



# TOEIC® Program fosters *positive washback*

Featuring authentic communicative tasks to enhance practical communicative competence

The TOEIC Program tests emphasize authentic English communication to help score users make meaningful interpretations about test takers' real-life communicative abilities. By aligning test content, tasks and scoring criteria with real-world language use, the TOEIC Program tests are designed to encourage positive "washback." In other words, appropriate use of TOEIC Program tests should have positive impacts on English teaching and learning worldwide.

# The TOEIC Program tests feature authentic English communication tasks from real-life contexts



*Communicative competence is the ability to "use" English to get things done*

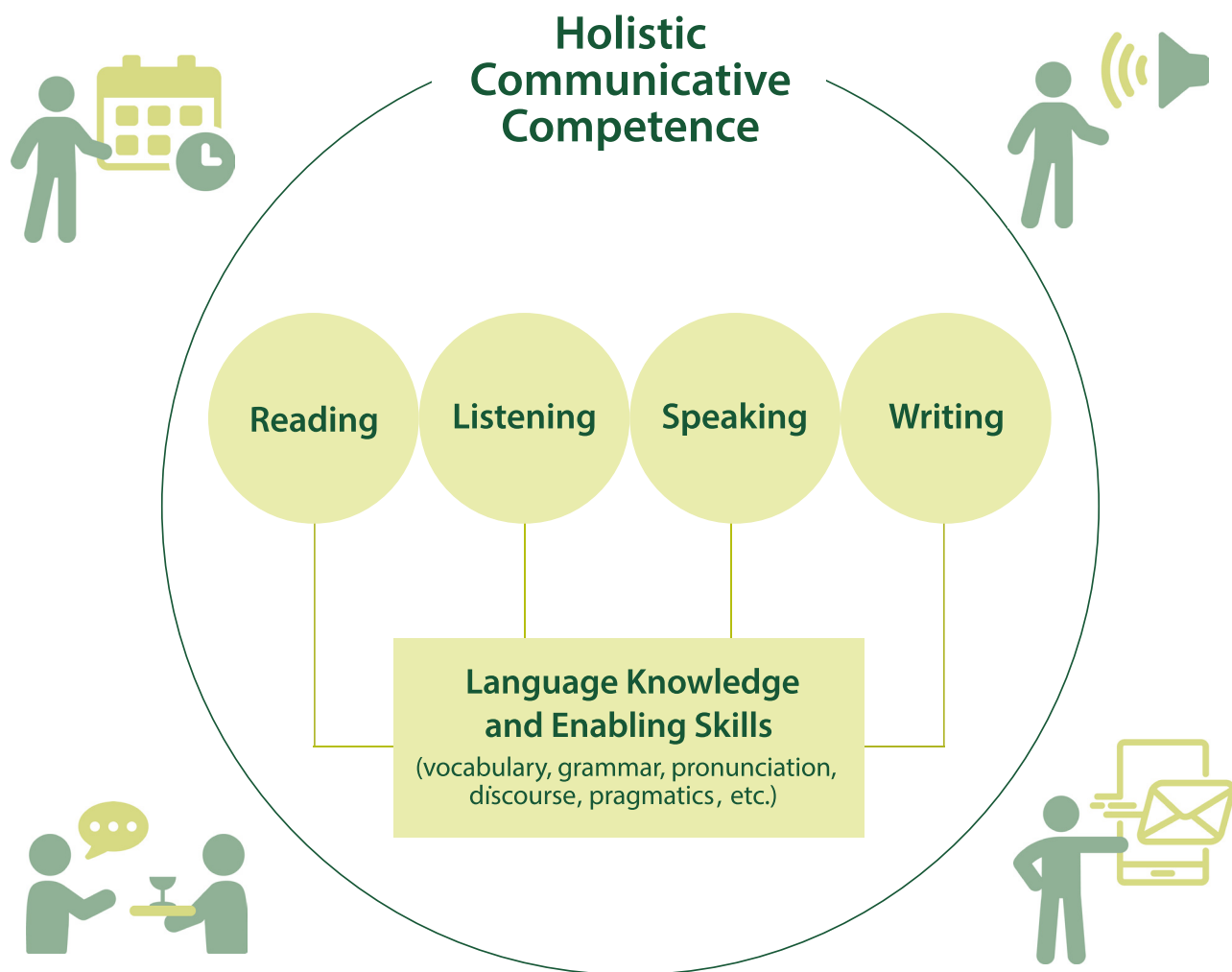
The TOEIC Program tests are designed to directly measure how well test takers can use their English knowledge and skills to communicate effectively in common daily life and workplace settings in the international environment.

As assessments of communicative competence, also referred to as ability for use, these tests are intended to measure more than just a test-taker's knowledge of English grammar, vocabulary or the skills that enable them to understand and produce language. They also test their skill in listening for and detecting English words, their skill in pronouncing English phonemes or their skill in constructing English sentences.

To be sure, all of these (and other) types of language knowledge and enabling skills are important components of a learner's communicative competence, but in and of themselves, they do not represent the actual ability to communicate.

Effective communication involves using language to get things done. It draws on the learner's language knowledge and enabling skills as they listen, read, speak and write in diverse settings with different kinds of texts and audiences for multiple purposes. It requires the ability to interpret and create meaning with English.

# Language is used in authentic ways to accomplish a communicative task



A useful way to think about ability for using English is to envision the kinds of tasks that involve communication. Communication tasks require the application of language abilities in order to get things done. For example:

- We listen to an announcement so that we can know important information and decide what to do next.
- We read a schedule in order to understand when and where something will happen.

- We speak at a restaurant to order what we want to eat.
- We write an e-mail to inform recipients about our ideas or needs or goals.

These tasks illustrate how language is used to communicate in authentic ways, that is, for a purpose, in a situation or context and often with a specific audience. The TOEIC and TOEIC Bridge tests are designed to assess English learners' abilities to communicate authentically in the real world on tasks like these.

<sup>1</sup> Discourse means the ways in which words and sentences are put together to create a coherent spoken or written text.


<sup>2</sup> Pragmatics means the ways in which language is used in a particular situation or context or with a particular audience.

# Assessing through tasks that are drawn from real-world daily life and workplace settings

In order to assess learners' English communication abilities in authentic ways, the TOEIC and TOEIC Bridge tests consist of multiple tasks that are drawn from real-world daily life and workplace settings (e.g., Ashmore et al., 2018; Park & Bredlau, 2018).

These tasks require test takers to interpret (listen and read) or produce (speak and write) language in a context and for a purpose, thereby reflecting the ways in which English is actually used to communicate.

For one example, a task on the TOEIC Speaking test might provide test takers with authentic input like the figure below, which shows information related to a series of seminars about starting a new business.

  
**STARTING AND MANAGING YOUR OWN BUSINESS**

**Date:** May 27  
**Location:** Bristol Office Building  
**Seminars:** 9:00 A.M. "Financing Your Business," Room 210—*Martha Ross, Certified Public Accountant*  
11:00 A.M. "How to Promote Your Own Business," Room 312—*Howard Brown, Brown Publishers*  
**OR**  
11:00 A.M. Planning for Profit, Room 318—*John Phillips, Phillips Associates*  
1:00 P.M. Lunch\*  
2:00 P.M. Sales Techniques Workshop, Room 248—*Helen King, West Side Consultants*  
4:00 P.M. General Discussion

**Registration Fee:** Individuals, \$95.00  
Members of the Business Information Center, \$75.00

\*Not included in registration fee

The test-taker's task is first to review and understand the information in the figure and then to answer a series of questions from a caller who is interested in attending the seminars (e.g., "Could you tell me what time the conference starts and how long it will last?").

The test takers, then, are embedded within a real-life workplace setting where they must deal with information in order to interact effectively with potential clients by providing accurate answers to their questions.

Clearly, this task involves knowledge about English vocabulary and grammar as well as the ability to read, listen and ultimately speak, just like tasks in the real world that often draw on multiple types of language knowledge and skills.

## Scoring is also aligned with real-world criteria

Test-taker performances are also scored according to real-world expectations, or criteria, for each task type on the TOEIC Program tests.

For example, performance on the above task would be scored not only for how well the test taker utilizes appropriate vocabulary and grammar and how intelligibly or clearly they speak but also for how accurately the test taker answers the questions and

how complete and socially appropriate the responses are.

Test performances are not scored based on whether the test taker uses language like a native speaker but rather on the comprehensibility or communicative effectiveness of the response from the perspective of a proficient user of English – who may be a native or nonnative speaker of English.

# The reasons why the TOEIC Program emphasizes authenticity:

## 1. It is a “must” for tests that directly measure communicative competence

So why does the TOEIC Program place such an emphasis on authentic English communication? There are several important reasons.

First, test users, including stakeholders in corporations and organizations, educators and others, generally want to know what a test taker can do in English and how well they can do it, not merely what the test taker knows about the English language system and structure. Test users look at English language proficiency test scores as indicators of test-takers’ communicative competence, that is, their ability to use English to communicate.

English language tests that do not include authentic communication tasks in each of the four skills are at best only indirect and incomplete measures of someone’s ability to get things done in English.

In order to support meaningful interpretations about test-takers’ communicative abilities, ethical and responsible language tests must include the kinds of authentic communication demands that reflect real-world language use (Norris & East, 2021).

## 2. To provide a positive influence on teaching and learning

A second and critically important reason why the TOEIC Program emphasizes authentic communication is that by doing so, English learners and teachers are provided a meaningful language learning goal and a trustworthy measure of progress toward that goal. As they develop learners’ English language knowledge and skills to succeed on the TOEIC Program tests, they are simultaneously working toward the meaningful goal of communicative competence in English.

One result is that learners will be able to accomplish authentic communication tasks in English, certainly a desirable outcome for language teaching. This relationship between the TOEIC Program tests and English language teaching and learning is explored in more detail in the next section.

# Appropriate choice of tests can lead to positive outcomes

One of the main purposes of the TOEIC Program is to produce language tests that can be used to support high-quality decision making.

Language tests produce information about language proficiency that may be used to make important decisions about individuals.

For example, an organization may want to incorporate information about candidates' language proficiency into a "hiring" or "promotion" decision.

A training institution may want to use information about students' language proficiency to make a "placement" decision.

The use of a language test to make course placement decisions is illustrated in the figure below.


The process begins when students take a test that provides a standardized sample of their language proficiency, or a 'test performance.'

A test performance is transformed into a score using a systematic scoring process. In the example below, the test taker receives a score of '3.'

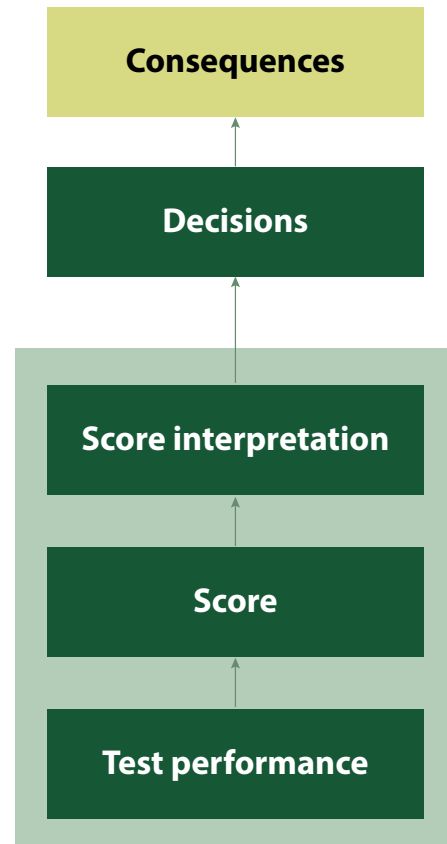
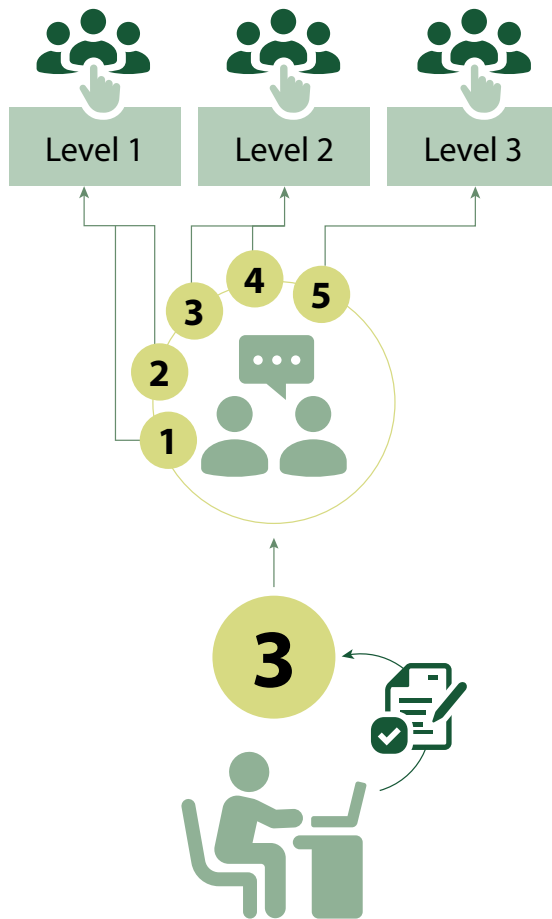
A score itself is only a number. To be useful, the score needs to be understood (or interpreted) in terms of a test taker's knowledge, skills, and abilities. In the example below, possible scores range from 1 to 5 and can be understood in terms of a test taker's speaking proficiency and communicative effectiveness. The test taker received a '3' and the meaning of this score should be communicated by the test developer.

A score user's understanding of the meaning of test scores should directly inform decision-making. In the example, students are placed into coursework based on their speaking proficiency. Students who receive a score of 1 or 2 are placed in a Level 1 course; students who receive a score of 3 or 4 are placed in a Level 2 course; and students who receive a '5' are placed in a Level 5 course. Thus, a student who received a score of '3' is placed in a Level 2 course.

In the example, setting up these decision rules required (a) an understanding of the meaning of test scores, or score interpretations; and (b) how score interpretations relate to the minimal requirements needed for each course.



*Selecting more relevant and higher quality tests can lead to better outcomes*



Decisions based on test scores have consequences (or outcomes) that can be positive or negative. For example, an organization may use language test scores to hire or promote a new employee into a position where language skills are needed or desired.

For a “hiring” decision, if the new employee can use their language skills appropriately we can say the decision had a positive outcome. If the new employee does not have the language skills needed, the decision had a negative outcome.

In the “placement” decision example, if the training institution classifies students into coursework appropriately, the decision has a positive outcome.

Making an accurate placement decision can contribute to effective teaching and learning, and, of course, the misclassification of students into inappropriate courses would likely detract from teaching and learning.

Because the choice of language test can impact the outcomes of decisions, test users should keep several critical points in mind.

First and foremost, the information about language proficiency provided by the language test should be relevant to the language skills, knowledge and abilities needed for the decision.

In many situations, test users are interested in communicative competence in English, which is best demonstrated by a test that directly measures listening, reading, speaking and writing skills (versus a test that may be used to predict communicative competence based on narrower or more indirect test tasks).

Secondly, the quality of the language test score(s) should be sufficient for the importance of the decision that will be made.

Two different tests may claim to measure the same skill(s) – for example, English speaking skills – but they will inevitably differ in how they define and measure those skills.

The two different tests will also differ in the level of quality of their measurement in reliability, validity, fairness. Ultimately, the best-case scenario for high-quality decision making is identifying a test that provides a high quality measurement of the most relevant skill(s) using a more direct approach.

## Washback of tests on teaching and learning

From an educational perspective, test-users' choice of language tests can have an impact on language teaching and learning (e.g., Hsieh, 2017). Especially when the decisions are extremely important – such as employment or selection decisions – research has shown that testing can have a large impact on teaching and learning.

When the language test is important, teachers and learners may spend a lot of time preparing for it. Consequently, teachers and learners focus on improving the language skills evaluated by the test. This consequence of using tests for important decisions is also called the “washback” of tests on teaching and learning.



# Washback can be positive or negative

It is important to realize that washback can be **positive** or **negative**. If the language test uses authentic and appropriate test tasks — realistic tasks designed to assess English learners' abilities to communicate authentically in the real world — and focuses on appropriate language skills, this can have a positive impact on teaching and learning.

On the contrary, if the language test uses inauthentic and inappropriate tasks, this can have a negative impact on teaching and learning by encouraging teachers and learners to focus on tasks that do not reflect how language is used in the real world (see discussion in Norris, 2018).

Also, if the language test focuses on a very limited aspect of language knowledge or skills, this can have a negative impact on teaching and learning by encouraging teachers and students to focus on narrow aspects of language ability instead of a broader communicative competence.

Now let us think about washback from two perspectives: authenticity of test tasks and a focus on which skills to measure.

## Factors that affect washback: Authenticity of test tasks

First, why is authenticity of tasks important? The extent to which a language test includes authentic communication tasks can create positive or negative washback.

For example, if a listening and reading test mostly consists of test items that measure knowledge of grammar and vocabulary, students may be encouraged to focus on memorizing grammatical rules and vocabulary.

While knowledge of grammar and vocabulary is an important foundational aspect of language ability, the ability to use listening and reading skills to comprehend language in real-life settings requires a broader set of knowledge and skills.

Authentic test tasks – or tasks that effectively simulate real-life communication situations – encourage teachers and students to focus on the skills needed for real-life communication.

The TOEIC and TOEIC Bridge tests are designed to facilitate positive washback with this principle in mind (Schmidgall et al., 2021).

For example the task below is from a sample test of the Reading section of the TOEIC Listening and Reading tests that requires test takers to refer to multiple passages to find an answer to the question in a limited period of time in a workplace context. If you can read through the three passages and collect the necessary pieces of information from each passage, you can find the answer.

This information processing – finding the information from various materials in a limited period of time – is exactly what people do in real-life situations, especially in workplace settings.

In such cases, the ability to capture necessary information and process it within a given timeframe is critical to meet the communication goal set for this kind of task.

Therefore, including such authentic tasks in the test can have positive washback. By practicing such tasks in preparing for the test, test takers can learn how to read English passages efficiently and process information efficiently, which is a useful skill in real life.

From:

Arun Phan <arun.phan@tnet.com>

To:

Customer Support <support@sparkypaints.com>

Date:

March 12

Subject:

Order #3397

Hello,

Thanks for sending my order #3397—it arrived this morning. Unfortunately, the paint was not the one I had asked for. I had selected color SP 944 but received SP 945 (Ocean Waves). They appear right next to each other on your Web site, so the two may have been confused at your end. Could you send me the correct paint, along with additional samples that are close in color to SP 722? That sample worked well in my house; the others looked too green on my walls.

Thank you,

Arun Phan

In his email, which color does Mr. Phan indicate that he likes?

- (A) Caspian Blue
- (B) Deep Sea Blue
- (C) Stormy Blue
- (D) Misty Gray

Answer: (C) Stormy Blue

Sparky Paints, Inc.

Sparky Paints, Inc., makes it easy to select the right colors for your home. Browse through hundreds of colors on our Web site, [www.sparkypaints.com](http://www.sparkypaints.com). Select your top colors, and we'll send free samples right to your door. Our color samples are three times larger than typical samples found in home-improvement stores and come with self-adhesive backing, allowing you to adhere them to your walls so you can easily see how colors will coordinate in your home. When you're ready to begin painting, simply select your chosen colors online, and we'll ship the paint of your choice to arrive at your home within 3-5 business days, or within 2 business days for an additional expedited shipping fee.

\*Actual colors may differ slightly from what appears on your monitor. For this reason, we recommend ordering several samples in similar shades.

http://www.sparkypaints.com/shoppingcart

Sparky Paints, Inc. 



Order Summary #3397

Customer: Arun Phan

Item	Size	Quantity	Price
Caspian Blue SP 237	n/a	1	\$0.00
Deep Sea Blue SP 298	n/a	1	\$0.00
Stormy Blue SP 722	n/a	1	\$0.00
Misty Gray SP 944	Gallon	2	\$50.00
Tax (8 percent)			\$4.00
Expedited shipping			\$18.99
<b>Total</b>			<b>\$72.99</b>

Proceed to Checkout

# Factors that affect washback:

## A focus on which skills to measure

Why is it important to consider which skills should be focused on and measured? In many cases, language test users are generally interested in communicative competence.

The premise of washback is that you need to use appropriate tests for your teaching and learning goals. If you aim for learners to acquire holistic communicative competence, you certainly should test all four skills. This will have positive washback on teaching and learning because four-skills testing leads to a balance of receptive and productive skills (CASE 1-1).

The type of test chosen by test users also sends a message to teachers and learners about which model of communication is valued.

As shown in CASE 1-1, a four-skills model has the potential to have a more positive washback on teaching and learning, where the goal is overall communicative competence.

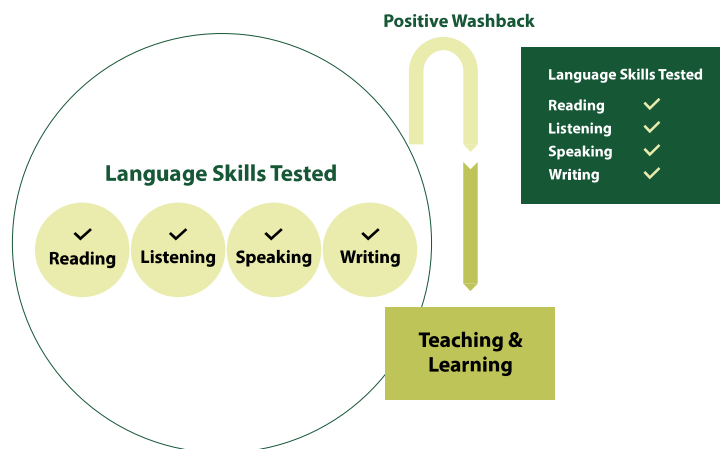
Teachers and learners are encouraged to focus on all aspects of language ability when they prepare for the test. In the long-term, this should lead to better language proficiency outcomes as teachers and learners focus on the broad set of competencies needed to communicate effectively.

Also, as you can see in CASE 1-2, if your primary goal is to acquire reading and listening skills, testing only reading and listening skills (using authentic listening and reading tasks) can still give you positive washback because it leads to teaching and learning listening and reading (although ultimately, four-skills testing will also contribute to development of listening and reading skills because the four skills are interrelated).

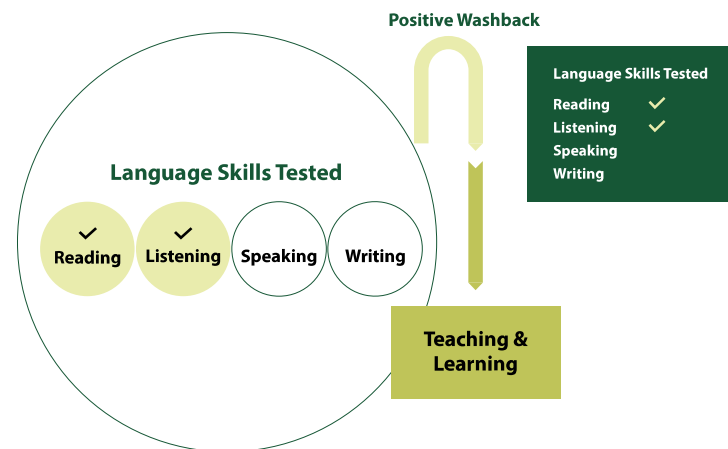
## Case 1

Use of an appropriate test for the teaching and learning goals => Positive Washback => appropriate teaching and learning

1. A test taker who wants to **acquire overall communicative competence** from the start of their learning and tests all four skills.



2. A test taker whose first priority is to **acquire only reading and listening skills** within a particular period of time and tests two skills: reading and listening.



Now consider what could happen if test users care about holistic communicative competence (four skills) but only use information about listening and reading skills, as shown in CASE 2-1.

If test users only have information about listening and reading skills, they have less information about overall language ability and a reduced model of communicative competence.

This reduces the quality of their language evaluation and could lead to more errors in decision-making, reducing their positive outcomes.

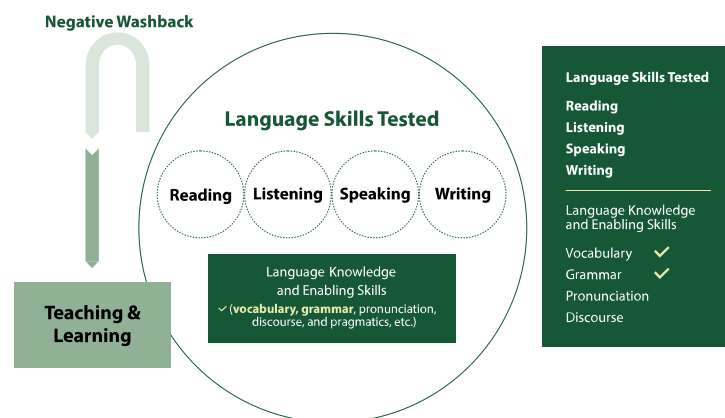
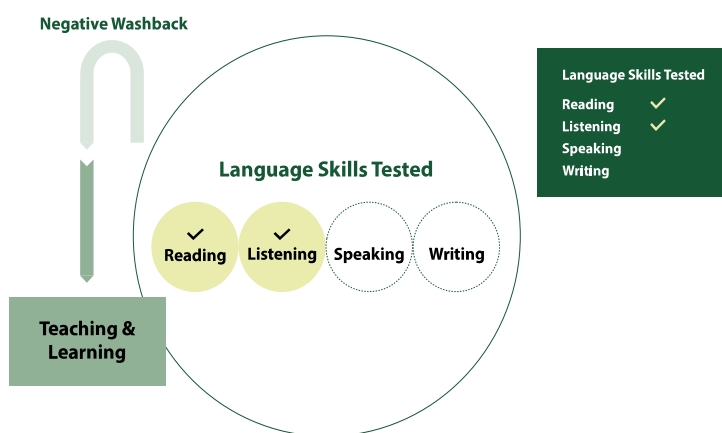
In addition, because of the washback effect, they are sending a message to teachers and learners that a four-skills model of communicative competence is not valued. Also, as in CASE 2-2, if test users care about listening and reading ability but test only grammar and vocabulary, it can lead to negative washback as well.

## Case 2

Use of an inappropriate test for the teaching and learning goals => Negative Washback => inappropriate teaching and learning

1. A test taker who wants to **acquire overall communicative competence** from the start of their learning but only tests listening and reading skills.

2. A test taker who wants to **acquire listening and reading skills** but only tests grammar and vocabulary skills.



As we have seen, using appropriate tests in terms of both (a) how authentic tasks are and (b) which skills to focus on and measure can lead to positive washback on teaching and learning.

The TOEIC and TOEIC Bridge tests measure all four skills with authentic test tasks, which encourages teachers and students to focus on the skills needed for real-life communication.

This is the very principle that drives the design of the TOEIC and TOEIC Bridge tests.

Ultimately, it's very important to learn all four skills and test all four skills even for those who want to acquire a particular set of skills because all four skills are intertwined.

# References

Ashmore, E., Duke, T., & Sakano, J. (2018). Background and goals of the TOEIC® Listening and Reading update project. In D. Powers & J. Schmidgall (Eds.), *The research foundation for the TOEIC® tests: A compendium of studies: Volume III* (pp. 3.1–3.8). ETS. <https://www.ets.org/s/toEIC/pdf/research-compendium.pdf>

Hsieh, C.-N. (2017). *The case of Taiwan: Perceptions of college students about the use of the TOEIC® tests as a condition of graduation* (Research Report No. RR-17-45). ETS. <https://doi.org/10.1002/ets2.12179>

Norris, J. M. (2018). Task-based language assessment: Aligning designs with intended uses and consequences. *JLTA Journal*, 21, 3–20. [https://doi.org/10.20622/jltajournal.21.0\\_3](https://doi.org/10.20622/jltajournal.21.0_3)

Norris, J. M., & East, M. (2021). Task-based language assessment. In M. J. Ahmadian & M. H. Long (Eds.), *The Cambridge handbook of task-based language teaching* (pp. 507–528). Cambridge University Press.

Park, E., & Bredlau, E. (2018). Expanding the question formats of the TOEIC® Speaking test. In D. Powers & J. Schmidgall (Eds.), *The research foundation for the TOEIC® tests: A compendium of studies: Volume III* (pp. 2.1–2.8). ETS.

Schmidgall, J., Cid, J., Carter Grissom, E., & Li, L. (2021). *Making the case for the quality and use of a new language proficiency assessment: Validity argument for the redesigned TOEIC Bridge® tests* (Research Report No. RR-21-20). ETS. <https://doi.org/10.1002/ets2.12335>