	Self-developed
	... haben mich frustriert.	Self-developed
	... habe ich unter viel Zeitaufwand vorbereitet.	Self-developed

TPACK – Pedagogical knowledge	... haben/hätten mich viel Vorbereitungszeit gekostet, um guten Unterricht zu machen.	Self-developed
	... haben mich schon in den Tagen davor beunruhigt.	Self-developed
	<i>Im Folgenden würden wir gerne erfahren, wie Sie sich selbst in Bezug auf Ihr pädagogisches, technologisches und technologisch-pädagogisches Können einschätzen.</i>	
	Ich weiß, wie ich feststellen kann, wo meine Schüler:innen gerade stehen.	Schmidt et al. (2009)
	Ich kann meinen Unterricht daran anpassen, was Schüler:innen bereits können oder noch nicht können.	Schmidt et al. (2009)
	Ich bin in der Lage, meinen Unterrichtsstil an unterschiedliche Arten von Lerner:innen anzupassen.	Schmidt et al. (2009)
	Ich kann den Lernerfolg meiner Schüler:innen auf unterschiedliche Weise erfassen.	Schmidt et al. (2009)
	Ich kann auf sehr unterschiedliche Weise unterrichten.	Schmidt et al. (2009)
	Ich bin mit typischen Fehlern und (Fehl-)Vorstellungen von Schüler:innen in Englisch vertraut.	Schmidt et al. (2009)
	Ich weiß, wie ich gutes Classroom Management begünstigen und aufrechterhalten kann.	Schmidt et al. (2009)
TPACK – Technology knowledge	Ich weiß mir bei technischen Problemen zu helfen.	Schmidt et al. (2009)
	Ich erlerne den Umgang mit neuen digitalen Medien schnell.	Schmidt et al. (2009)
	Ich bin immer auf dem aktuellsten Stand der wichtigsten neuen digitalen Medien.	Schmidt et al. (2009)
	Ich spiele häufig mit digitalen Medien herum.	Schmidt et al. (2009)
	Ich kenne viele verschiedene digitale Medien.	Schmidt et al. (2009)
	Ich habe die technischen Fertigkeiten, um digitale Medien einzusetzen.	Schmidt et al. (2009)
	Ich hatte bereits ausreichend Gelegenheiten, mit verschiedensten digitalen Medien zu arbeiten.	Schmidt et al. (2009)

TPACK – Pedagogical content knowledge	Ich bin in der Lage, angemessene Unterrichtsmethoden auszuwählen, um das fachspezifische Denken und Lernen der Schüler:innen in Englisch anzuleiten.	Adapted to the domain of English from Schmidt et al. (2009)
TPACK – Technological pedagogical knowledge	Ich bin in der Lage digitale Medien auszuwählen, die den Lehransatz einer Unterrichtsstunde verbessern.	Schmidt et al. (2009)
	Ich bin in der Lage digitale Medien auszuwählen, die den Lernprozess der Schüler:innen in einer Stunde verbessern.	Schmidt et al. (2009)
	Meine Ausbildung hat mich dazu angeregt, mich tiefer damit auseinander zu setzen, wie der Einsatz von digitalen Medien meine Lehransätze beeinflussen kann.	Schmidt et al. (2009)
	Ich prüfe kritisch, wie ich digitale Medien in meinem Unterricht einsetzen kann.	Schmidt et al. (2009)
	Ich kann den Einsatz von digitalen Medien, die ich neu kennenlerne, an verschiedene Unterrichtsaktivitäten anpassen.	Schmidt et al. (2009)
TPACK – Technological content knowledge	Mir sind verschiedene digitale Medien bekannt, mit denen ich das Verstehen und Anwenden von Englisch unterstützen kann.	Adapted to the domain of English from Schmidt et al. (2009)
TPACK – Technological pedagogical content knowledge	Ich bin in der Lage so zu unterrichten, dass Fachinhalte, digitale Medien und verschiedene Unterrichtsmethoden sinnvoll aufeinander abgestimmt sind.	Adapted from Schmidt et al. (2009)
	Ich bin in der Lage digitale Medien auszuwählen, die das, was ich unterrichte und was meine Schüler:innen lernen, verbessern.	Adapted from Schmidt et al. (2009)
	Ich kann Kolleg:innen dabei unterstützen, den Einsatz von Fachinhalt, digitalen Medien und Unterrichtsmethoden sinnvoll aufeinander abzustimmen.	Adpted from Schmidt et al. (2009)
	Ich bin in der Lage digitale Medien so auszuwählen, dass sie dem Ziel meiner Unterrichtsstunde zuträglich sind.	Schmidt et al. (2009)
TPACK (FeedBook-specific)	Ich bin in der Lage so zu unterrichten, dass Fachinhalte, das FeedBook-System und verschiedene Unterrichtsmethoden sinnvoll aufeinander abgestimmt sind.	Adapted from Schmidt et al. (2009)

Instructional quality – Cognitive activation	Ich bin in der Lage Themen und Aufgaben im FeedBook auszuwählen, die das unterstützen, was ich unterrichte und was meine Schüler:innen lernen.	Adapted from Schmidt et al. (2009)
	Ich bin in der Lage Themen und Aufgaben im FeedBook auszuwählen, die meine Schüler:innen bei ihren (Haus-)Aufgaben unterstützen.	Adapted from Schmidt et al. (2009)
	Ich bin in der Lage Themen und Aufgaben im FeedBook so auszuwählen, dass sie dem Ziel meiner Unterrichtsstunde zuträglich ist.	Adapted from Schmidt et al. (2009)
	<i>Bitte schätzen Sie im Folgenden Ihre Klasse und den Englischunterricht in dieser Klasse ein.</i>	
	Mir ist wichtig, dass die Schüler:innen regelmäßig üben.	Adapted from Jaekel et al. (2021), originally adapted from Baumert et al. (1997) and Ramm et al. (2006)
	Übungsphasen sind für die Schüler:innen eine gute Gelegenheit, um zu sehen, ob sie den Unterrichtsstoff beherrschen.	Adapted from Jaekel et al. (2021), originally adapted from Baumert et al. (1997) and Ramm et al. (2006)
	Im Unterricht üben und wiederholen die Schüler:innen das Gelernte, bis es sitzt.	Adapted from Jaekel et al. (2021)
	Die Schüler:innen üben regelmäßig, damit sie die Aufgaben immer schneller und sicherer bearbeiten können.	Adapted from Jaekel et al. (2021), originally adapted from Baumert et al. (1997) and Ramm et al. (2006)
Instructional quality – Classroom management – Monitoring	Ich weiß im Unterricht genau, was in der Klasse vor sich geht.	Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

**Instructional quality –
Classroom management
– Disturbances**

Ich greife ein, bevor Unruhe und Störungen überhaupt erst entstehen.

Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

Ich achte sehr darauf, dass alle aufpassen.

Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

Ich merke sofort, wenn Schüler:innen beginnen, etwas anderes zu treiben.

Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

Im Unterricht ist es manchmal laut und alles geht durcheinander.

Adapted from Kunter et al. (1997)

Der Unterricht wird oft gestört.

Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

Manchmal muss ich zu Beginn der Stunde lange warten, bis Ruhe eintritt.

Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

Im Unterricht wird fortwährend laut gequatscht.

Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

Im Unterricht wird manchmal auch Zeit vertrödelte

Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

Im Unterricht dauert es meistens sehr lange, bis alle zur Arbeit bereit sind.

Jaekel et al. (2021), originally adapted from Baumert et al. (1997)

Instructional quality – Support – Handling mistakes	Ich bin geduldig, wenn jemand im Unterricht einen Fehler macht.	Ramm et al. (2006)
	Für mich ist es nichts Schlimmes, wenn jemand im Unterricht einen Fehler macht.	Ramm et al. (2006)
	Wenn jemand im Unterricht etwas falsch macht, gebe ich ihm/ihr die Möglichkeit, es zu korrigieren oder noch einmal anzufangen.	Jaekel et al. (2021)
	Ich achte darauf, dass im Unterricht niemand ausgelacht wird, der/die einen Fehler macht.	Ramm et al. (2006)
	Wenn ein:e Schüler:in etwas falsch macht, erkläre ich ihr oder ihm, warum es falsch ist.	Jaekel et al. (2021)
Instructional quality – Support – Internal differentiation	Ich stelle unterschiedlich schwere Fragen, je nachdem, wie gut ein:e Schüler:in ist.	Baumert et al. (1997)
	Im Unterricht gebe ich den einzelnen Schüler:innen verschiedene Aufgaben.	Baumert et al. (1997)
	Ich Sorge dafür, dass die Schüler:innen Materialien haben, mit denen sie gut lernen können.	Jaekel et al. (2021)
	Ich verlange von den guten Schüler:innen deutlich mehr.	Baumert et al. (1997)
	Ich gebe den Schüler:innen für die Bearbeitung von Aufgaben genügend Zeit.	Jaekel et al. (2021)
Instructional quality – Support – Enthusiasm	Ich unterrichte mit großer Begeisterung.	Adapted from Kunter et al. (2008, 2011)
	Mir macht das Unterrichten großen Spaß.	Adapted from Kunter et al. (2008, 2011)
	Es bereitet mir immer große Freude, Schüler:innen etwas Neues beizubringen.	Adapted from Kunter et al. (2008, 2011)
	Mir bereitet der Kontakt mit den Schüler:innen große Freude.	Adapted from Kunter et al. (2008, 2011)

Instructional quality – Support – Feedback	Es ist eine Freude zu unterrichten.	Adapted from Kunter et al. (2008, 2011)
	Die Schüler:innen erfahren von mir immer wieder, wo sie mit ihren Leistungen stehen.	Jaekel et al. (2021), originally adapted from Klimczak et al. (2012)
	Ich gebe den Schüler:innen regelmäßig Rückmeldung, was sie schon können.	Jaekel et al. (2021), originally adapted from Klimczak et al. (2012)
	Ich gebe den Schüler:innen regelmäßig Rückmeldung, was sie noch nicht so gut können.	Jaekel et al. (2021), originally adapted from Klimczak et al. (2012)
	Ich zeige den Schüler:innen regelmäßig, wie sie sich verbessern können.	Jaekel et al. (2021), originally adapted from Klimczak et al. (2012)
Instructional quality – Support – Autonomy support	<i>Es ist mir wichtig...</i>	
	... meinen Schüler:innen Wahlmöglichkeiten und Optionen anzubieten.	Adapted from student perspective based on Williams & Deci (1996)
	... meinen Schüler:innen stets Verständnis entgegenzubringen.	Adapted from student perspective based on Williams & Deci (1996)
	... meinen Schüler:innen Vertrauen in ihre Englisch-Fähigkeiten zu vermitteln.	Adapted from student perspective based on Williams & Deci (1996)

	... meine Schüler:innen zu ermutigen, Fragen zu stellen.	Adapted from student perspective based on Williams & Deci (1996)
	... mir anzuhören, was meine Schüler:innen gerne machen würden.	Adapted from student perspective based on Williams & Deci (1996)
	... erst zu verstehen, wie meine Schüler:innen etwas sehen, bevor ich einen neuen Weg vorschlage.	Adapted from student perspective based on Williams & Deci (1996)
Satisfaction with job	<i>Wie zufrieden sind Sie mit Ihrem Beruf?</i>	
	In unserem Beruf ist es schwer, glücklich zu werden.	Kunter et al. (2008)
	Für mich überwiegen in unserem Beruf eindeutig die Vorteile.	Kunter et al. (2008)
	Wenn ich noch einmal wählen könnte, würde ich sofort wieder Lehrer:in werden.	Kunter et al. (2008)
	Ich habe mir nicht nur einmal überlegt, ob es für mich nicht besser gewesen wäre, einen anderen Beruf zu ergreifen.	Kunter et al. (2008)
	Für mich gibt es keinen besseren Beruf.	Kunter et al. (2008)
	Ich bereue manchmal, Lehrer:in geworden zu sein.	Kunter et al. (2008)
Exhaustion	<i>Inwieweit treffen folgende Belastungen auf Sie zu?</i>	
	Ich fühle mich in der Schule oft erschöpft.	Baumert et al. (2009)
	Ich empfinde es als Problem, in so vielen Klassen unterrichten zu müssen.	Baumert et al. (2009)
	Ich fühle mich insgesamt überlastet.	Baumert et al. (2009)
	Ich merke öfter in der Schule, wie lustlos ich bin.	Baumert et al. (2009)
	Manchmal bin ich am Ende des Schultags echt deprimiert.	Baumert et al. (2009)
Changing profession	Würden Sie in einen anderen Beruf wechseln, wenn Sie die Möglichkeit dazu hätten?	Voss et al. (2014)

Homework	Gibt es an Ihrer Schule das Konzept der Hausaufgaben?	Self-developed
	Wie oft erteilen Sie Hausaufgaben in Englisch?	Self-developed
	Wie oft haben Sie in Task Cycle [1/2/3/4] Hausaufgaben in Englisch erteilt (egal in welcher Form)?	Self-developed
	Wie oft haben Sie in Task Cycle [1/2/3/4] Hausaufgaben im FeedBook erteilt?	Self-developed
	Wie oft kontrollieren Sie, ob die Schüler:innen ihre Hausaufgaben tatsächlich gemacht haben?	
	Wie oft haben Sie während Task Cycle [1/2/3/4] kontrolliert, ob die Schüler:innen ihre Hausaufgaben tatsächlich gemacht haben?	Self-developed
	Wie viel Zeit verwenden Sie normalerweise durchschnittlich zur Besprechung der Hausaufgaben?	Self-developed
	Wie viel Zeit haben Sie während Task Cycle [1/2/3/4] durchschnittlich zur Besprechung der Hausaufgaben verwendet?	Self-developed
	Nehmen Sie sich normalerweise regelmäßig Zeit, allen Schüler:innen individuelle Rückmeldungen zu den Hausaufgaben zu geben?	Self-developed
Task persistence	Haben Sie sich während Task Cycle 1 regelmäßig Zeit genommen, allen Schüler:innen individuelle Rückmeldungen zu den Hausaufgaben zu geben?	Self-developed
	Ich versuche etwas immer und immer wieder, auch wenn ich beim ersten Mal keinen Erfolg hatte.	Steinberg et al. (2007), originally adapted from TPQ by Cloninger (1987)
	Ich arbeite häufig an Dingen weiter, wenn Andere schon aufgegeben haben.	Steinberg et al., (2007), originally adapted from TPQ by Cloninger (1987)
Personality - Conscientiousness	<i>Im Folgenden werden Ihnen einige Aussagen präsentiert, die verschiedene Facetten von Personen beschreiben können. Inwiefern stimmen Sie den folgenden Aussagen in Bezug auf Ihre eigene Person zu?</i>	

Personality - Extraversion	Ich erledige meine Aufgaben gründlich.	BFI-K, Rammstedt & John (2005)
	Ich bin bequem, neige zu Faulheit.	BFI-K, Rammstedt & John (2005)
	Ich bin tüchtig und arbeite flott.	BFI-K, Rammstedt & John (2005)
	Ich mache Pläne und führe sie auch durch.	BFI-K, Rammstedt & John (2005)
	Ich bin eher zurückhaltend, reserviert.	BFI-K, Rammstedt & John (2005)
	Ich bin begeisterungsfähig und kann andere leicht mitreißen.	BFI-K, Rammstedt & John (2005)
Personality - Agreeableness	Ich bin eher der/die "Stille", wortkarg.	BFI-K, Rammstedt & John (2005)
	Ich gehe aus mir heraus, bin gesellig.	BFI-K, Rammstedt & John (2005)
	Ich neige dazu, andere zu kritisieren.	BFI-K, Rammstedt & John (2005)
	Ich schenke anderen leicht Vertrauen, glaube an das Gute im Menschen.	BFI-K, Rammstedt & John (2005)
Personality - Neuroticism	Ich kann mich kalt und distanziert verhalten.	BFI-K, Rammstedt & John (2005)
	Ich kann mich schroff und abweisend anderen gegenüber verhalten.	BFI-K, Rammstedt & John (2005)
	Ich werde leicht deprimiert, niedergeschlagen.	BFI-K, Rammstedt & John (2005)

	Ich bin entspannt, lasse mich durch Stress nicht aus der Ruhe bringen.	BFI-K, Rammstedt & John (2005)
	Ich mache mir viele Sorgen.	BFI-K, Rammstedt & John (2005)
	Ich werde leicht nervös und unsicher.	BFI-K, Rammstedt & John (2005)
Personality - Openness	Ich bin vielseitig interessiert.	BFI-K, Rammstedt & John (2005)
	Ich bin tief sinnig, denke gerne über Sachen nach.	BFI-K, Rammstedt & John (2005)
	Ich habe eine aktive Vorstellungskraft, bin phantasievoll.	BFI-K, Rammstedt & John (2005)
	Ich schätze künstlerische und ästhetische Eindrücke.	BFI-K, Rammstedt & John (2005)
	Ich habe nur wenig künstlerisches Interesse.	BFI-K, Rammstedt & John (2005)
FeedBook use	Haben Sie das FeedBook in den letzten Wochen während des vorgesehenen Task Cycles eingesetzt?	Self-developed
	Von wann bis wann haben Sie Task Cycle [1/2/3/4] mit Ihrer Klasse bearbeitet?	Self-developed
	Haben Sie die Target Task ([Target Task 1/Target Task 2/Target Task 3/Target Task 4]) wie vorgesehen durchgeführt?	Self-developed
	Falls nein: Ich habe die Target Task wie folgt abgewandelt:	Self-developed
	Ich fand das Angebot an Aufgaben im FeedBook gut.	Self-developed
	Ich fand das Angebot an Aufgaben im FeedBook sinnvoll zum Erlangen der vorgesehen sprachlichen Kompetenzen für meine Schüler:innen.	Self-developed

	Ich finde die Handhabung des FeedBooks ausreichend verständlich.	Self-developed
	Ich finde die Nutzung des FeedBooks intuitiv.	Self-developed
	Ich habe mir die Video-Tutorials angesehen.	Self-developed
	Ich fand die Video-Tutorials sinnvoll.	Self-developed
	Ich habe folgende Vorschläge zur Verbesserung des Nutzungserlebnisses für Lehrkräfte:	Self-developed
	Ich habe folgende Vorschläge zur Verbesserung der Aufgaben(auswahl):	Self-developed
	Ich habe folgende Vorschläge, um das FeedBook besser in meinen Unterricht oder für die Vergabe von Hausaufgaben einbinden zu können:	Self-developed
	Möchten Sie uns zum FeedBook noch etwas mitteilen?	—
	<i>Wie haben Sie das FeedBook während des letzten Task Cycles in dieser Klasse eingesetzt bzw. Aufgaben an die Schüler:innen dieser Klasse vergeben?</i>	
	Ich ließ den Schüler:innen komplette Freiheit in der Nutzung des FeedBooks (sie durften also selbst entscheiden, welche Aufgaben sie bearbeiten wollten).	Self-developed
FeedBook use 2	Ich ließ den Schüler:innen komplette Freiheit in der Nutzung des FeedBooks (sie durften also selbst entscheiden, welche Aufgaben sie bearbeiten wollten).	Self-developed
	Ich vergab Untereinheiten bestimmter Sections/Bereiche an die Schüler:innen, wie z.B. „Words and Phrases – Expressing opinion“.	Self-developed
	Ich vergab meist ganz konkrete Aufgaben an die Schüler:innen, wie z.B. „Die ersten 3 Aufgaben in der Rubrik „Words and Phrases – Expressing opinion“.	Self-developed
	Gab es Sections/Bereiche im letzten Task Cycle, die Sie in dieser Klasse nicht vergeben haben bzw. die nicht von Ihren Schüler:innen dieser Klasse bearbeitet werden sollten?	Self-developed
	Falls ja: Welche Sections/Bereiche haben Sie in dieser Klasse ausgelassen?	Self-developed
	Ich habe mich während des letzten Task Cycles stark an den mir bereitgestellten Stundenentwürfen orientiert.	Self-developed
Material use		

	Ich fand die Stundenentwürfe sinnvoll.	Self-developed
	Ich fand die Stundenentwürfe hilfreich.	Self-developed
	Ich habe während des letzten Task Cycles die bereitgestellten Tafelbilder eingesetzt.	Self-developed
	Ich fand die Tafelbilder sinnvoll.	Self-developed
	Ich fand die Tafelbilder hilfreich.	Self-developed
	Ich habe die Unterrichtsmaterialien für den letzten Task Cycle eingesetzt.	Self-developed
	Ich fand die Unterrichtsmaterialien sinnvoll.	Self-developed
	Ich fand die Unterrichtsmaterialien hilfreich.	Self-developed
	Ich habe die Übungen für den letzten Task Cycle eingesetzt.	Self-developed
	Ich fand die Übungen sinnvoll.	Self-developed
	Ich fand die Übungen hilfreich.	Self-developed
	Ich habe die Listening Dateien für den letzten Task Cycle eingesetzt.	Self-developed
	Ich fand die Listening Dateien sinnvoll.	Self-developed
	Ich fand die Listening Dateien hilfreich.	Self-developed
	Wie viel Zeit haben Sie durchschnittlich vor jeder Stunde aufgewendet, um sich mit den bereitgestellten Materialien vertraut zu machen?	Self-developed
	Fanden Sie diese Vorbereitungszeit rückblickend angemessen?	Self-developed
	Falls nein: Stattdessen wäre diese Zeit vermutlich eher angemessen/notwendig/ausreichend gewesen:	Self-developed
	Möchten Sie uns zu den Materialien noch etwas mitteilen?	Self-developed
Gender	Sind sie männlich, weiblich oder divers?	—
Age	Wie alt sind Sie?	—
Contact time with class	Wie viele Wochenstunden unterrichten Sie Englisch in dieser Klasse?	—

Home teacher?]	Sind Sie Klassenlehrer:in dieser Klasse?	—
Other subjects in this class	Unterrichten Sie in dieser Klasse weitere Fächer?	—
Teaching experience	Wie viele Jahre insgesamt (inklusive Referendariatszeit/Vorbereitungszeit) haben Sie einschließlich dieses Schuljahres unterrichtet?	—
Experience with FeedBook or sibling systems	Haben Sie in der Vergangenheit bereits mit Schwestersystemen des FeedBooks von der Uni Tübingen gearbeitet?	Self-developed
Subjects studied	Welche Fächer haben Sie studiert?	—
Graduation	In welchem Jahr haben Sie den Abschluss gemacht?	—
Subjects teaching	Welche Fächer unterrichten Sie?	—
Favorite subject	Welches Ihrer Unterrichtsfächer bereitet Ihnen am meisten Freude?	—
Impact by pandemic	War Ihr Englischunterricht durch die Pandemie in den letzten Wochen beeinflusst?	Self-developed
Comment	Nutzen Sie das folgende Feld, falls Sie uns noch etwas mitteilen möchten.	—

Note. Not all items were used at all time points.

Appendix B

Overview of Measures Included for Strand 2 (Ambulatory Assessment Data)

Table B.1

Student Short Questionnaires/Ambulatory Assessment (After Each Logout)

Scale/construct	Item (German)	Source
	<i>Wie sehr treffen diese Aussagen auf dich und deine Aufgabenbearbeitung im FeedBook heute zu?</i>	
State motivation – Self-concept	Ich hatte den Eindruck, dass ich die Aufgaben im FeedBook gut bewältigen konnte.	Adapted from trait-like self-concept; originally adapted from Gaspard et al. (2017)
	Ich bin zufrieden mit meiner Leistung bei den Aufgaben.	Adapted from trait-like self-concept; originally adapted from Gaspard et al. (2017)
State motivation – Intrinsic value	Die Bearbeitung der Aufgaben hat mir Spaß gemacht.	Adapted from trait-like intrinsic value; originally adapted from Gaspard et al. (2017)
State motivation – Attainment value	Was ich heute durch die Aufgaben gelernt habe, hat für mich eine Bedeutung.	Adapted from trait-like attainment value; originally adapted from Gaspard et al. (2017)
State motivation – Utility value	Was ich heute mit dem FeedBook gelernt haben, werde ich auch anderswo gebrauchen können.	Adapted from trait-like utility value; originally adapted from Gaspard et al. (2017)

State motivation – Cost	Es war anstrengend für mich, mich mit den Aufgaben im FeedBook zu beschäftigen.	Adapted from trait-like cost; originally adapted from Gaspard et al. (2017)
	Die Beschäftigung mit den Aufgaben hat mich gelangweilt.	Adapted from trait-like cost; self-developed
	Die Beschäftigung mit den Aufgaben hat mich frustriert.	Adapted from trait-like cost; self-developed
	Ich bin enttäuscht, wie ich bei den Aufgaben abgeschnitten habe.	Adapted from trait-like cost; self-developed
Effort	Ich habe mein Bestes bei den Aufgaben gegeben.	Adapted from trait-like effort; originally adapted from Gaspard et al. (2017)

Appendix C

Interview Guides Included for Strand 4 (Qualitative Data)

Guided interviews were and will be conducted with the participating teachers in Hamburg at the beginning and at the end of the school year as well as before and after the Task Cycles. The interviews will allow to analyze in which way the work with the FeedBook influences the teaching process, interactions, perceived motivation, and perceived performance. There are five interviews per teacher planned. In the following, we will provide the interview guides for the first three of these interviews. The other two interview guides will be published in an amendment to this pre-registration (amendments will be accessible via PsychArchives (<https://doi.org/10.23668/psycharchives.5366>) at “Related PsychArchives Objects”).

First Interview

Berufsbiographische Einstiegserzählung

- Sie sind Lehrkraft für Englisch an einem Gymnasium in Hamburg. Könnten Sie mir sagen, wie es dazu gekommen ist, dass sie Englischlehrer:in dort geworden sind?
- Wie lange sind Sie bereits Lehrkraft?
- Was macht Ihnen als Englischlehrkraft besonders viel Freude?

TBLT und FeedBook

- Kennen Sie den TBLT Ansatz?
- Was sind Ihre Erwartungen oder Wünsche an die Arbeit mit dem Feedbook?
 - Was sind Ihre Erwartungen an ein digitales Tool im Englischunterricht?
- Welche Rolle spielt Ihrer Meinung nach das Üben im Englischunterricht?

Interaktion im Unterricht

- Wie würden Sie eine typische bzw. eine ideale Interaktion im Englischunterricht definieren?

- Wenn Sie mit eine/r Schüler:innen sprechen, was geht Ihnen durch den Kopf, was sind Ihre Ziele bei diesen Gesprächen?

Diagnostik und Heterogenität

- Wenn Sie feststellen, dass ein/e Schüler:in ein Problem (in Bezug auf den Englischunterricht) hat, wie gehen Sie vor?
- Wie würden Sie beurteilen, ob ihre Einschätzung eines Schülers/ einer Schülerin (SuS) richtig ist?

Second Interview

Umgang mit dem FeedBook

- Erzählen Sie wie Sie den Task Cycle 1 erlebt haben?
 - Liegt der Fokus auf Arbeit mit dem FeedBook?
 - Liegt der Fokus auf dem Unterrichtsgeschehen?
 - Liegt der Fokus auf dem Lernerfolg der SuS?
 - Liegt der Fokus auf organisatorischen Angelegenheiten?
- Was hat Ihrer Meinung nach besonders gut geklappt? Warum?
- Was hat Ihrer Meinung nach nicht gut geklappt? Warum?
- Hat die Arbeit mit FeedBook Ihre Interaktion im Unterricht beeinflusst? Wenn ja, wann war das und warum?

Nachfragen zur Diagnostik (kognitiven Leistungsdispositionen und Performanz)

- Erzählen Sie, wie Sie das Erlernen der Grammatik bei den SuS wahrgenommen haben?
- Wie sind sie mit Schwierigkeiten umgegangen?
 - Haben Sie hier das FeedBook genutzt?
 - Haben Sie den Unterricht aufgrund von Schwierigkeiten angepasst?

- Hat sich Ihre diagnostische Arbeit durch die Nutzung des FeedBooks verändert?
- Wie wären Sie ohne FeedBook vorgegangen?

Einleitungsfragen zum Konzept des **Computational Thinking**

- Auf einer Skala von 1-10: Wie hoch ist ihr Vertrauen in ein digitales Tutorensystem?
 - Können Sie ihre Entscheidung begründen?
- Sehen Sie einen Zusammenhang zwischen der Informatik und dem Sprachenlernen?
 - Wenn ja, wo genau und warum?
- Wie gehen Sie vor, wenn Sie SuS einen komplexen Sachverhalt erklären müssen?
 - Wederholen Sie Ihre Strategie?

Third Interview

Umgang mit dem **FeedBook**

- Erzählen Sie wie Sie den Task Cycle 2 erlebt haben?
 - Liegt der Fokus auf Arbeit mit dem FeedBook?
 - Liegt der Fokus auf dem Unterrichtsgeschehen?
 - Liegt der Fokus auf dem Lernerfolg der SuS?
 - Liegt der Fokus auf organisatorischen Angelegenheiten?
- Was hat Ihrer Meinung nach besonders gut geklappt? Warum?
- Was hat Ihrer Meinung nach nicht gut geklappt? Warum?
 - Wie sind Sie mit Schwierigkeiten umgegangen?
- Hat die Arbeit mit FeedBook Ihre Interaktion im Unterricht beeinflusst? Wenn ja, wann war das und warum?
- Haben Sie das FeedBook für die Planung Ihres Unterrichts genutzt?

- Haben Sie die Zusatzmaterialien gut nutzen können? Gab es Schwierigkeiten (bspw. Stundenverlaufspläne)?
 - Haben Sie Anpassungen vorgenommen?

Verknüpfungen zwischen „**gutem Unterricht**“ und TBLT im **digitalen Setting**

- Waren Sie mit der Target Task zufrieden?
 - Woran machen Sie ihre Antwort feste?
- Erzählen Sie, wie sie das Erlernen der Grammatik bei den SuS wahrgenommen haben?
- Sind Sie der Meinung, dass digitales Üben zu einer besseren Performanz der SuS im Unterricht führt?
 - Hat sich Ihre persönliche Performanz im Unterricht verändert?
- Nehmen Sie das Tutorensystem als Co-Teacher war?
 - Wie würden Sie Ihre Zusammenarbeit mit dem System beschreiben?
- Wie würden Sie mit dem digitalen System arbeiten, wenn es nicht mit der Unit im Buch verbunden wäre?
 - Wie würden Sie die Unterschiede und Gemeinsamkeiten zwischen Ihrer Arbeit mit dem Workbook und der Software beschreiben?
- Hat sich ihre Rolle als Lehrkraft durch die Arbeit mit dem FeedBook verändert?
 - Können Sie sich vorstellen, dass sich Ihre Rolle in Zukunft verändern wird?
Wenn ja, inwiefern?

Ausblick

- Was sind Ihre Wünsche für den dritten Task Cycle?