



*The Voxy Proficiency Assessment® (VPA)<sup>1</sup> is an objective, in-product proficiency instrument designed to measure learners' general level of English. The first VPA was released in August 2012, with subsequent forms added every three months.*

*Taken before learners begin using Voxy, the VPA is used as a benchmark tool to determine each learner's current proficiency level, and then to measure how much the learner has improved at a later time when another form of the VPA is taken. The following paper will briefly review the literature that informed test development, offer information on how learners and institutions can interpret their VPA scores, explain some of the key research studies Voxy has conducted with VPA score data, and conclude with a review of upcoming research initiatives and planned VPA developments.*

## **VPA Components**

There are multiple forms of the VPA, and each form is comprised of test items at different levels of difficulty. Depending upon how learners self-assess their proficiency, they receive versions of the test with items that are more or less difficult that range from tests designed for advanced learners to those intended for 0-beginners, who are learners with no previous experience learning English. Each form of the VPA has multiple versions, each of which has an equal number of grammar constructs and comprehension question types at each level of difficulty. For example, each form of the VPA includes an easy item on irregular past tense verbs in the grammar section, and a hard item on the purpose of a paragraph in a reading passage for the reading section. Depending on the level of the test and the proficiency level of the learner, it takes anywhere from 20 to 80 minutes to complete the VPA, with most learners spending 30 to 45 minutes.<sup>2</sup>

Each form of the VPA assesses a learner's reading comprehension, listening comprehension, and grammatical knowledge.<sup>3</sup> There are a total of 63 multiple choice (MC) items for each form, except for the 0-beginner forms, which have a total of 21 MC items. There are three components of the VPA, each of which assesses a different skill or competency.

---

<sup>1</sup> In keeping with language assessment reliability and validity standards, all VPA forms have a Cronbach's alpha of at least 0.80.

<sup>2</sup> The actual average time spent taking the VPA was 32 minutes (SD=20) for 3087 learners.

<sup>3</sup> Currently being developed is a speaking section, which will be completed synchronously with a live examiner, and a writing section, which will be completed asynchronously.

## Reading Comprehension

Second language reading ability has been conceptualized in terms of language skill, knowledge resources, and general cognitive skills (Carpenter & Just, 1986; Carr & Levy, 1990; Grabe, 1991; Rayner & Pollatsek, 1989). Building on earlier findings (Alderson, 2000; Grabe, 2009, Khalifa & Weir, 2009; Adlof, Perfetti, & Catts, 2011), Grabe and Jiang (2014) identified a number of factors that affect reading ability and account for individual differences in reading comprehension, such as vocabulary knowledge and the ability to recognize discourse structuring and genre patterns.

These factors were considered when creating the reading section, which has a total of 21 MC items (seven for the 0-beginner forms). These factors were also considered while selecting passages for the reading section. Comprehension items based on these passages—which are all authentic reading materials, such as job advertisements, news articles, and how tos—assess a learner’s knowledge of vocabulary within a given context, the ability to identify the main ideas of a text, the ability to interpret text meaning critically, and understanding of discourse structure. In addition, these passages cover not only a variety of formats but also a range of topics, such as business, technology, entertainment, and health.

## Listening Comprehension

Some researchers posit that second language listening ability can be measured by the ability to process both explicit and implicit information during listening tests (Hansen & Jensen, 1994; Nissan, Devnicenzi, & Tang, 1996; Wagner, 2002). Wagner’s (2002) model of second language listening ability was used as a reference when creating the listening section of the VPA.

The listening section has 15 MC items (five for 0-beginner forms) that assess a learner’s ability to listen for specific details and facts, identify the purpose of the spoken discourse, make inferences about speakers’ attitudes and pragmatic meanings, and decode words through context. As with the reading section, the listening section covers an array of recorded content, such as speeches, candid conversations, advertisements, and interviews.

## Grammatical Knowledge

While grammatical knowledge is not specifically one of the four skills that determines individual competence in a second language, it is usually conceptualized within the broader context of communicative language ability (Canale & Swain, 1980; Bachman & Palmer, 1996; Purpura, 2004). For this reason, the VPA contains a section that assesses grammatical knowledge.

The grammar section has a total of 27 MC items (nine for 0-beginners forms). These items assess a learner's command of English word formation and sentence formation. Items include questions about modals, prepositions, verbs, and tense. Knowledge of cohesion (e.g., use of pronouns and lexical repetition) and information management (e.g., using connectors) is assessed as well. The Appendix lists the constructs tested at each difficulty level of the assessment.

## Scoring and Mapping to Other Proficiency Scales

When learners first set up their Voxy accounts, they are able to self-assess their proficiency level, which informs the difficulty level of the VPA that is assigned to each individual learner. The results of this test ultimately determine the difficulty level of the learner's course.

Each test item, which has four answer options and an additional "I don't know" option, is weighted according to its relative level of difficulty as determined by a Rasch analysis. Easier items are weighted the least and the hardest items are weighted the most, so a learner's performance on the more difficult items has a relatively greater influence on the final score than his or her performance on the less difficult items. The "I don't know" option allows learners to clearly indicate what they do not know instead of guessing and possibly being placed in a level does not accurately represent their true proficiency. In addition, there is a test for 0-beginners, or true novices, that is much shorter and does not require them to answer questions that will be too difficult.

A learner's final score on a VPA is mapped either to one of Voxy's 15 levels, or to the proprietary scale of one of Voxy's partner institutions. Voxy's 15 levels are grouped into seven different proficiency bands:

Beginner, High Beginner, Low Intermediate, Intermediate, High Intermediate, Low Advanced, and Advanced. Because partner institutions often have in-house, custom proficiency scales with more or fewer levels, the VPA score can be calibrated to the levels of any institution to allow Voxy's content, instruction, and assessment tool to reflect any partner's own proficiency levels.

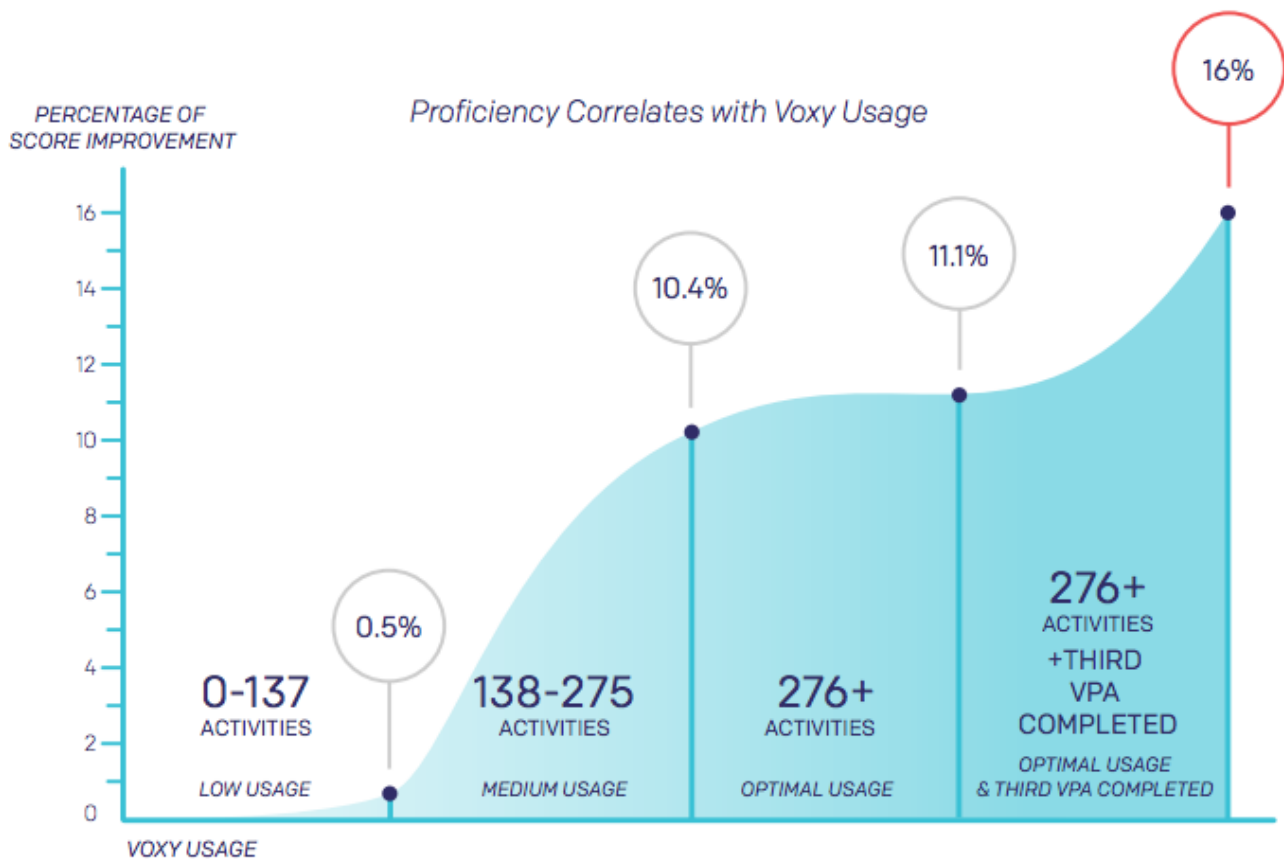
In order to allow Voxy learners to understand what their VPA scores mean in a broader context, all scores are mapped to various widely-known proficiency scales. This was done by mapping Voxy's Can-Do Statements to those scales' Can-Do Statements and/or scoring rubric descriptions. Learners can see how their VPA scores are aligned to The Common European Framework of Reference (CEFR)—probably the most familiar scale to learners worldwide—as well as the Global Scale of English (GSE). VPA scores are also aligned to the scales of high-stakes tests like the TOEFL, IELTS, and TOEIC.

## Research that Uses the VPA

The VPA has provided not only a way to measure global proficiency and proficiency improvement over time for Voxy learners, but also an opportunity for Voxy to conduct research regarding learner performance, particularly in a technology-mediated learning setting. One particular [study](#) was conducted to examine the impact blended learning has on learner performance and engagement. After considering two different groups—one that participated in private tutoring sessions and one that did not—it was clear that Voxy learners who worked with tutors demonstrated an overall increase in their post-test scores compared to those who did not work with tutors. Furthermore, the higher the number of private tutoring sessions, the higher the post-test scores. These findings have served as supporting evidence that synchronous instruction in addition to asynchronous study is most beneficial, and has encouraged many learners to engage with Voxy more in this way.

VPA data have also been used to show what may predict proficiency gain, maintenance, or loss over time. For this [study](#), the types of Voxy resources learners engaged with were examined to determine if the number of resources they completed could predict the outcomes of a second VPA. A follow-up study is underway using a more robust data set that will provide more insight into how this may be accomplished.

In addition to contributing to online testing research, VPA data are also used to determine how much Voxy usage is ideal for learners to improve their proficiency over time. Early VPA and Voxy engagement data showed that on average, learners had to complete at least 276 activities (the equivalent of 92 lessons) to improve their proficiency level. Internal research conducted since then continues to confirm that this level of Voxy usage is ideal if learners want to see improvement. A recent study that examined existing VPA data revealed that for 870 Voxy learners who took two VPAs, scores increased by 11% when they completed at least 276 activities. Of these learners, 147 went on to take a third VPA and increased their scores by 16%. Figure 1 shows mean score improvement for learners based on the amount of Voxy usage.



*Learners who complete at least 276 activities and take three Voxy Proficiency Assessment™ exams at regular intervals see an average score improvement of 16%.*

Figure 1. Mean score improvement

## Future Research

The VPA and other VPA related features are continuously being revised and improved based on new research findings and the needs of Voxy learners. For example, Voxy's original 7-level proficiency scale was recently expanded to 15 levels so that learners can see their progress more quickly. While moving from Level 1 to Level 2 can happen relatively quickly, once learners reach intermediate proficiency, more hours are required to progress from one level to the next. The new scale was created to make this progress more transparent to learners. With this new scale, each proficiency band has multiple sub-levels within it in order to better demonstrate progress to both learners and partner institutions.

While using the VPA to show learners how they are progressing is important, it is also crucial for the VPA to serve as a teaching tool. In the future, every distractor for each MC item will be weighted depending on how close it is to the correct answer (Rasch, 1961). Doing this will allow Voxy to give learners valuable feedback on what exactly they need to review based on how much difficulty they had with certain skills. In addition, a future version of the VPA will include more specific, granular feedback on the steps learners need to take to improve. For instance, instead of telling learners to simply spend more time in the Grammar Guide, the new VPA results page will recommend that they "focus on the present progressive tense" or "review the difference between adverbs and adjectives."

In addition to these VPA changes, more research that uses VPA data will be conducted. A major study will examine how language learners in a traditional classroom setting, blended learning environment, or online-only class improve their proficiency. The findings of this study will not only be able to further inform product improvements for Voxy, but also help higher education institutions, private language institutions, and corporations make more informed decisions about how best to provide their students or employees with language education.

## APPENDIX. GRAMMAR CONSTRUCTS



	EASY	MEDIUM	HARD	VERY HARD
<b>ADJECTIVES</b>	<i>Comparative</i>	<i>Quantifier</i>	<i>Adjective Order</i>	<i>Adjective Order</i>
	<i>Simple</i>			<i>Comparative</i>
	<i>Superlative</i>			
<b>ADVERBS</b>	<i>Degree</i>	<i>Indefinite Frequency</i>	<i>Degree</i>	<i>Conjunctive</i>
				<i>Degree</i>
				<i>Indefinite Frequency</i>
				<i>Simple</i>
<b>ARTICLES</b>	<i>Definite</i>	<i>Definite</i>	<i>Definite</i>	
	<i>Indefinite</i>	<i>Zero</i>		
<b>CONDITIONALS</b>	<i>First</i>	<i>Second</i>	<i>Second</i>	<i>Third</i>
	<i>Zero</i>		<i>Third</i>	
<b>CONJUNCTIONS</b>	<i>Coordinating</i>	<i>Coordinating</i>	<i>Subordinating</i>	<i>Subordinating</i>
<b>NOUNS</b>	<i>Compound</i>		<i>Joint Possessive</i>	<i>Countable</i>
	<i>Countable</i>			
	<i>Uncountable</i>		<i>Uncountable</i>	
	<i>Plural Possessive</i>			
<b>PHRASES</b>	<i>Phrasal Verb</i>	<i>Phrasal Verb</i>	<i>Phrasal Verb</i>	<i>Phrasal Verb</i>
	<i>Simple</i>	<i>Simple</i>		<i>Simple</i>



	<b>EASY</b>	<b>MEDIUM</b>	<b>HARD</b>	<b>VERY HARD</b>
<b>PART OF SPEECH</b>		<i>Common Root Word</i>	<i>Common Root Word</i>	<i>Common Root Word</i>
<b>PREPOSITIONS</b>	<i>Location</i>	<i>Spatial Relationships</i>	<i>Direction</i>	<i>Place</i>
		<i>Movement</i>	<i>Place</i>	<i>Location</i>
			<i>Time</i>	
<b>PRONOUNS</b>	<i>Pronoun Antecedent Agreement</i>	<i>Pronoun Antecedent Agreement</i>	<i>Pronoun Antecedent Agreement</i>	<i>Pronoun Antecedent Agreement</i>
	<i>Relative</i>	<i>Relative</i>	<i>Relative</i>	<i>Relative</i>
<b>QUESTIONS</b>		<i>Compound Tenses</i>	<i>Tag</i>	<i>Indirect</i>
		<i>Negative</i>		<i>Negative</i>
		<i>Tag</i>		
<b>TENSES</b>	<i>Present Progressive</i>	<i>Present Perfect Progressive</i>	<i>Past Progressive</i>	<i>Simple Future</i>
	<i>Simple Past</i>			
	<i>Simple Present</i>	<i>Simple Past</i>		
	<i>Simple Future</i>			
<b>VERBS</b>	<i>Gerund</i>	<i>Gerund</i>	<i>Gerund</i>	<i>Gerund</i>
	<i>Infinitive</i>	<i>Infinitive</i>	<i>Past Participle</i>	<i>Infinitive</i>
	<i>Present Participle</i>	<i>Past Participle</i>	<i>Present Participle</i>	<i>Past Participle</i>
		<i>Present Participle</i>		<i>Present Participle</i>
<b>VOICE</b>	<i>Active</i>	<i>Active</i>	<i>Passive</i>	<i>Active</i>
		<i>Passive</i>		<i>Passive</i>