![ETS — Measuring the Power of Learning.®]

Research Memorandum
ETS RM–18-11

A Good Fit?
Examining the Alignment Between the *TOEFL Junior*® Standard Test and the English as a Foreign Language Curriculum in Berlin, Germany

Veronika Timpe-Laughlin

October 2018

ETS Research Memorandum Series

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Memorandum series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Memorandum series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

A Good Fit?

Examining the Alignment Between the *TOEFL Junior*® Standard Test and the English as a Foreign Language Curriculum in Berlin, Germany

Veronika Timpe-Laughlin

Educational Testing Service, Princeton, New Jersey

October 2018

Corresponding author: V. Timpe-Laughlin, E-mail: VLaughlin@ets.org

Abstract

Over the past years, students in Berlin, Germany, have taken the *TOEFL Junior*® Standard test as part of an English language contest called Go4Goal! Given consistently increasing numbers of participating students, schools have asked whether it would be possible to use the TOEFL Junior Standard test in their English as a foreign language (EFL) instruction in the form of a *Klassenarbeitsersatz* (i.e., a substitute for 1 of the 4 mandatory written classroom assessments). To ensure that an external assessment constitutes an appropriate measurement tool for a given educational context and specific test use, it is imperative to conduct alignment studies and gauge the fit of the external assessment for the specific local use. This study investigates the alignment between the official EFL curriculum mandated by the ministry of education in the state of Berlin and the competencies and language skills assessed by the TOEFL Junior Standard test. To gauge the fit, curricula were reviewed and 40%–50% of the activities included in 4 different textbooks for Grades 7, 8, 9, and 10 were coded systematically. The frequencies were tallied to show the match between (a) what the test measures and (b) what is being taught in secondary-level EFL classes in Berlin. Additionally, the results from the coding exercise were triangulated with teacher perceptions of the TOEFL Junior Standard test. The results from the systematic coding and the teachers' perceptions suggest that the TOEFL Junior Standard test is an appropriate measure for secondary-level EFL learners in Berlin.

Key words: alignment, *TOEFL Junior*® Standard test, curriculum, English as a foreign language, Germany

Table of Contents

To ensure that an external assessment constitutes an appropriate measurement tool for a given educational context and specific test use, experts have oftentimes highlighted the need to conduct alignment studies (e.g., Herman & Webb, 2007; Roach, Niebling, & Kurz, 2008; Webb, 2007). *Alignment* is thereby defined as "the extent to which curricular expectations and assessments are in agreement and work together to provide guidance for educators' efforts to facilitate students' progress toward desired academic outcomes" (Roach et al., 2008, p. 158). In other words, it is important to evaluate the general match and coherence between an external, potentially large-scale test and the local curriculum and classroom instruction to ensure that the test is appropriate for the local context and is conducive to students' educational achievements. In language testing, studies are oftentimes conducted to investigate the alignment between an assessment and a given set of standards or a framework (e.g., Harsch & Hartig, 2015; Papageorgiou, 2016). To a lesser extent, the field has seen alignment studies that evaluate the match of a large-scale assessment for a particular educational context (e.g., Hsieh, 2015).

Adding to the latter strand of alignment research within the field of language assessment, this study aimed at investigating the fit of the *TOEFL Junior*® Standard test for a particular local context and use: secondary-level English as a foreign language (EFL) education in Berlin, Germany. Over the past 5 years, secondary-level students at different schools in Berlin have participated in an English language contest called Go4Goal![1] In the context of Go4Goal!, students, together as a class, take the TOEFL Junior Standard test, a test of EFL designed for secondary-level EFL learners. Given consistently increasing numbers of participating students, schools have asked whether it would be possible to use the TOEFL Junior Standard test in their EFL instruction in the form of a *Klassenarbeitsersatz* (a substitute for one of the four mandatory written classroom assessments).

In EFL instruction in Germany, students traditionally have to take four written, paper-delivered classroom assessments (i.e., *Klassenarbeiten*) in their English classes per term or semester. These norm-referenced classroom assessments are usually designed and/or compiled by individual teachers based on their instruction and the textbook used in class. The assessments are graded by the respective classroom teachers who mark students' performances on a scale ranging from 1 (*sehr gut* [very good]) to 6 (*ungenügend* [insufficient/fail]).[2] Together with other performances, such as oral in-class participation, students' performances on these four written classroom assessments are then factored into their final mark at the end of the term. As per the

education acts mandated by the ministry of education in the state of Berlin, schools and/or teachers are allowed to substitute one of these written classroom assessments with a performance assessment that is to a large degree equivalent to a *Klassenarbeit* in that it assesses the competences and skills that EFL teachers promote in their secondary-level learners.[3] In late 2017, the Research and Development division at Educational Testing Service (ETS) was approached by Language and Testing Service GmbH (LTS) to investigate whether the TOEFL Junior Standard test could serve as a *Klassenarbeitsersatz* for secondary-level learners in schools throughout Berlin, Germany.

Hence this study investigated the alignment between the official EFL curriculum mandated by the ministry of education in the state of Berlin and the competences and language skills assessed by the TOEFL Junior Standard test. With regard to alignment, Resnick, Rothman, Slattery, and Vranek (2004) argued that it is

> not a yes-or-no question; rather, it consists of a number of dimensions that collectively tell the story of the degree of match between the expectations states have for students' performance and the measure used to gauge whether students are meeting those expectations. (p. 6)

Following Resnik et al.'s advice to aggregate evidence from multiple sources to finally make an informed decision with regard to the degree of alignment, this study drew on both curricular documents and teaching materials as well as expert opinions. That is, in addition to a close and systematic matching of materials used in EFL instruction with the content of the test, we also regarded it as important to consider the views of teachers at different schools in Berlin whose students have been engaged in Go4Goal! on the basis of the same test. In sum, the aim of this study was to systematically examine whether the TOEFL Junior Standard test is a good fit for assessing the skills, abilities, and competences fostered in the context of secondary-level EFL education in Berlin, Germany.

The TOEFL Junior Standard Test

The TOEFL Junior Standard test was designed for EFL learners between the ages of 11 and 17 years. The TOEFL Junior Standard test is not based on a specific English curriculum. Instead, it was designed to measure learners' English skills and competences at the middle school level (for the complete test design framework, see So et al., 2015). While some test tasks

"assess underlying enabling skills, such as grammatical and lexical knowledge, the main emphasis of the test is the measurement of communicative competence, that is, the ability to use language for communicative purposes" (So et al., 2015, p. 2).

All tasks included in the TOEFL Junior Standard test correspond to tasks in the target language use (TLU) domain of English-medium middle school environments. This TLU domain is further subdivided into three subdomains: academic, social–interpersonal, and navigational (see Figure 1).
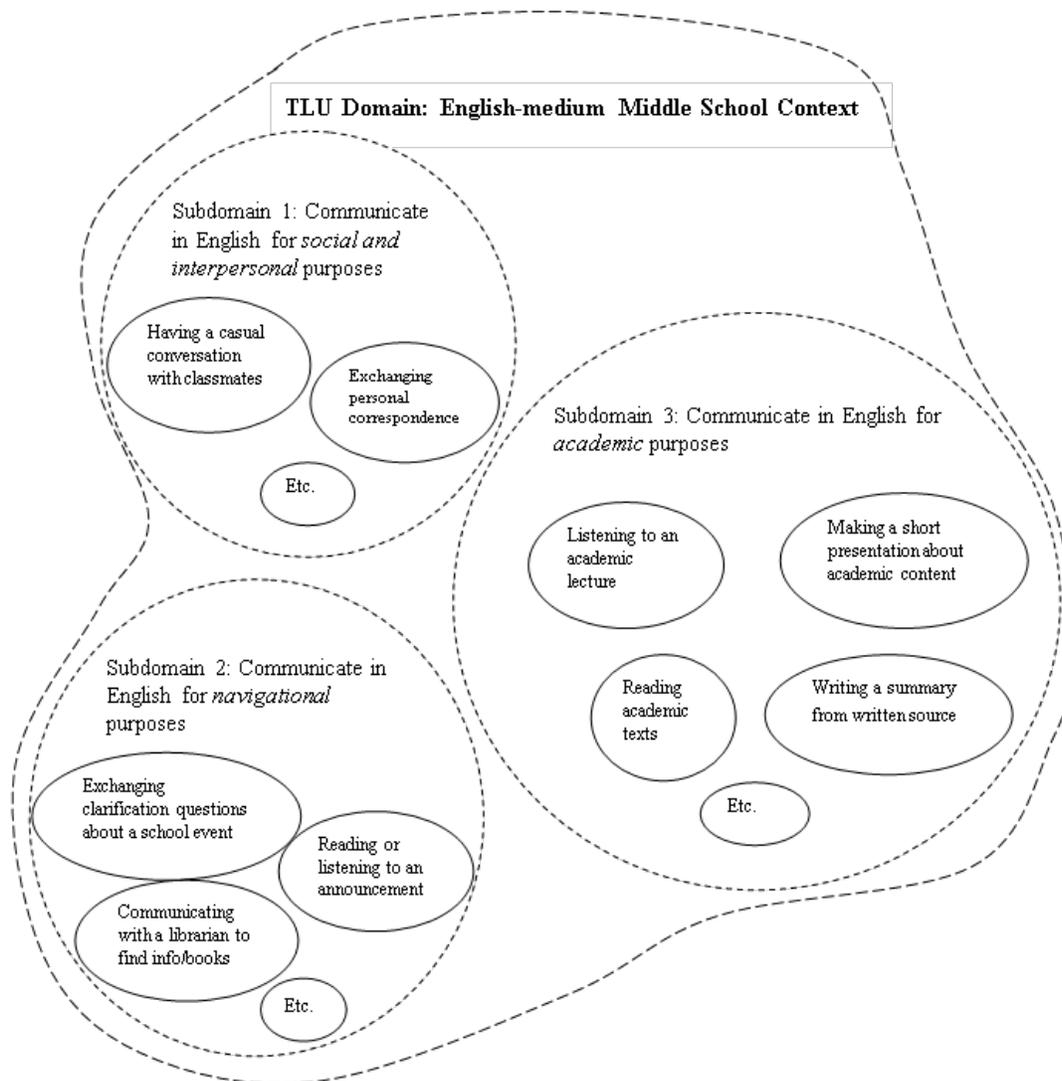

Figure 1. TOEFL Junior Standard test's target language use domain. Adapted from So et al. (2015). Adapted from *TOEFL Junior Design Framework* (TOEFL Junior Research Report No. TOEFL JR-02) by Y. So, M. K. Wolf, M. C. Hauck, P. Mollaun, P. Rybinski, D. Tumposky, & L. Wang, 2015. Copyright 2015 by Educational Testing Service.

As shown in Table 1, the test includes three sections that measure listening comprehension, language form and meaning, and reading comprehension. Each test section consists of 42 multiple-choice items.

Table 1. Structure of the TOEFL Junior Standard Test

Section	Time	No. items
Listening Comprehension	40 minutes	42
Language Form and Meaning	25 minutes	42
Reading Comprehension	50 minutes	42
Total	1 hour 55 minutes	126

In the language form and meaning section, students are provided with reading passages in which four to eight words or collocations are missing. From a list of four possible answer options, test takers have to select the answer option that completes the sentence correctly either from a grammatical perspective (language form) or from a semantic perspective (language meaning). With regard to the language form–oriented items, the following grammatical aspects are included in the assessment: subject–verb agreement, correct subject–object form, verb tense and aspect, active/passive voice, relative clauses, word order, and comparative/superlative forms. The language meaning–oriented items include these parts of speech: verbs, nouns, adjectives, determiners, adverbs, conjunctions, and prepositions (for more detailed information, see So et al., 2015).

In the listening comprehension section, test takers listen to aural input. Reflective of the three TLU subdomains, the listening passages include short conversations, classroom instructions, and academic texts. As presented in Table 2, each of the three subdomains assesses a number of listening subskills. Similar to the listening comprehension section, the reading comprehension section provides test takers with written texts followed by multiple-choice items. Table 2 shows the correspondence between type of stimulus material and the TLU subdomain targeted by each stimulus type in the listening and the reading comprehension sections of the TOEFL Junior Standard test (So et al., 2015).

Table 2. Domain–Skill–Stimulus Correspondence for the Reading and Listening Sections of TOEFL Junior Standard Test

Skill/TLU subdomain	Stimulus material	Subskills
Listening		
Social and interpersonal	Short conversations	Comprehending the main idea
		Identifying salient details
		Making inferences
		Making predictions
		Identifying speaker's purpose
		Understanding meaning conveyed by prosodic features
Navigational	Classroom instructions (monologic)	Comprehending the main idea
		Identifying salient details
		Making inferences
		Making predictions
Academic	Academic (monologic/dialogic)	Comprehending the main idea
		Identifying salient details
		Making inferences
		Making predictions
		Identifying speaker's purpose
Reading		
Social and interpersonal	Correspondence	Comprehending main idea
Navigational	Nonlinear text, journalism	Identifying important supporting factual information
		Making inferences
Academic	Expository	Discerning the meaning of low-frequency words or expressions from context
		Recognizing an author's purpose or use of particular rhetorical structures
		Understanding figurative and idiomatic language from context

Note. TLU = target language use.

The TOEFL Junior Standard Test in Berlin

Berlin's Secondary School System

As shown in Figure 2, the educational system in Berlin consists of an elementary school (mandatory for children from 6 to 12 years of age) and two main types or tracks of secondary schools: (a) *Gymnasium*[4] and (b) *integrierte Sekundarschule*[5] (*ISS*). Both the *Gymnasium* and the *ISS* begin with Grade 7. The *Gymnasium* is an academic secondary school that prepares students for a university education. After Grade 12, students graduate with a diploma called *Abitur* by passing oral and written examinations. By contrast, in the *ISS*, students are more supported depending on their individual needs than in the context of a *Gymnasium*. At an *ISS*, students can complete various qualifications after Grade 10: a certificate of vocational education (*Berufsbildungsreife*), an extended certificate of education (*erweiterte Bildungsreife*), or a

certificate of intermediate education (*mittlerer Schulabschluss*). After Grade 10, they can also enter the *Gymnasium* track within the *ISS* and graduate with *Abitur*[6] after Grade 13.[7]

In addition to these two types of secondary schools, the *Gemeinschaftsschule* constitutes a rather recent, new type of school. Still in the pilot phase, *Gemeinschaftsschulen* aim to provide education to both primary- and secondary-level learners. *Gemeinschaftsschulen* can be loosely compared to *ISS*s with the difference that *Gemeinschaftsschulen* tend to operate under a more inclusive approach in that they provide education to students irrespective of their socioeconomic background and level of performance/achievement, two aspects that traditionally dictate the track into which a student is placed.
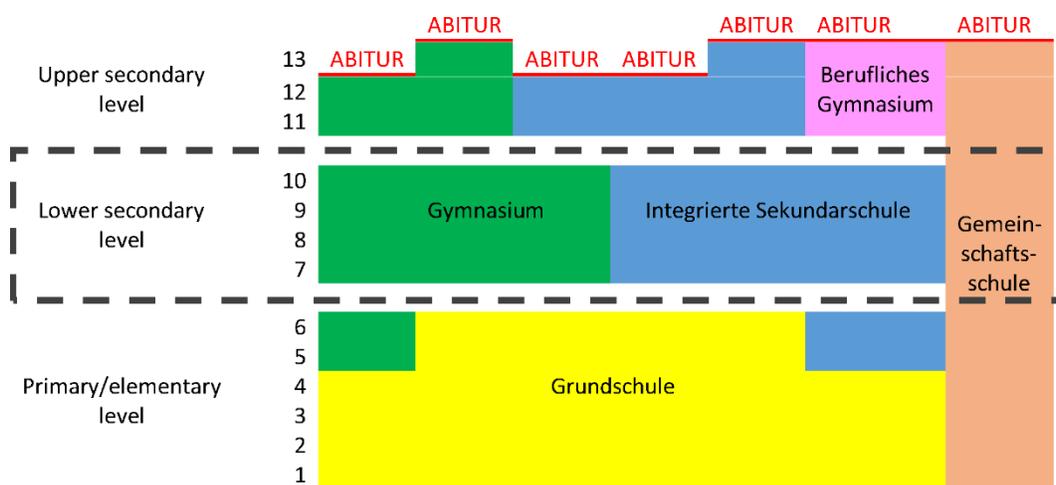
Figure 2. Overview of the school system in Berlin. Some elementary schools run from Grades 1 to 4, others from Grades 1 to 6. In both cases, students can proceed to either *Gymnasium* or *integrierte Sekundarschule* and graduate with *Abitur* either after 12th or 13th grade. Note that this study focuses on the lower secondary level only. Adapted from http://www.berlin.de/sen/bildung/schule/bildungswege/

EFL is a mandatory, core subject in both elementary- and secondary-level education in Berlin. Most students begin learning English in third grade (although some elementary schools offer it in first grade). English remains a mandatory subject throughout students' secondary-level education at the *Gymnasium*, the *ISS*, and the *Gemeinschaftsschule*, all the way to the *Abitur* (baccalaureate).

Go4Goal!

The TOEFL Junior Standard test is currently used in Berlin within Go4Goal!, an English-language competition for secondary-level EFL learners offered by Go4Goal! e.V. To participate in the contest, secondary-level schools in Germany, Austria, and Switzerland can sign up their EFL classes (Grades 7–10). All students in a participating class who have received parental consent then take the TOEFL Junior Standard test. Classes compete against other classes at the same grade level; that is, all eighth-grade classes compete against one another, all ninth-grade classes compete against one another, and so on. Classes as well as individual learners who achieve the highest scores within their grade level win small awards. Most schools participating in Berlin are *Gymnasien*; however, some *Gemeinschaftsschulen* are among the contestants as well.

Given the steady growth of the contest in recent years, schools have been inquiring if the TOEFL Junior Standard test could be used beyond the context of the contest. In particular, they have highlighted a potential fit of the standardized assessment as a substitute for one of the four mandatory assessments per term (i.e., *Lernersatzleistung für eine Klassenarbeit* or *Klassenarbeitsersatz*).

Research Questions

The superordinate question that motivated the alignment study was whether the TOEFL Junior Standard test constitutes a suitable *Klassenarbeitsersatz* for secondary-level EFL learners in Berlin. Guided by this overarching question, the study aims to respond systematically to the following research questions:

1. To what extent does the TOEFL Junior Standard test align with the competences and skills promoted in Berlin's secondary-level EFL curricula?

2. What are teachers' opinions of the TOEFL Junior Standard test as an assessment for their EFL learners?

Methodology

Evidence for the investigation of the alignment was collected in two parts. First, the researcher investigated the extent to which the skills and competences assessed by the TOEFL Junior Standard test were also covered in Berlin's EFL curriculum. Second, teachers' opinions

were collected to triangulate the quantitative estimates and gauge in more detail the suitability of TOEFL Junior Standard test as an assessment for secondary-level EFL learners in Berlin.

Content Coverage

This first part of the alignment study focused on the content coverage and overlap between the EFL curriculum in Berlin and the TOEFL Junior Standard test. To that end, curricular documents and materials used for instruction were sampled and investigated, as follows (see Appendix A).

Evaluation materials. In a first step, the researcher reviewed the *Rahmenlehrpläne für die erste Fremdsprache* English (i.e., the curricula or frameworks for the first foreign language taught), including the following documents:

- *Rahmenlehrplan 1–10 kompakt—Themen und Inhalte des Berliner Unterrichts im Überblick*

- *Teil A: Bildung und Erziehung in den Jahrgangsstufen 1–10*

- *Teil B: Fachübergreifende Kompetenzentwicklung*

- *Teil C: Moderne Fremdsprachen Jahrgangsstufen 1–10*

Given that these documents constitute the general framework for EFL instruction, they provide broad, general guidelines rather than the detailed content that is being taught in the EFL classrooms. It is explicitly stated in the framework/curriculum that the "Beitrag der Fächer zur Förderung von Sprach- und Medienkompetenz wird im Rahmen des *schulinternen Curriculums* abgestimmt" [the contribution of the subjects to promote learners' language and media competence is determined by school-internal curricula]; Part C, 2015c, p. 13). Although school-internal curricula tend to be more fine grained, they usually do not reflect in detail the tasks and activities conducted in the actual EFL classrooms. In other words, *Rahmenlehrpläne* and curricula are too broad to serve as points of reference for a systematic alignment analysis.

The theoretical underpinnings as well as the pedagogical frameworks, principles, and content requirements put forth in the curricula are usually operationalized in the textbooks (*Lehrbücher*) that up to this day constitute the core medium of EFL education in Germany (Finkbeiner, 1999; Haß, 2016). In fact, school-internal curricula tend to be strongly informed if not entirely based on textbook contents (see Appendix B for an example of a school-internal

curriculum provided by a *Gymnasium* that participated in this alignment study). Therefore we chose *English G Access*, one of the main textbook series in lower secondary EFL education in Berlin, as the point of reference for the alignment. Published by Cornelsen, the following four textbooks from the series were used, as they corresponded to the four grade levels of lower secondary EFL education: *English G Access 3* for Grade 7, *English G Access 4* for Grade 8, *English G Access 5* for Grade 9, and *Context Starter* for Grade 10.

Sample units. To obtain an in-depth understanding of how the content of Grades 7, 8, 9, and 10 curricula aligns with the TOEFL Junior Standard test, we randomly selected—following Hsieh (2015)—40% to 50% of the units from the four textbooks for the alignment analyses. Table 3 presents the total of eight units, two from each textbook, that were used in the alignment. All units are structured similarly. Each unit centers around a general topic. For example, Unit 4 ("Faces of South Dakota") in *English G Access 4* includes topics such as school life, school system, sport, family and friends, and Native Americans—all of which are wrapped into stories that take place in South Dakota. Units usually begin with warm-up activities, followed by reading passages (e.g., magazine articles, websites, reports, excerpts from a novel, poems), practice activities, and study tips. Moreover, they tend to have a particular focus with regard to enabling skills and language phenomena. For instance, Unit 4 ("Faces of South Dakota") in *English G Access 4* has an emphasis on vocabulary, grammar, and pronunciation.

Table 3. Textbooks and Units Used in Alignment

Textbook	Grade level	Total no. textbook units	Units included in alignment	Percentage of content used in alignment
English G Access 3	7	5	1_This Is London 3_A Weekend in Liverpool	40
English G Access 4	8	5	2_New Orleans 4_Faces of South Dakota	40
English G Access 5	9	4	2_Relationships 3_Big Dreams, Small Steps	50
Context Starter	10	4	2_Communicating in the Digital Age 4_Going Places	50

Alignment procedures. To examine the alignment, each activity in the selected textbook units was reviewed and, if applicable, matched with the corresponding domain, skill, and task type from the TOEFL Junior Standard test. For example, in the practice section (Part B) of Unit 2, about New Orleans, students are asked to "read a story about something that went wrong and tell [their] partner about it" (Rademacher, 2016, p. 43).

As shown in Figure 3, this activity was assigned the domain "academic" as well as the task type (or skill) to which it corresponds (i.e., reading, or, in more detail, reading for the main idea, i.e., R_Main idea). Additionally, the activity requires students to report orally the content of the story to a partner. Therefore, it received a second code under the heading "other task types," which specifies the skill (speaking) and function (summarizing) that is to be performed. Derived from the test blueprint as well as So et al. (2015), Appendix C provides an overview of all coding criteria, including the domains, skills, and TOEFL Junior task types that were used as alignment criteria and thus as codes in the coding process.

Unit	Unit part	Unit component	Activity	Domain	TOEFL Junior task type and skill	Other task types
2_New Orleans	Part B Practice	5_Something went wrong	Read a story and tell your partner about it.	Academic	R_Main idea	S_summarizing

Figure 3. Coding sample.

Following Hsieh (2015), the alignment coding was performed by the author and a trained assistant from LTS in Berlin. The researcher coded all activities included in the eight different units shown in Table 3, while the LTS representative coded all activities in the units selected from *English G Access 3* and *4* (i.e., 384 language-related activities). In other words, to ensure a reliable coding procedure, 50% of the content was double coded. Before individually coding the textbook activities, the researcher provided the LTS representative with background knowledge about the TOEFL Junior construct and trained the representative how to apply the different codes in two 1-hour training sessions. At first, the obtained simple agreement between the codings was 83%, including the main and supplemental codes as well as the TLU domains. A closer inspection of the discrepancies revealed that the two coders differed with regard to the application of two particular codes: the LFM codes, both LFM_LM and LFM_LF, as well as the code "listening for detail." To adjudicate, a TOEFL Junior specialist from ETS's assessment development office was contacted to review each discrepancy and determine the appropriate code. Upon completion of the coding process, frequency counts and percentages were calculated for each alignment criterion per level.

Alignment results. As a first step to obtain insights into the alignment, I tallied—within the portion of activities coded—the number of language-related activities per grade level. For example, as shown in Table 4, grade level 7 (*English G Access 3*, Units 1 and 3) contained 184 language-related activities.[8] Additionally, I calculated the domain coverage per grade level. As presented in Table 4, the language-related activities coded for each level represented all three domains—academic, social/interpersonal, and navigational—with the academic domain being the most frequent across all levels.

Table 4. Domain Coverage Across Levels for Textbooks

Grade level	Total no. language-related activities coded, *n* (%)	Academic, *n* (%)	Social/interpersonal, *n* (%)	Navigational, *n* (%)
7	184 (100)	115 (62.5)	57 (30.98)	12 (6.52)
8	200 (100)	124 (62)	70 (35)	6 (3)
9	233 (100)	135 (57.94)	97 (41.63)	1 (0.43)
10	234 (100)	215 (91.88)	15 (6.41)	4 (1.71)

Second, I tallied the main skills coded for each activity across the selected units at the four levels. As presented in Table 5, all grade levels included activities that address the four skills of reading, listening, speaking, and writing as well as enabling skills of grammar and vocabulary knowledge (LFM). Overall, the receptive and enabling skills—listening, reading, and LFM—that constitute the focus of the TOEFL Junior Standard test (shaded in Table 5) are addressed by nearly 50% of the coded activities at each grade level.

Table 5. Number of Language-Related Activities per Grade Level for Textbooks

Grade level	Total no. language-related activities, *n* (%)	Listening, *n* (%)	Reading, *n* (%)	LFM, *n* (%)	Speaking, *n* (%)	Writing, *n* (%)
7	184 (100)	23 (12.5)	40 (21.74)	28 (15.22)	61 (33.15)	32 (17.39)
8	200 (100)	24 (12)	39 (19.5)	36 (18)	67 (33.5)	34 (17)
9	233 (100)	15 (6.44)	45 (19.31)	46 (19.74)	84 (36.05)	43 (18.45)
10	234 (100)	27 (11.54)	54 (23.08)	22 (9.4)	72 (30.77)	59 (25.21)

Note. Shaded cells show the receptive and enabling skills of the TOEFL Junior Standard test.

As a third and more fine-grained step, I calculated the frequencies of the subskills underlying the three language skills also included in the TOEFL Junior Standard test: listening, reading, and LFM. Table 6 summarizes the coverage of task types across skills and grade levels. All skills and subskills were represented by at least one activity across the four grades, except for making predictions when listening to aural texts and identifying a pronoun referent in a reading

passage. Although grammar- and vocabulary-oriented activities feature prominently across all levels, the two most frequently represented task types targeting listening and reading are skimming and scanning, in terms of reading strategies, as well as listening for gist and listening for specific information—activities that also feature prominently on the TOEFL Junior Standard test (So et al., 2015).

Table 6. Coverage of TOEFL Junior Task Type per Textbook Grade Level

Task type	7	8	9	10
Listening	23	24	15	27
L_Main idea	6	2	5	9
L_Detail	15	19	9	14
L_Speaker's purpose	1	0	0	0
L_Rhetorical device	1	0	0	2
L_Prediction	0	0	0	0
L_Inference	0	1	0	1
L_Prosody	0	2	1	1
Reading	40	39	45	54
R_Main idea	22	21	13	19
R_Factual information	12	9	20	27
R_Vocabulary in context	2	3	3	2
R_Inference	3	4	7	1
R_Author's purpose/Rhetorical structure	1	2	2	5
R_Pronoun referent	0	0	0	0
Language form and meaning	28	36	46	22
LFM_LM_Can demonstrate knowledge of vocabulary meaning, in academic and nonacademic texts of varying complexity	12	12	20	10
LFM_LF_Can identify the correct, meaningful use in context of the following grammatical structures	16	24	26	12

Note. The data in row R_Vocabulary in context includes both figurative and idiomatic language.

In general, the TOEFL Junior Standard test appears to cover and thus align very well with the receptive competences in the EFL curriculum in Berlin as operationalized via the textbooks—an observation that we aimed to triangulate by conducting interviews and/or focus groups with teachers in Berlin to gauge their perspectives about whether the TOEFL Junior was a fitting assessment for their students.

Teacher Focus Group Interviews

Given that teachers implement the activities in their EFL classrooms, we also conducted interviews and/or focus groups with teachers at two different school types in Berlin to explore further whether instructors considered the TOEFL Junior Standard test a suitable assessment for their learners.

Participants. A focus group and an interview were conducted with a total of four teachers at two schools, a *Gymnasium* (*n* = 3) and a *Gemeinschaftsschule*[9] (*n* = 1). All teachers had participated with their classes in the Go4Goal! competition for several years prior to the interviews. Also, they had participated in a continuous education seminar (*Lehrerfortbildung*) that had focused on standardized assessment in general and the TOEFL Junior in particular. Thus they were largely familiar with the TOEFL Junior in terms of its structure and administration. Moreover, three of the four teachers also used the same textbook series evaluated for Research Question 1.

Procedures and analysis. The participating teachers were recruited by LTS. The researcher met with the teachers in Berlin in March 2018 to conduct the focus group and the interview at the local schools. During both meetings, the researcher followed the same semistructured interview approach featured in the focus group protocol (see Appendix D). Thus they provided more detailed information about four distinct aspects: (a) how they utilized the TOEFL Junior Standard test at their respective schools, (b) the impact they saw on students and teachers, (c) their perception of the test's alignment with the curriculum, and (d) its potential use as a *Klassenarbeitsersatz*.

Both the focus group and the interview were carried out in German, audio recorded, and then transcribed verbatim. Aiming for triangulation of the textbook coding results, the transcripts were then analyzed qualitatively to identify main themes and obtain more detailed insights into teachers' perceptions about the fit of TOEFL Junior for their respective learner groups.

Results

In the following pages, teachers' opinions and perceptions will be depicted following the four main themes discussed in the focus groups: (a) utilization of the TOEFL Junior Standard test at respective schools, (b) perceived impact of the test on students and teachers, (c) perceived alignment of the test with curriculum and instruction, and (d) the TOEFL Junior Standard test's potential use as a *Klassenarbeitsersatz*.

How the TOEFL Junior Standard Test Is Used at the Schools

With regard to test use, the two schools started participating in the Go4Goal! contest in 2014 and 2015, respectively. Since then, the number of students and classes participating in the contest has grown considerably. For example, the *Gymnasium* began participating with two

classes in 2014. Back then, participation among students was voluntary. With annually increasing numbers of students wanting to participate, the school has made participation mandatory for all upper secondary-level classes (i.e., all students in Grades 11 and 12 participate every year). Additionally, participation was recently expanded to lower grades. A similar course and growth of the Go4Goal! participation was described by the teacher from the *Gemeinschaftsschule*. However, she noted that the school was unable to make the assessment mandatory for all students because of the cost of the assessment and the socioeconomic background of her learners. In sum, the *Gymnasium* administers the TOEFL Junior to all levels (Grades 7–12), whereas the *Gemeinschaftsschule* participates with Grades 9 and 10.

In terms of test administration, both the *Gymnasium* and the *Gemeinschaftsschule* incorporate the test into the students' daily routine; that is, they schedule the administration of the test in the morning during regular English classes. Learners usually take the test in the school's largest auditorium (*Aula*). If there are too many students, such as in the case of the *Gymnasium* now, teachers administer the test per cohort or grade level (i.e., all seventh graders, all ninth graders, etc.). Both schools emphasized that they tended to schedule test administrations for younger students earlier in the morning (e.g., between 8:00 and 10:00 a.m.), while older students take it in the later morning hours (10:00 a.m. to 12:00 p.m.). On the day of the test administration, teachers receive the test booklets in a sealed envelope and then administer the assessment to their learners. Hence, schools and teachers have been incorporating the TOEFL Junior as a standardized assessment into their learners' regular school day.

Perceived Impact on Students and Teachers

At both schools, teachers reported that the TOEFL Junior Standard test appeared to have a motivational effect on their learners, which also boosted their ambition and competitiveness with regard to learning English. For example, one of the *Gymnasium* teachers noted that "die Schüler freuen sich eigentlich immer darauf, wenn TOEFL angesagt ist" [the students are always looking forward to when TOEFL is on the agenda] (ID01). Her colleague added that

> da wird dann tatsächlich verglichen, wie viel Punkte hast du, wie viel Punkte habe ich
> und man guckt dann auch hinten auf den Bereich, ist es jetzt B1, ist es B2 oder bin ich
> eben doch noch im A2-Bereich. Und die Schüler wissen in der Regel schon, wie sie mit
> so einem Score report umzugehen haben beziehungsweise entnehmen da gewisse

Informationen. [they really do compare, how many points do you have and how many points do I have and then they check the information on the back if it's a B1 or a B2 or am I still at the A2 level. And the students in general, they do know how to handle such a score report and they know how to extract certain information.] (ID02)

Along similar lines, the teacher from the *Gemeinschaftsschule* highlighted her students' increased ambition, commenting that

sie sind auch gespannt und sie finden sich dann auch so ein bisschen besser als so die anderen. Ach, ich habe dran teilgenommen und ach, guck mal und ich habe die B1 geschafft und du hast nur die A2 . . . da entwickeln sie dann doch so ein bisschen Ehrgeiz" [they are really excited and they find themselves a little better than the others like Ha! I participated in this and look, I made the B1 level and you only have an A2 . . . so they do get a little competitive]. (ID04)

Hence, the TOEFL Junior Standard test seems to provide a level of comparability that appears to be viewed as positive among EFL learners.

Instructors identified two particular aspects that they regarded as potentially underlying the increased motivation and ambition they observed in their students. First, all teachers highlighted the opportunity to monitor one's own progress across time as a particularly beneficial aspect. Although the TOEFL Junior Standard test is currently only used in Grades 9 and 10 at the *Gemeinschaftsschule*, teacher ID04 used this particular point to argue for expanding the use of TOEFL Junior across all levels from Grades 7 to 10. She regarded an annual administration of the assessment as useful:

Weil man dann dieses Entwicklungspotenzial jetzt nicht nur neunte, zehnte hat, also wenn man in der siebten schon anfangen würde. Und dann sehen die eben, ich habe, was weiß ich, von A1 bin ich auf Niveau B2, das ist natürlich viel motivierender, als wenn die jedes Jahr eine Note kriegen, die zwischen drei und vier liegt, wo man ja nicht tatsächlich ein Entwicklungspotenzial erkennen kann mit den Noten. [Because you don't just have the progression in 9th or 10th grade if you could start in 7th. And then they could see ah I made it from an A1 to a B2 level; that is much more motivating than just getting your regular English mark or grade every year and you find yourself in-between 3 and 4 each year. That mark system doesn't really show you any potential progress.] (ID04)

In addition to monitoring their own progress, teachers emphasized the practical benefits of the TOEFL Junior Standard test beyond the actual classroom. Teacher ID03 noted that students liked the TOEFL Junior because they can use the certificate they receive to apply for internships or even some programs at colleges and universities. Teacher ID04's view confirmed the real-life benefits, arguing that the TOEFL Junior Standard test "ist ein Test, mit dem [man] im Leben etwas anfangen [kann]" [a test that you can use in real life]. Along those lines, *Gymnasium* teacher ID01 shared anecdotal evidence that in 2016 she had a student who was selected for a highly prestigious 1-year internship in a politically oriented nonprofit organization because of her high scores on the TOEFL Junior Standard test she had taken in the context of Go4Goal! Hence the TOEFL Junior seems to have positive effects on students' motivation and self-determination, which according to teachers participating in this study, may be due to the increased transparency as well as the assessment's use beyond the classroom—factors that, as they argued, have contributed to students taking a more active role in their English language learning.

In addition to positive effects the TOEFL Junior Standard test seems to have on learners, teachers identified three benefits for themselves as well. First, they argued that a standardized English assessment like the TOEFL Junior Standard test provided comparability not only at the class level but moreover across schools. Teacher ID04, for example, highlighted the following: "Vergleichbarkeit. Finde ich auch ganz wichtig, dass man auch so im Standard der anderen Schulen auch sieht, wo man steht, weil wir ja unsere Arbeiten eigentlich selber entwerfen" [Comparability. I find that very important that you can see yourself in comparison to other schools to see where we stand because we usually design and come up with the written classroom assessments ourselves.] In addition to comparability, teachers emphasized the quality of tasks included in the TOEFL Junior, in particular the listening comprehension tasks. They argued that it was next to impossible for individual teachers to design these types of tasks for a regular classroom assessment. *Gymnasium* teacher ID03, for example, spoke in more detail to this point:

> Der Umfang der Aufgaben, die da abgefragt sind, das kann man als Fachlehrer so für sein Lehrwerk gar nicht erstellen. Das geht ja soweit in die Breite. Und gerade was die Hörtexte betrifft, ich finde das immer sehr schön, dass da auch verschiedene Akzente und Dialekte zu hören sind, dass man die Schüler wirklich eben auch daran gewöhnt, dass

nicht Oxford-Englisch gesprochen wird. Und die kommen mittlerweile eigentlich ganz gut damit zurecht. In einer Klassenarbeit, da habe ich mein Material, meinen Klassenarbeitstrainer, wo ich dann entsprechende Texte auswählen kann, aber dann habe ich drei. Ja und da habe ich dann einen Sprecher und der spricht einen Akzent oder eben keinen Akzent. Und dadurch ist die Breite der Sprache doch sehr eingeschränkt. Und das finde ich hier bei TOEFL wunderbar. [The scope and breadth of the tasks that are included in the assessment. You can't develop something like that as a regular classroom teacher. That is so comprehensive. And especially the listening comprehension texts, I really like that they feature various accents and dialects so the students are getting used to listening to different accents and not only Oxford English. And they are really fine with it by now. In a written classroom assessment, I have my materials, my "classroom assessment coach"[10] from which I can select certain texts, but then I only have three. Yes, and then I have one speaker and he speaks either one or no accent. So the breadth of the language is quite limited. So I like that (the breadth of it). I find that wonderful in TOEFL.]

In addition to comparability and the wider construct that can be assessed by means of the TOEFL Junior Standard test, teachers also outlined a more practical impact the test would have, if officially implemented, on their line of work: saving time. Teachers are provided with total and section scores at individual and class levels. Hence they do not have to grade their students' performances and can therefore focus on other job-related duties such as preparing their classes in more detail—a point made by all teachers at both schools. For instance, *Gymnasium* teacher ID02 commented that "bei der Auszeichnungsveranstaltung war das immer bisher so, dass die Schüler das sofort dann in die Hand bekommen haben, und demzufolge hatten wir die Auswertung nicht" [at the award ceremony it was always the case so far that students got this right away so therefore we did not have to handle the grading] (ID02). The teacher at the *Gemeinschaftsschule* commented along the same lines, quantifying the time she would save as "vier, fünf Stunden Ersparnis, aber mindestens, ja, Arbeitsersparnis" [four, five hours of time saved at least, yes saving of work] (ID04).

Perceived Correspondence Between the TOEFL Junior Standard Test and the English as a Foreign Language Curriculum

Overall, all teachers were quite positive about the test and regarded it as a fitting assessment that measures the receptive skills, strategies, and abilities they promote in their EFL classes. Given that teachers do not get to see the actual items (they were familiar with the practice materials on the ETS website, however), they highlighted that they based their judgment primarily on two aspects: (a) the audio texts used in the listening comprehension section, which they listen to during test administration, and (b) the observed relationship between test scores and students' performance in class. For example, when asked how she would evaluate the fit of the TOEFL Junior Standard test for the assessment of her students' language abilities, one teacher at the *Gymnasium* argued as follows:

> Ja, das kann ich ja nur anhand der Hörtexte beurteilen. Und da, muss ich sagen, gibt es nicht immer Überschneidungen. Das finde ich aber eigentlich eher positiv. Denn das, was die Lehrbuchinhalte betrifft, das testen wir ja im Unterricht sozusagen. Und die Kompetenzen, die sie erwerben sollen, die sind ja von diesen Inhalten unabhängig. Und deshalb finde ich das eigentlich ganz gut, dass das so Querbeet geht. Und das ist ja bei VERA 8 und bei den MSA-Prüfungen genau das Gleiche. Und das ist für mich eher ein positiver Aspekt. [Yes, well, I can only judge that by the listening comprehension texts. And I gotta say they don't always match [[in terms of topics]]. But I find that rather positive because we assess the topics in the textbook in class so to speak. And the competences they are supposed to acquire, they are independent of the topics anyways. And that's why I actually like that because it's all over the map. And that's the same with VERA 8 and the MSA examinations. So that for me is a rather positive aspect.] (ID01)

They pointed out discrepancies with regard to the topics covered but highlighted a match between the skills, strategies, and competences that are promoted in EFL education—an observation that underscores the actual design and construct of the TOEFL Junior.

In addition to highlighting TOEFL Junior Standard test's fit as a measure for the competences, students are supposed to acquire and learn in EFL classes, teachers noted the strong relationship between the test scores and learners' in-class performances. The following comments were made by three of the four participating teachers to that effect:

Also die, die jetzt gut im Unterricht sind, die haben auch bei dem Test mindestens eine B2 gehabt. [So those who are good in class they also test at least B2.] (ID04)

Ja, aber im Endeffekt deckt sich das so mit den Beobachtungen, die man im Unterricht macht. Also man weiß schon, wer sind die sehr guten Schüler, die kriegen den höchsten Score. Und die man als schlecht einteilt oder einschätzt, die kriegen dann halt relativ wenig Punkte. Also das ist schon so, wie man eigentlich das erwartet. [Yes, at the end it matches with the observations in class. So you know exactly who the very good students are and they get the highest score. And those that you would evaluate or judge as rather bad, those tend to get fewer points. So it is exactly how you would expect it.] (ID02)

Also gerade bei dem Leistungskurs, da vergleiche ich immer, was ist meine Benotung und was ist da rausgekommen. Und es war jetzt zumindest zweimal, alle zwei haben jetzt einen Leistungskurs gehabt und zweimal hat sich das fast eins zu eins mit den Einsen gedeckt, die auch wirklich dann im C1-Bereich schon waren. Und, also ne, also durchaus. [So especially in the advanced courses, there I always compare what is my grade and what did the test say. And it has been a match twice so far, I had two advanced classes and both times it was almost a 100% match with the straight A students, those that really were at the C1 level already. So, yes, definitely.] (ID01)

In sum, they highlight a fairly strong apparent correspondence between the students' achievements in their English classes and the scores students obtain on the TOEFL Junior Standard test.

Although, overall, teachers were very positive about the TOEFL Junior Standard test and did not see any disadvantages in terms of using the test with their learners, they identified two areas that may require additional investigation and conceptualization before the TOEFL Junior Standard test can be implemented as a regular classroom assessment (i.e., *Klassenarbeitsersatz*). First, they pointed out that it would be helpful if teachers could obtain more detailed diagnostic information with regard to the weaknesses of a given individual learner (i.e., what types of items the learner got incorrect). According to the instructors, more feedback and diagnostic information would be relevant to provide required transparency vis-à-vis the students' parents or legal guardians. Moreover, it would be helpful because learners actively seek feedback to improve their abilities and scores. For example, a teacher at the *Gymnasium* commented that "es

gibt immer wieder Schüler, die geben sich damit zufrieden zu wissen, was sie am Ende erreicht haben, aber es gibt eben auch viele ehrgeizige Schüler, die wollen sich verbessern und die möchten dann eben gerne wissen, wo habe ich jetzt hier einen Fehler gemacht" [there are always students who are OK with only knowing what they achieved, but then there are many others who are more ambitious, who want to improve and those want to know where they made mistakes] (ID01).

Another *Gymnasium* teacher recalled the following episode:

Was ich von meinen Schülern gefragt wurde, war zum Beispiel mal, ob sie denn eine Auswertung bekommen könnten, wo sie was falsch gemacht haben. Also im Endeffekt erfährt man ja nur, man hat so und so viele Punkte und für die Schüler wäre es zumindest punktuell interessant, wo habe ich denn einen Fehler gemacht. Und da habe ich gesagt: "Das ist natürlich schwierig für uns, weil wir die Tests selber ja gar nicht kennen. Und für die Auswertung wir ja eigentlich auch keine anderen Informationen haben, als das, was uns dann in der Übersicht dargestellt wird." Ob man das machen könnte, ist vielleicht noch eine Frage. [What I was asked by my students was, for example, if they could get an analysis or more detailed evaluation where they got something wrong. So, basically, you only get to know that you have that many points and for the students it would be at least in some instances be of interest where did I make a mistake. And so I said: "That is obviously quite difficult for us to say because we do not know the exact test [[form]]. And for the evaluation, we do not have more information than what is printed on the back." If you could do that (i.e., include more information vis-à-vis the students' weaknesses or mistakes), that is maybe a question.] (ID02)

A teacher at the *Gemeinschaftsschule* shared her personal solution to the challenge. Once students receive their score reports, she holds individual consulting sessions with each test taker. In personal conversations, she then goes over their score reports and points out potential weaknesses based on the section scores and "can do" statements presented on the students' score reports. She recounted this process as follows:

Mache ich mit jedem einzelnen ein persönliches Gespräch und werte mit ihnen das aus und gucke, wo, aber eben eher auf der allgemeinen Basis. Also guck mal da, in Listening, da hast du so und so viele Punkte und bei dem Reading hast du so und so viele Punkte.

Im letzten Jahr, also ich versuche immer die zu überzeugen, die in der neunten Klasse dran teilgenommen haben, in der zehnten auch nochmal, damit man so sehen kann, habe ich einen Entwicklungsfortschritt gemacht oder nicht. . . . Sprechen wir eben Möglichkeiten durch, was sie selbstständig auch tun können zuhause zum Beispiel, um auf dieses Niveau zu kommen. [I schedule individual appointments with every student to go over the score report and then we see in general. So look, in listening you got that many points and in reading you got that many points. Last year—I always try to convince those that have participated in 9th grade to participate again in 10th grade so you can see if there was development and progress or not. . . . We talk about possible options and opportunities they can do at home, for example, to get to the next level.] (ID04)

Additionally, teachers pointed out that more fine-grained diagnostic information—both at the class and the individual level—would be highly beneficial as it could be used to inform the curriculum and concrete classroom instruction. For example, teacher ID04 noted the following: "Ich sehe ja jetzt auch nicht, wo jeder Schüler seinen Fehler gemacht hat, um daraus so ein individuelles Curriculum zu machen" [I cannot see where each student made mistakes in order to really create a curriculum based on the test results] (ID04). In addition to acknowledging the potential constraints of making test items publicly available, teachers also proposed a potential solution to foster each individual's English-language competency based on the test results. Teacher ID04, for instance, proposed the following:

Also wenn dann Übungen mitkommen würden, dass man sagt also: "Der hat noch mehr Schwächen bei dem Teil des Lesens, beim Skimming oder Scanning oder was weiß ich," dass man dann gleich auch eine Vorlage, also Übungsbeispiele vielleicht bekommt, die ich ihm dann gleich mitgeben könnte und sagen, dann bringt es auch Sinn . . . wenn jeder Schüler auch so sein Feedback bekommt, so dass wir da gleich so ein Übungsmaterial auch zusammenstellen können, woran die selber dann individuell arbeiten können. . . . So ein paar Vorschläge und ich brauche das nur ausdrucken und dem einheften. [Yes, if you could provide exercises so that I could say, "He has weaknesses in reading, skimming or scanning or whatnot" that you have a template, that is exercises, activities, and materials that I could give the students after the test so they can practice. Then it would make sense . . . if every student could get individualized feedback so we would have exercises and

materials that we could package for each learner and then they could individually work

on those exercises. . . . Just a few suggestions and all I need to do is print it out and put it

in a folder for each learner.]

They argued it would be helpful to obtain activities linked to particular score levels that students could then work on individually in class with the goal of improving specific competences.

In addition to wanting more feedback and diagnostic information, teachers pointed out the issue of setting cut scores and/or providing concordances to translate the scores into school grades. They argued that it would need to be clarified (a) whether certain cut scores would be set at the state level that would then be made mandatory for each school form or (b) whether individual schools and teachers were tasked with translating the scores into grades themselves.

Discussion and Conclusion

The overall goal of this study was to examine the alignment (i.e., match) between the TOEFL Junior Standard test content and the secondary-level EFL curriculum in Berlin, Germany. Therefore, curricula were reviewed and 40% to 50% of the activities included in four different textbooks for Grades 7, 8, 9, and 10 were coded systematically. Additionally, the results from the coding exercise were triangulated with teacher perceptions of the TOEFL Junior and its fit as an assessment for their EFL learners. Thus, we aimed to evaluate the match and correspondence between (a) what the test measures and (b) what is being taught in secondary-level EFL classes in Berlin to explore in a first attempt whether it would be appropriate to consider the TOEFL Junior Standard test as a potential candidate for a *Klassenarbeitsersatz*.

The results from the systematic coding suggest that the TOEFL Junior Standard test is an appropriate measure for secondary-level EFL learners in Berlin. The domain coverage, shown in Table 4, indicates that the majority of tasks and activities in EFL education are rooted in the academic domain, followed by tasks in the social–interpersonal and navigational domains. Similarly, most items on the TOEFL Junior represent the academic domain, with fewer numbers of items in the social–interpersonal and navigational domains (So et al., 2015). In short, the academic domain features most prominently in the EFL curriculum as well as on the TOEFL Junior.

In addition to domain coverage, the TOEFL Junior Standard test assesses to a large extent the receptive skills and competences included in functional communicative competence, one of

the key components of EFL instruction in Berlin (Senatsverwaltung für Bildung, Jugend und Wissenschaft, Berlin Ministerium für Bildung, Jugend und Sport des Landes Brandenburg, 2015a, p. 5–7). The curriculum states explicitly that "Voraussetzung für die Realisierung der einzelnen Kompetenzen ist das Verfügen über angemessene sprachliche Mittel und kommunikative Strategien" [a necessary condition for the realization of individual competences is the disposition of appropriate linguistic resources and communicative strategies] (Senatsverwaltung für Bildung, Jugend und Wissenschaft, Berlin Ministerium für Bildung, Jugend und Sport des Landes Brandenburg, 2015b, p. 10). The TOEFL Junior covers receptive skills and subskills. For instance, as shown in Table 5, almost 50% of all language-related activities in the textbooks represent skills and competences assessed in one of the three sections of the TOEFL Junior Standard test. Hence, the TOEFL Junior Standard test covers receptive skills, reading and listening, and its subskills (see Tables 5 and 6) as well as a considerable portion of enabling skills—all of which are represented in the EFL tasks and curriculum.

By contrast, productive skills, speaking and writing, which constitute the remaining 50% of activities, are not directly measured by the TOEFL Junior Standard test. However, the fact that the TOEFL Junior was not designed to measure productive skills directly may not necessarily disqualify the assessment from being a good candidate for a *Klassenarbeitsersatz* for the following reasons. First, writing is evaluated in a kind of receptive format; that is, critical skills needed to produce good writing in English are covered in the language form and meaning section in the form of editing tasks. In fact, two out of five LFM items on each test form are editing tasks that require students to read a piece of writing and then identify and correct errors. Hence, while the test certainly does not cover the entire construct of writing, it assesses editing, a key part of the writing process. If the TOEFL Junior Standard test were to be used as a *Klassenarbeitsersatz*, it would only constitute one out of four mandatory assessments per term. Thus, if adopted, it could complement the three additional assessments, which could then feature a more prominent emphasis on assessing writing. Second, speaking is not covered by the TOEFL Junior.[11] However, speaking as a skill tends not to be assessed in the context of traditional written classroom assessments. Instead, speaking skills are usually evaluated by the teacher in the context of students' oral contributions during regular, everyday class participation. Hence, although it does not cover the entire construct of productive skills of speaking and writing, the TOEFL Junior Standard test aligns very well with the EFL curriculum in Berlin insofar as it

assesses the receptive and enabling skills needed for functional communicative competence in English—an alignment that teachers highlighted in their perceptions as well.

In addition to attesting to the alignment, teachers raised two aspects in need of further discussion. First, they highlighted the benefits of additional diagnostic information and learning materials that would allow them to use the test scores to inform their instruction and foster students' skills in more individualized ways—both aspects that are currently being investigated by ETS. Moreover, they emphasized that they would most likely be required to give students marks for their respective performances on the TOEFL Junior Standard test. Papageorgiou (2016) pointed out that standard setting is

> a critical component of any alignment process if test scores are to be used to classify test takers into different proficiency levels, [adding that] even if an assessment provides adequate content coverage in relation to a standard [or in this case, a curriculum], setting cut score too low or too high will have important unintended consequences. (p. 335)

Hence, a study needs to be conducted to empirically establish concrete and justifiable cut scores in line with the German marking system (from 1 *sehr gut* [very good] to 6 *ungenügend* [insufficient/fail]) for each grade level.

A mapping study is inevitable, and Table 7 provides a first point of reference. It juxtaposes the information from two sources: (a) the Common European Framework of Reference for Languages (CEFR) categorization for each grade level, outlined in the EFL curricula, and (b) the TOEFL Junior Standard test scores as they were mapped to the CEFR levels (see Papageorgiou & Baron, 2017).

Table 7. Juxtaposition of Grade Level, Common European Framework of Reference for Languages (CEFR) Categorization, and TOEFL Junior Standard Test Score

Grade level (*Schuljahr*)	Alphabetical level (*Niveaustufe*)	CEFR level (*GER*)	TOEFL Junior total score
10	H	B1+	795–845
9	G	B1	745–790
8	F	A2	645–740
7	E	A1+	<640

Note. No score mapping study has been conducted yet, although one is necessary to directly link the *Niveaustufen* with TOEFL Junior total scores.

In the context of future research, a score mapping study would need to be conducted to link the *Niveaustufen* outlined in Berlin's EFL curricula and the TOEFL Junior Standard test's

total and section scores. Additionally, standard-setting workshops with a representative group of teachers from each school type in Berlin would most likely need to be conducted to establish concrete cut scores that may provide empirical backing to matching particular scores with respective marks from 1 (*very good*) to 6 (*insufficient*). Thus concrete marks provided for specific performances would be based on rigorous empirical investigation.

Finally, some limitations of the current alignment study need to be addressed. First, the sample of participating teachers was very small, and a bigger, more representative sample of EFL teachers from all institutions that offer secondary-level EFL instruction may provide a more comprehensive picture of the fit of the TOEFL Junior Standard test for all types of secondary-level learners in Berlin. Second, the activities used for the alignment may not reflect the total scope of activities used in secondary EFL education in Berlin. Owing to practical constraints, only a selection of the activities from a specific textbook series (*English G Access*) was coded. Although it is one of the most widely used textbook series in EFL education in Berlin, no generalizations can be made with regard to coverage in other textbooks. On the other hand, teachers tend to supplement textbooks with additional materials and activities, so, for example, the lack of coverage regarding pronoun referent in reading and the prediction in listening found in this study may very well be covered in the regular classroom.

Moreover, the study did not account for topic alignment. The reason is that, as a large-scale English proficiency test, the TOEFL Junior Standard test was designed to be curriculum independent. Thus it does not include topics specific to a given curriculum but instead measures competences, skills, and abilities in English as a second/foreign language. Given that the TOEFL Junior Standard test is used by many different English learners around the word, the topics included in the assessment are carefully selected to avoid bias, cultural or otherwise. In line with the teachers' perception, the researcher observed that the topics covered in the TOEFL Junior Standard test are to a large extent different from the topics in the textbook units. For example, some units included texts about dating in high school or pictures of religious signs, topics that would not be included in the TOEFL Junior. Hence the coding focused on the alignment of the strategies, competences, and language skills needed for English-language ability.

To summarize, the study found a fairly strong alignment between the secondary-level EFL curriculum in Berlin and the TOEFL Junior Standard test. To a large extent, the competences, strategies, and skills that are measured by the test match the ones promoted in EFL

education in Berlin. Also, teachers were positive toward the assessment and highlighted its fit as a *Klassenarbeitsersatz*. In particular, they identified ninth grade as a potentially good grade level for the assessment given that (a) students would be at a fitting age and ability level and (b) ninth grade constitutes the year without any standardized assessments, so students would not be overwhelmed by an unreasonable number of external assessments.

References

Finkbeiner, C. (1999). Lehrplan–Lehrwerk–Stoffverteilungsplan–Unterricht. In J.-P. Timm (Ed.), *Englisch lernen und lehren: Didaktik des Englischunterrichts* (pp. 36–44). Berlin, Germany: Cornelsen.

Harsch, C., & Hartig, J. (2015). What are we aligning tests to when we report test alignment to the CEFR? *Language Assessment Quarterly, 12*, 333–362. https://doi.org/10.1080/15434303.2015.1092545

Haß, F. (2016). *Fachdidaktik Englisch: Tradition, Innovation, Praxis*. Stuttgart, Germany: Klett.

Herman, J. L., & Webb, N. M. (2007). Alignment methodologies. *Applied Measurement in Education, 20*, 1–5. https://doi.org/10.1080/08957340709336727

Hsieh, C. (2015). *Examining the use of TOEFL Primary for young English learners in the context of English-medium instruction*. Unpublished manuscript.

Papageorgiou, S. (2016). Aligning language assessments to standards and frameworks. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 327–340). Berlin, Germany: Mouton de Gruyter. https://doi.org/10.1515/9781614513827-022

Papageorgiou, S., & Baron, P. A. (2017). Using the Common European Framework of Reference for young learners' English language proficiency assessments. In M. K. Wolf & Y. G. Butler (Eds.), *English language proficiency assessments for young learners* (pp. 136–152). New York, NY: Routledge.

Rademacher, J. (Ed.). (2016). *English G Access Band 4*. Berlin, Germany: Cornelsen.

Resnick, L., Rothman, R., Slattery, J. B., & Vranek, J. L. (2004). Benchmarking and alignment of standards and testing. *Educational Assessment, 9*(1–2), 1–27.

Roach, A. T., Niebling, B. C., & Kurz, A. (2008). Evaluating the alignment among curriculum, instruction, and assessments: Implications and applications for research and practice. *Psychology in the Schools, 45*, 158–176. https://doi.org/10.1002/pits.20282

Senatsverwaltung für Bildung, Jugend und Wissenschaft, Berlin Ministerium für Bildung, Jugend und Sport des Landes Brandenburg. (2015a). *Rahmenlehrplan für die Jahrgangsstufen 1–10 der Berliner und Brandenburger Schulen: Teil B. Fachübergreifende Kompetenzentwicklung*. Staßfurt, Germany: Salzland Druck.

Senatsverwaltung für Bildung, Jugend und Wissenschaft Berlin, Ministerium für Bildung, Jugend und Sport des Landes Brandenburg. (2015b). *Rahmenlehrplan für die*

Jahrgangsstufen 1–10 der Berliner und Brandenburger Schulen: Teil C. Moderne Fremdsprachen Jahrgangsstufen 1–10. Staßfurt, Germany: Salzland Druck.

So, Y., Wolf, M. K., Hauck, M. C., Mollaun, P., Rybinski, P., Tumposky, D., & Wang, L. (2015). *TOEFL Junior design framework* (TOEFL Junior Research Report No. TOEFL JR-02). Princeton, NJ: Educational Testing Service. https://doi.org/10.1002/ets2.12058

Webb, N. L. (2007). Issues related to judging the alignment of curriculum standards and assessments. *Applied Measurement in Education, 20,* 7–25. Retrieved from http://www.cehd.umn.edu/edpsych/C-BAS-R/Docs/Webb2007.pdf

Appendix A. Curricula and Textbooks Used

Curricula Used

Kölle, B. (2017). *Rahmenlehrplan 1–10 kompakt: Themen und Inhalte des Berliner Unterrichts im Überblick*. Berlin, Germany: Senatsverwaltung für Bildung, Jugend und Familie.

Senatsverwaltung für Bildung, Jugend und Wissenschaft, Berlin Ministerium für Bildung, Jugend und Sport des Landes Brandenburg. (2015). *Rahmenlehrplan für die Jahrgangsstufen 1–10 der Berliner und Brandenburger Schulen: Teil A. Bildung und Erziehung in den Jahrgangsstufen 1–10*. Staßfurt, Germany: Salzland Druck.

Senatsverwaltung für Bildung, Jugend und Wissenschaft, Berlin Ministerium für Bildung, Jugend und Sport des Landes Brandenburg. (2015). *Rahmenlehrplan für die Jahrgangsstufen 1–10 der Berliner und Brandenburger Schulen: Teil B. Fachübergreifende Kompetenzentwicklung*. Staßfurt, Germany: Salzland Druck.

Senatsverwaltung für Bildung, Jugend und Wissenschaft Berlin, Ministerium für Bildung, Jugend und Sport des Landes Brandenburg. (2015). *Rahmenlehrplan für die Jahrgangsstufen 1–10 der Berliner und Brandenburger Schulen: Teil C. Moderne Fremdsprachen Jahrgangsstufen 1–10*. Staßfurt, Germany: Salzland Druck.

Textbooks Used

Rademacher, J. (Ed.). (2016). *English G Access Band 3*. Berlin, Germany: Cornelsen.

Rademacher, J. (Ed.). (2016). *English G Access Band 4*. Berlin, Germany: Cornelsen.

Rademacher, J. (Ed.). (2017). *English G Access Band 5*. Berlin, Germany: Cornelsen.

Schwarz, H., & Leithner-Brauns, A. (Eds.). (2016). *Context starter*. Berlin, Germany: Cornelsen.

Appendix B. School-Internal Curriculum (Gymnasium)
Klasse 7

Lehrwerk/Materialien: Access 3 (Cornelsen), Workbook, Videos, interaktive Tafelbilder

Leistungsmessung: 4 Klassenarbeiten

Unit 1 – This is London
London, sights and songs, holiday postcards and messages

Lesen: reading street maps / tube maps, background files on London

Schreiben: sending messages and holiday postcards, writing good sentences

Sprachliche Strukturen/Grammatik: simple past / present perfect (revision), present perfect progressive including for and since

Sprechen: small talk, different stresses in German and English words, talking about sights, making plans

Übergreifende Themen und Fächer:

　　　Medienkompetenz/ Geografie: Karten auswerten

　　　Musik: Songs über London

　　　Interkulturelle Kompetenz: Gepflogenheiten in alltäglichen Konversationen

Empfohlene Literatur/Film: Medienforum Berlin: Vielfältige Audio-, Video- und Unterrichtsmaterialien (*London Files*, *Teens in London*)

Unit 2 – Welcome to Snowdonia
Town and country, moving house, volunteer work, sports, Cardiff, Wales

Lesen: background file: *Wales, Snowdonia at Night, Everything is Wrong, Climbing Mount Snowdon*

Schreiben: writing paragraphs and topic sentences, writing about moving away

Sprachliche Strukturen/Grammatik: present progressive, will future (revision), conditional I (revision); conditional II

Sprechen: planning a trip, talking about jobs

Übergreifende Themen und Fächer:

　　　Medienkompetenz: Verwendung eines English-German Wörterbuchs

　　　Geografie: Wales

　　　Interkulturelle Kompetenz: Access to cultures: similar words

Empfohlener Film: Medienforum Berlin (Online-Video) *Me and My Future: Rob and Volunteering* [Llanrumney, ein Vorort von Cardiff]

Unit 3 – A weekend in Liverpool
Visiting a museum, football in a stadium, football on the radio, sport, The Beatles

Lesen/Hören: reading *Morgan's Plan, At the Slavery Museum, At the Anfield Road—A Liverpool Hero,* listening to a song *You'll Never Walk Alone*

Schreiben: writing about a famous person

Sprachliche Strukturen/Grammatik: relative pronouns, relative clauses (revision), contact clauses, the passive

Sprechen: talking about famous people

Übergreifende Themen und Fächer:

>   Interkulturelle Kompetenz: Access to cultures: Free national museums
>
>   Musik: *The Beatles*
>
>   Sport: Football

Empfohlener Film: Medienforum Berlin (Online-Video) *Me and My Future: Tom and Dancing* [job (dancer); Liverpool Institute for Performing Arts]

Unit 4 – My Trip to Ireland

Legends, natural wonders, crossing borders, life on a farm, Irish history

Lesen/Sehen: reading *The Horse Ride, Belfast and the Northern Coast, Lewis's Travel Post, Arrival in the Burren, Yu Ming is Ainm Dom, Making the Film: Feelings* (Background file: *Moments in Irish History*)

Schreiben: using time markers in a story, making a brochure

Sprachliche Strukturen/Grammatik: modals (revision); modal substitutes; simple past (revision), past perfect

Sprachmittlung: at the hotel: conversation between a tour guide and tourists

Sprechen: pronunciation: consonants, accents

Übergreifende Themen und Fächer:

>   Interkulturelle Kompetenz: Irische Sprache und Kultur
>
>   Geografie: Irland
>
>   Literatur: Irish Legends
>
>   Deutsch/Ethik/Sozialkunde: eine Reisebroschüre erstellen
>
>   Study skills: Teamwork
>
>   Empfohlene Literatur/Film: Medienforum Berlin: Vielfältige Audio-, Video- und Unterrichtsmaterialien (*Great Britain and Ireland*)

Unit 5 – Extra-Ordinary Scotland

Festivals, a scientific project, helping friends in trouble

Lesen/Sehen: reading *Extraordinary West Coast, Star Performers, Missing, Background file: this is Scotland, Escape to . . . Scotland*

Schreiben: basic writing skills

Sprachliche Strukturen/Grammatik: reflexive pronouns / each other, indirect speech

Sprachmittlung: conducting interviews about a hobby, presenting a poster

Sprechen: planning and conducting an interview, presenting facts

Übergreifende Themen und Fächer:

      Medienkompetenz: A good presentation

      Interkulturelle Kompetenz: Access to cultures, sheep dog trials

      Ethik/Musik: schottische Musik und Kultur, Tartans, Dudelsackmusik

Empfohlener Film: Medienforum Berlin (Online-Video) *Marketing Scotland*

Klasse 8

Lehrwerk/Materialien: Access 4 (Cornelsen), Workbook, Videos, interaktive Tafelbilder

Leistungsmessung: 3 Klassenarbeiten/ Vera 8 Vergleichsarbeit/2 LEKs

Unit 1 – New York
Sights, multicultural NYC, immigration, description of places

Lesen: excerpt from a book *Faces of America*

Schreiben: writing a dialogue or a picture story

Sprachliche Strukturen/Grammatik: gerunds

Sprechen: speaking about sights

Übergreifende Themen und Fächer:

      Interkulturelles Leben: Living in New York

      Präsentationskompetenz: presentation of a special place/giving feedback to a presentation

      Sprachbewusstheit: Varietäten der englischen Sprache

      Musik: *Songs about NYC*

      Film: New York I Love You

Empfohlene Literatur: Underground New York (Cornelsen)

Unit 2 – New Orleans
History: slavery; Civil Rights; African-Americans; the South: music, food; Thanksgiving

Hör-/she-verstehen: watching *A musician in NO* (video)

Lesen: reading *Access to Culture*

Schreiben: writing a text about New Orleans

Sprachliche Strukturen/Grammatik: conditional I-III / verbs with two objects

Sprechen: speaking about something which is important in your life

Übergreifende Themen und Fächer:

      Methodenkompetenz: skimming a text

      Sprachbewusstheit: Synonyme

      Interkulturelles Leben: Cajun culture, Mardi Gras

      Politik/Geografie: Hurricane Katrina

      Musik: Jazz

Empfohlene Literatur/Film: *The Help*

<u>Unit 3 – California</u>
The Gold Rush movement, entertainment, national parks, volunteering

Hör-/sehverstehen: watching *Californian Road Trip*
Schreiben: writing a review/making an outline for a report/writing a report
Sprachliche Strukturen/Grammatik: passive
Sprachmittlung: mediation course
Übergreifende Themen und Fächer:
       Sprachbewusstheit: Wordbuilding (-ful, -less)
       Interkulturelle Kompetenz: being polite
       Geschichte: Gold Rush, Summer of Love
       Geografie: Silicon Valley
       Musik: Songs about California
Empfohlene Literatur/Film: *The Pursuit of Happyness*; Videos–national parks

<u>Unit 4 – Small Town America</u>
Schools in the states, sports, family and friends, first love, Native Americans

Hör-/sehverstehen: watching a short film about families / information about film language, shots
Schreiben: writing about photos, inner monologues (feelings)
Sprachliche Strukturen/Grammatik: *simple- progressive forms, indirect speech*
Sprechen: talking about school life and rules, giving a presentation about your own city
Sprachmittlung: mediation (rules)
Übergreifende Themen und Fächer:
       Deutsch/ Literatur: Storytelling
       Filmsprache
       Kunst: Kultur der indigenen Bevölkerung
       Literatur: The Absolutely True Diary of a Part-time Indian Sherman Alexie

<u>Unit 5 – In the Southwest</u>
Jobs, employment, child labor, migrant workers, the wilderness

Hör-/sehverstehen: watching and analyzing a documentary, working with film stills
Schreiben: writing a summary, writing a poem, a formal letter or e-mail
Sprachliche Strukturen/Grammatik: relative and participle clauses (revision)
Sprechen: commenting on child labor / dialogues about migrant workers
Übergreifende Themen und Fächer:
       PW: Menschenrechte/Kinderarbeit, Mexican Wall
       Musik: *This Land Is Your Land*
       Interkulturelle Kompetenz: Patriotismus

Klasse 9

Lehrwerk/Materialien: English G 21 A5 (Cornelsen), Workbook, Videos, interaktive

Tafelbilder;eine Lektüre nach Wahl

Leistungsmessung: 4 Klassenarbeiten

Unit 1 – Australia
Life in Australia, aboriginal people of Australia, life of a teenager in Sydney

Hör/-sehverstehen: watching (scenes from) a film: *Rabbit-Proof Fence, School of Air* (video), listening to a radio program: *Rob–sports-mad in Sydney*

Lesen: talking about statistics, reading an e-mail: *Jeannie - in the middle of nowhere*, reading a newspaper article: *The stolen generation*, reading literature: *In the Outback*

Schreiben: writing an e-mail

Sprachliche Strukturen/Grammatik: to-infinitives instead of relative clauses

Sprechen: having a conversation

Übergreifende Themen und Fächer:

      Medienkompetenz: making a poster

      Bilinguales Modul: Biologie: ecosystems in Australia (Great Barrier Reef)

      Film: Long Walk Home

Unit 2 – The Road Ahead
Young people's plans and expectations, talking about one's personality/life, schools in New Zealand

Hör/-sehverstehen: watching scenes from a reality TV show

Lesen: taking a quiz

Schreiben: writing formal letters

Sprachliche Strukturen/Grammatik: gerund with its own subject, gerund vs. infinitive

Sprachmittlung: talking about schools in New Zealand

Sprechen: asking for, confirming, giving information

Übergreifende Themen und Fächer:

      Deutsch: Vergleich mit formalen Anforderungen eines Bewerbungsschreibens / eines Lebenslaufes

      Medienkompetenz: visuell gestützte Präsentationen geben

Unit 3 – Life in the big city
Hong Kong, Johannesburg, Mumbai

Hör/-sehverstehen: watching scenes from a video/film: Hong Kong: *Live it. Love it!* (promotional video), *Slum Dog Millionaire, The Man Who Knew Infinity*

Lesen: reading a report: *Mumbai slums*; reading literature: *Mumbai homes*

Schreiben: summarizing a nonfictional text

Sprachliche Strukturen/Grammatik: participles, participle clauses

Sprachmittlung: talking about changing your flight

Sprechen: giving an oral summary

Übergreifende Themen und Fächer:

       Medienkompetenz: digitale und analoge Medien selbstständig zur Informationsbeschaffung nutzen; eine Präsentation unter Einbeziehung digitaler Medien halten

       Geografie: Städtepräsentationen

       Musik: *Give me hope Joanna* (Johannesburg)

Empfohlene Literatur/Film: Hong Kong: *Live it. Love it!* (promotional video), *Slum Dog Millionaire, The Man Who Knew Infinity, The Pigman, Speak, Holes, Looking for Alaska, The Fault in Our Stars, Big Mouth Ugly Girl, Hidden Figures*

Klasse 10

Lehrwerk/Materialien: Context Starter (Cornelsen); Videos, Podcasts/Audios, Interaktive Tafelbilder

Leistungsmessung: 3 Klassenarbeiten, 1 MSA schriftl., 1 MSA mdl.

Unit 1 – The Time of Your Life
Life plans; aims and ambitions, hopes and fears

Hör/-sehverstehen: listening to a podcast, radio report *Sailor Girl*; viewing a film, *Make It Count*

Lesen: reading nonfiction, narrative prose, (extracts from) a novel *The Midnight Driver*, using reading strategies: skimming and scanning a text

Schreiben: making and taking notes, writing a speech, writing a comment, writing a new ending for a short story, writing an analysis/ characterization

Sprechen: talking about aims and ambitions, hopes and fears, describing events, past or present, expressing personal ideas, opinions, talking in dialogues, monologues,, having a discussion

Text- und Medienkompetenz: analyzing narrative prose, working with charts and graphs, making a poster, creating a comic

Übergreifende Themen und Fächer:

       Sozialkunde, Deutsch: Jugendproblematik

       Ethik: Zusammenleben, Gruppenzwang

Empfohlene Literatur/Film: *Make it Count; The Midnight Driver; The Road Not Taken* (short story); *Free!* (short story)

Unit 2 – Communicating in the Digital Age
Importance of the internet, the influences and impacts of digital media on our daily lives, cyberbullying

Hör/-sehverstehen: listening to a podcast *How to Deal with Cyberbullying*, viewing and analyzing a film *The Cyberbullying Virus*

Lesen: reading nonfiction, narrative prose, reading extracts from a novel *B4: I C U, Thirteen Reasons Why*; analyzing text structure, examining text types, narrative perspective, working with visuals (photos), working with charts and graphs working with cartoons

Schreiben: practicing note taking, mind mapping, writing a comment, writing a summary, writing a well-structured text

Sprechen: talking about the impact and influences of digital media on our professional, social and personal lives, speaking in dialogues; monologues, having a discussion

Text- und Medienkompetenz: making a poster, conducting research on the Internet, giving a presentation

Übergreifende Themen und Fächer:

>Informatik/Sozialkunde/Deutsch/Ethik: über die Bedeutung und Einflüsse digitaler Medien im allgemeinen und auf unser Kommunikationsverhalten diskutieren, den sozialen Umgang im Internet aus verschiedenen Blickwinkeln betrachten, Mobbing

Empfohlene Literatur/Film: *How to Deal with Cyberbullying, The Cyberbullying Virus, B4: I C U, Thirteen Reasons Why, Wonder*

Unit 3 – Living in the Global Village
Globalization and its challenges, climate change, global warming, pollution, sustainability

Hör/-sehverstehen: listening to an interview *The Fairphone*; viewing films, e.g. *Smartphones and Sustainability*, *Tomorrow*; viewing a news report

Lesen: reading a fictional text *A Vision of the Future in China*; reading an online article, examining text types, narrative perspective, working with visuals (photos), charts, graphs and cartoons; using different reading strategies: skimming and scanning of a text

Schreiben: making and taking notes, writing an article for a school magazine, writing a personal letter, a guided report

Sprechen: talking about the impact and influences of digital media on our professional, social and personal lives; speaking in dialogues, monologues; having a discussion

Text- und Medienkompetenz: analyzing narrative perspective, working with charts and graphs, interpreting statistics and infographics, conducting research on the Internet, analyzing a cartoon

Übergreifende Themen und Fächer:

>Sozialkunde/Geografie/Ethik/Deutsch: verschiedene Aspekte der Globalisierung diskutieren, wirtschaftliche Faktoren der globalen Vernetzung beleuchten, Fragen zu Umweltverschmutzung und Klimawandel besprechen, Handeln auf lokaler Ebene, globales (Um)Denken bewirken

Empfohlene Literatur/Film: *The Fairphone, A Vision of the Future in China*

Appendix C. Coding Categories (Domains, [Enabling] Skills, and Task Types or Subskills)

Domains
Academic
Social and interpersonal
Navigational

Skills
Listening
Reading
Speaking
Writing
Language Form and Meaning

TJS task type	Goals
Listening Comprehension	
L_Main idea	Can identify the main idea
L_ Speaker's purpose	Can make inferences based on what is implied but not explicitly stated
	Can recognize rhetorical structures (e.g., compare/ contrast, cause and effect)
L_Prediction	Can correctly interpret pragmatic meaning (intonation, contrastive stress, intended meaning in specific contexts, idiomatic speech, etc.)
	Can make predictions based on what is stated
L_Inference	Can identify speaker's purpose or purpose of a message
L_Rhetorical device	Can make inferences based on what is implied but not explicitly stated
L_Prosody	Can make predictions based on what is stated
L_Inference	Can make predictions based on what is stated
L_ Detail	Can identify important details
L_Prediction	Can recognize a speaker's purpose
Reading Comprehension	
R_Main idea	Can understand main ideas
R_Factual information	Can understand key details and/or supporting information
R_Vocabulary in context	Can understand negative factual information in which three choices are true and one is not
R_Vocabulary figurative language	Can make inferences based on what is implied but not explicitly stated
R_Vocabulary idiomatic language	Can demonstrate understanding of the meaning of middle school–level words in context
R_Pronoun referent	Can figure out the meaning of challenging words from context
R_Inference	Can understand referential relationship between a pronoun and the subject or object that matches it
R_Author's purpose/Rhetorical structure	Can understand the author's purpose in a variety of genres
Language Form and Meaning	

LFM_LF   Can identify the correct, meaningful use in context of the following grammatical
- Verb form   structures:
- Subject   - noun/ pronoun, gerund, or infinitive clause
- Object   - verb tense
- Passive voice   - adjective phrase
- Adjective or   - adverb phrase
  adjective phrase   - comparatives/superlatives
- Adverb or adverb   - noun clause or noun phrase
  phrase   - subject and verb combination
- Relative clause   - word form
- Word order   - word order
- Noun phrase or noun   - relative pronoun/ clause
  clause   - passive voice
- Expletive   - appropriate lexical unit (e.g., noun, verb, modifier, conjunction, preposition)
- Object complement
- Comparative or
  superlative
- Subject and verb
- Predicate nominative
- Word form
LFM_LM   Can demonstrate knowledge of vocabulary meaning in academic and nonacademic
- Conjunction   texts of varying complexity
- Preposition
- Noun
- Modifier
- Determiner

Writing activities
W_summary
W_dialogue
W_note taking
W_sentences
W_email
W_essay
W_commenting

Speaking activities
S_read aloud
S_opinion
S_short answer
S_explaining
S_presenting
S_discussion
S_predicting

Appendix D. Focus Group Protocol

Ort/location:

Zeit/time:

Durchführung/facilitator:

Anzahl der Teilnehmer/number of participants:

---

1. Vorstellungsrunde und Ziele der *Focus group*?

Introduction and what is a focus group?

A focus group needs to start with something easy to get the participants to talk. Therefore I plan to introduce myself and tell them a little bit about the alignment study and the purpose of conducting this focus group.

2. Wie gut kennen Sie den TOEFL Junior Standard Test?

How familiar are you with the TOEFL Junior Standard?

The goal of Question 2 is to familiarize the participants with the TJS. Although they have administered the test over the past years in the context of the competition, I was told that teachers know very little about the test itself. So the goal of this question is to obtain an understanding of their knowledge of the TJS and to gauge their general knowledge and familiarity with standardized assessments. Depending on how little they know, I may present them with some basic knowledge about the TJS and with the ability to ask questions about the test.

      a. Haben Sie Fragen zum TJS?

      Do you have any questions regarding the TJS?

      b. Aufbau und Ziele des TOEFL Junior Standard (TJS)

      Structure and claims of the TOEFL Junior Standard

3. Erfahrungen mit dem TJS

Experiences with TJS

Given that the teachers will have administered the assessment (yet presumably know very little about the TJS), I would like to obtain an understanding about their experiences with test administrations and their general use of the test results.

      a. Wann, wie und zu welchem Zweck verwenden Sie an Ihrer Schule den TJS?

      When, how, and for which purpose do you use the TJS at your school?

  b. Können Sie die Abläufe während der Testdurchführung beschreiben?

  Can you describe the steps taken during test administration?

  c. Wer wertet die Testergebnisse aus?

  Who scores the test responses?

  d. Wie nutzen Sie die Ergebnisse im Unterricht und/oder im weiteren Schulkontext?

  How do you use the test results in your instruction and/or in the larger school context?

4. Denken Sie, dass der Test in ausreichender Weise die Inhalte der Lehrbücher widerspiegelt?

Do you think that the test content adequately reflects what is being taught and addressed in the textbooks?

The textbooks are the basis for what gets taught. The curricula contain very broad standards (if they can even be called that). Therefore textbooks have served as the reference in the alignment study and will thus also be used here as the point of reference.

  a. Warum? Warum nicht?

  Why? Why not?

  b. Deckt der Test die im Unterricht behandelten sprachlichen Fähig- und Fertigkeiten ausreichend ab?

  Does the test cover the different language skills that are being taught?

5. Spiegeln die Testergebnisse des TJS Ihrer Meinung nach die unterschiedlichen Leistungsniveaus der Schüler/Innen wider?

Do you think that the test scores appropriately reflect your students' different proficiency levels?

Questions 5 aims at obtaining a deeper understanding, drilling down in terms of different skills covered on the test and different levels of proficiency assessed.

  a. Warum/warum nicht?

  Why? Why not?

6. Welche Vorteile ergeben sich aus der Nutzung des TJS als Lernersatzleistung?

What are advantages to using the TJS as a classroom assessment?

7. Gibt es potenzielle Nachteile, die sich aus der Nutzung des TJS als Lernersatzleistung ergeben?

Are there potential disadvantages to using the TJS as a classroom assessment?

8. Sie weitere Anregungen oder Fragen bzgl. des TJS oder der Studie?

Do you have any further questions, comments, or concerns regarding the TJS or the study?

To conclude, I would like to give teachers the opportunity to raise any additional questions or comments they may have.

Vielen Dank für Ihre Zeit und Beteiligung an der Studie.

Thank you for your time and participation in this study.

Notes

[1]http://go4goal.eu/

[2]The full scale of the German grade system includes the following levels: 1 = *sehr gut* [very good]; 2 = *gut* [good]; 3 = *befriedigend* [satisfactory]; 4 = *ausreichend* [sufficient]; 5 = *mangelhaft* [insufficient]; 6 = *ungenügend* [insufficient/fail].

[3]For more information about Berlin's education acts, see https://www.berlin.de/sen/bildung/schule/rechtsvorschriften/

[4]https://www.berlin.de/sen/bildung/schule/bildungswege/gymnasium/

[5]https://www.berlin.de/sen/bildung/schule/bildungswege/integrierte-sekundarschule/

[6]*Abitur* denotes the final school-leaving examination or diploma from German secondary school, which serves as qualification for university admission or matriculation.

[7]See, for more information about the two core school types in Berlin, https://www.berlin.de/willkommenszentrum/en/families/schools/

[8]Note that only activities related to English-language learning or practice were coded. Activities that explicitly asked students to read about or discuss a topic in German (e.g., *Context Starter*; Unit 2) or to shoot a photo or video (e.g., *English G Access 3*; Unit 3) were excluded from the analysis.

[9]The instructor from the *ISS* was unable to participate due to illness.

[10]She is referring to a specific portfolio of materials provided by the publisher of the textbook series. Teachers oftentimes use these supplementary materials to select tasks for the written classroom assessments.

[11]Please note that there is a separate TOEFL Junior Speaking test available: https://www.ets.org/toefl_junior/prepare/speaking_sample_questions/responses/