The Language Experience and Proficiency Questionnaire (LEAP-Q): Ten years later

Margarita Kaushanskaya¹, Henrike K. Blumenfeld² and Viorica Marian³

¹University of Wisconsin-Madison; ²San Diego State University and ³Northwestern University

Abstract

The Language Experience and Proficiency Questionnaire (LEAP-Q) is a validated questionnaire tool for collecting self-reported proficiency and experience data from bilingual and multilingual speakers ages 14 to 80. It is available in over 20 languages, and can be administered in a digital, paper-and-pencil, and oral interview format. The LEAP-Q is used by researchers across various disciplines (Psychology, Neuroscience, Linguistics, Education, Communication Sciences & Disorders, etc.) to provide a comprehensive description of their bilingual participants, to substantiate a division of bilinguals into groups (e.g., early vs. late bilinguals), and to screen participants for adequate or threshold levels of language proficiency. Best practices for using the LEAP-Q include administration of the full questionnaire, consideration of acquisition and history of language use together with self-ratings of proficiency, and supplementation of self-reported data with objective language measures whenever possible.

The LEAP-Q can be downloaded at no cost at https://bilingualism.northwestern.edu/leapq/.

How does it work?

The LEAP-Q enables collection of self-reported language proficiency and experience data for any number of languages spoken by an individual. The LEAP-Q was designed to collect both broad measures of language dominance, language exposure, and language preference, and specific measures associated with each language (ages of acquisition and ages of attained fluency; length of immersion in different contexts; estimates of proficiency in speaking, reading, and understanding; ratings of how different contexts contribute to the acquisition of the language; extent of exposure to the language in different contexts; and degree of accent). See Figure 1 for a visual schematic. An 11-point (0–10) Likert scale is used for all the questions requiring estimates of degree and strength, with each point of the scale anchored to a descriptive label. The digital version of the LEAP-Q can be completed in 15 minutes by speakers of two languages. Knowledge of another language adds approximately 5 minutes to the completion of the questionnaire. The administration of the questionnaire as an oral interview extends administration time by 5–10 minutes.

The LEAP-Q was validated in two studies with two different samples of bilingual speakers. In Study 1, the LEAP-Q was administered to 52 multilingual speakers with diverse language backgrounds, representing 34 different languages. The internal validity of the questionnaire was tested with a visual schematic. An 11-point (0–10) Likert scale is used for all the questions requiring estimates of degree and strength, with each point of the scale anchored to a descriptive label. The digital version of the LEAP-Q can be completed in 15 minutes by speakers of two languages. Knowledge of another language adds approximately 5 minutes to the completion of the questionnaire. The administration of the questionnaire as an oral interview extends administration time by 5–10 minutes.

The LEAP-Q was validated in two studies with two different samples of bilingual speakers. In Study 1, the LEAP-Q was administered to 52 multilingual speakers with diverse language backgrounds, representing 34 different languages. The internal validity of the questionnaire was tested with a visual schematic. An 11-point (0–10) Likert scale is used for all the questions requiring estimates of degree and strength, with each point of the scale anchored to a descriptive label. The digital version of the LEAP-Q can be completed in 15 minutes by speakers of two languages. Knowledge of another language adds approximately 5 minutes to the completion of the questionnaire. The administration of the questionnaire as an oral interview extends administration time by 5–10 minutes.
was established via a factor analysis that revealed logical clusters of questions reflecting fundamental bilingual dimensions (native language competence; second language competence; etc.). In Study 2, the LEAP-Q was administered to 50 bilingual speakers of English and Spanish. The internal validity of the questionnaire was replicated with this more homogenous sample of bilingual speakers. More importantly, criterion-based validity was established by confirming a relationship between self-reported data on the LEAP-Q and performance on objective, behavioral speech and language measures administered in the bilinguals’ two languages (Reading Fluency, Oral Comprehension, Passage Comprehension, Productive Vocabulary, and Sound Awareness subtests of the Woodcock-Johnson/ Woodcock-Muñoz Tests of Achievement; Peabody Picture Vocabulary Test/ Test de Vocabulario en Imágenes Peabody; and grammaticality judgment tasks).

Correlation analyses linked self-ratings of proficiency in the L1 and the L2 to performance scores on speech and language measures. The general pattern of correlations indicated stronger relationships for the L2 than for the L1, most likely because of the wider range of values for the L2, both in the self-reported proficiency data and in the behavioral performance data. Broad measures of language ability (reading and oral comprehension) correlated more strongly with the self-ratings of proficiency than more specific measures of language ability (vocabulary; sound awareness; etc.). This indicates that bilinguals’ self-ratings of their proficiency reflect gestalt approximations of language skill. In the L1, self-ratings of reading proficiency were the strongest predictors of performance on speech and language measures. Conversely, in the L2, self-ratings of speaking proficiency were the strongest predictors of performance on speech and language measures.

How is it used?

Most frequently, the LEAP-Q is used by researchers to provide a comprehensive description of their bilingual participants (e.g., Ettlinger et al., 2015; Leonard, Torres, Travis, Brown, Hagler Jr, Dale, Elman & Halgren, 2011), and to substantiate a division of bilinguals into groups and subgroups. For example, LEAP-Q data often form the basis for assigning and/or confirming the assignment of bilinguals to early/simultaneous vs. late/sequential bilingual groups (e.g., Dimitropoulou, Duñabeitia & Carreiras, 2011; Pelham & Abrams, 2014; Shi, 2010). The LEAP-Q is also used by researchers to screen their bilingual participants for adequate levels of language proficiency (e.g., Conrad et al., 2011), to confirm native-speaker status (e.g., Hespos & Piccin, 2009), high levels of language proficiency (e.g., Lidií, Palmer, Peretz & Morningstar, 2011), or to document differences in L1 vs. L2 language skills (e.g., Mor, Yitzhaki-Amsalem & Prior, 2014). In some cases, it is required that threshold levels of proficiency be established for participating in a study. For instance, in studies requiring that only highly proficient bilinguals complete the experimental tasks, only participants who report proficiency levels of greater or equal to 7 (Stocco & Prat, 2014) or 8 (Krizman, Marian, Shook, Skoe & Kraus, 2012) may be recruited.

At the same time, in a world where monolingualism is becoming less common, the LEAP-Q is increasingly often used to institute threshold levels of knowledge in a second language in defining monolingual participants. As is the case with using threshold levels to document high language proficiency, the levels used to document low levels of language proficiency also fluctuate. Thus, in some studies, a proficiency rating of 3 or below qualifies participants as monolinguals (e.g., Pelham & Abrams, 2014; Pelham & Abrams, 2014; Shi, 2010).

Who is it for?

The original LEAP-Q was designed to be used in research settings, and to be administered to adult and adolescent bilingual and multilingual speakers representing a wide variety of language experiences and proficiency levels. In published studies, the youngest participants who successfully completed the LEAP-Q were 14 years of age (Krizman, Skoe, Marian & Kraus, 2014), and the oldest participants were 80 years of age (Ansaldo, Ghazi-Saidi & Adrover-Roig, 2015; Blumenfeld, Schroeder, Bobb, Freeman & Marian, 2016; Schroeder & Marian, 2012). The 11-point rating scales make the LEAP-Q exquisitely sensitive to fluctuations in bilingual experience, and as a result, it can be successfully completed by highly proficient bilingual speakers (e.g., Conrad, Recio & Jacobs, 2011; Mercier, Pivneva & Titone, 2014; Pelham & Abrams, 2014), as well as by very inexperienced L2 learners (Ettlinger, Morgan-Short, Faretta-Studenberg & Riera, 2017). Recently, the LEAP-Q was modified to enable collection of language background data from parents regarding their children (Rochanavibhata & Marian, in preparation). All translations and adaptations of the LEAP-Q are available on https://bilingualism.northwestern.edu/leapq/. It is important to note that only the original LEAP-Q was validated against objective measures of speech and language ability.
of proficiency understanding than speaking L2. Such discrepancies in proficiency ratings may pose challenges to aggregating across proficiency scores. While we ultimately leave the choice of a particular analytical strategy to each individual researcher, we recommend that researchers conduct factor or correlational analyses of their LEAP-Q data before combining responses obtained from the LEAP-Q into a single score. It is appropriate to combine variables into a single index only when they correlate with each other (e.g., Rubin, 2012). Across the two experiments in the initial validation study, we found that although general constructs captured by the LEAP-Q were stable, there was also variability in how the questions clustered together depending on the specific characteristics of the bilingual sample.

Can it be modified?

Because only the original LEAP-Q was validated against behavioral measures of speech and language ability, our recommendation is to administer the LEAP-Q in its entirety. We therefore discourage researchers from changing the wording of the questions, changing the question order, or inserting / deleting questions from the LEAP-Q. However, we encourage researchers to modify the LEAP-Q in ways that suit their purposes by adding questions at the end of the questionnaire. Several such modifications have been implemented. For example, Libben and Titone (2009) modeled their proficiency questions after the questions on the LEAP-Q, and used the 11-point proficiency scale to also collect bilinguals’ self-ratings of their translating ability, grammatical ability, and fluency. Reichle (2010) adapted the LEAP-Q to probe for pronunciation ability, in addition to speaking, understanding, and reading proficiency. Prior (2012) added questions about parental education to the original LEAP-Q in order to obtain information about participants’ socioeconomic status. The LEAP-Q was also adapted to collect language proficiency and exposure data in a trilingual individual with aphasia before and after the stroke (Faroqi-Shah & Waked, 2010).

Best practices for using the LEAP-Q

Because self-reported proficiency data collected via the LEAP-Q correlate with behavioral measures of speech and language performance (as reported in the original study), it is tempting to rely solely on bilinguals’ self-reported proficiency when describing a bilingual sample. Self-reported proficiency data alone can sometimes meet the needs of a particular study (e.g., Athanasopoulos, Damjanovic, Burnand & Bylund, 2015; Garbin, Costa, Sanjuan, Forn, Rodriguez-Pujadas, Venura, Belloch, Hernandez & Avila, 2011; Martin, Strijkers, Santestaban, Escera, Hartsuiker & Costa, 2013; Sullivan & Schatz, 2009). However, a comprehensive approach to assessing bilinguals is more likely to yield a reliable picture of bilinguals’ language profiles. In the case of the LEAP-Q, the correlations between self-reported proficiency levels and performance on behavioral speech and language measures were moderate-to-strong for the L2, and weak-to-moderate for the L1. In the regression models where the behavioral measures of speech and language were used to predict self-rated levels of proficiency, the $R^2$ values ranged from 0.70 to 0.23: that is, there was not a perfect alignment between self-ratings of proficiency and objective measures of speech and language ability. A study by Shi (2011) illustrates this point.

Shi (2011) tested the ability of LEAP-Q proficiency ratings to yield reliable estimates of bilinguals’ performance on an English...
listening task. Those bilinguals who self-rated their English proficiency as 7 or above (highly proficient) tended to perform most similarly to the monolinguals on the listening task. However, there was a large number of false-positive cases — i.e., many participants rated themselves as highly proficient while performing rather poorly on the listening task. Crucially, consideration of other aspects of history (language dominance and age of acquisition) improved upon the use of proficiency scores in assigning bilinguals into groups that would align with their behavioral performance.

In the original publication, we suggested that researchers use aspects of bilinguals’ language acquisition history (and not just self-ratings of proficiency) to characterize the bilingual participants. We continue to believe that this is the best practice. Furthermore, in our own work, we administer the LEAP-Q to characterize our bilingual participants, but whenever possible, we use objective metrics of language performance to confirm the proficiency or dominance profiles revealed by the LEAP-Q. We strongly encourage other researchers to do the same.

Looking to the future

It is now standard to include measures of proficiency and experience in any research protocol that involves bilingual participants. Over the past 10 years, the LEAP-Q has been employed as such a measure across a wide range of academic disciplines, ranging from psycholinguistic studies (e.g., Bartolotti & Marian, 2012; Blumenfeld, Bobb & Marian, 2016; Dimitropoulou et al., 2011; Mercier, Pivneva & Titone, 2014; Shook & Marian, 2012; Van Engen, 2010; Whitford & Titone, 2012) to neuroimaging (e.g., Conrad et al., 2011; Leonard et al., 2011; Marian et al., 2014; Reichle & Birdsong, 2014; Scherer, Fonseca, Amiri, Adrover-Roig, Marcotte, Giroux, Senhadji, Benali, Lesage & Chahi, and Bialystok (2018), and the Language Exposure and Social Background Questionnaire by Anderson, Mak, Conrad et al., 2011; Leonard et al., 2011; Marian et al., 2014; Mercier, Pivneva & Titone, 2014; Shook & Marian, 2012; Van Engen, 2010; Whitford & Titone, 2012) to neuroimaging (e.g., Conrad et al., 2011; Leonard et al., 2011; Marian et al., 2014; Reichle & Birdsong, 2014; Scherer, Fonseca, Amiri, Adrover-Roig, Marcotte, Giroux, Senhadji, Benali, Lesage & Ansaldo, 2011; Yi, Maddox, Mumford & Chandrasekaran, 2014) and clinical research (e.g., Farooqi-Shah & Waked, 2010; Mor et al., 2014; Summers, Gonzales & Pechak, 2015). Together with other questionnaire tools such as the Language History Questionnaire by Li and colleagues (2006, 2014), the Language and Social Background Questionnaire by Anderson, Mak, Chahi, and Bialystok (2018), and the Language Exposure Assessment Tool for children by DeAnda, Bosch, Poulin-Dubois, Zesiger, and Friend (2016), the LEAP-Q has become part of the standard battery of background assessments in bilingualism research.

To date, the original LEAP-Q article has been cited over 800 times (see google scholar metrics at goo.gl/eDJxhn). As the field moves forward, we envision that the LEAP-Q will continue to play a central role in multifaceted examinations of the bilingual experience. It is likely that the use of LEAP-Q data will shift in step with the re-conceptualizations of bilingualism as a continuum (e.g., Luk & Bialystok, 2013; Dunn & Tree, 2009), and with changing trends in the research questions posed within the broad realm of bilingualism research. For example, specific aspects of bilingual proficiency and experience obtained from the LEAP-Q are used increasingly in a graded manner, often serving as continuous predictors of bilinguals’ performance (e.g., Krizman et al., 2014; Mercier et al., 2014; Reichle & Birdsong, 2014). This approach, when appropriate, may ultimately yield more consistent findings across studies and populations.

Although the existing version of the LEAP-Q and its many translations and adaptations can continue to serve the needs of many researchers for some time, it is important to recognize the need to re-norm the questionnaire in the future, as well as to update it in relevant ways. For example, possible updates to the LEAP-Q that we are currently considering include: addition of questions regarding code-switching practices and exposure to code-switching; addition of questions regarding language exposure through internet and media; and a validation of empirically-derived thresholds in proficiency, exposure, and ages-of-acquisition that would reliably separate bilinguals into sub-groups. Any of these, or other changes, would require another validation of the questionnaire against objective measures of language ability, as was done for the original LEAP-Q.

In conclusion, together with other instruments, the LEAP-Q can be a helpful tool to bilingualism researchers in identifying profiles of language proficiency and exposure that influence linguistic and cognitive processes and their neural underpinnings. As the LEAP-Q continues to be used, translated, adapted, and modified by scientists to meet their research needs, we will continue to maintain a publicly available repository of all LEAP-Q versions so that we can contribute to validation and replicability efforts in the bilingualism research community.

Author ORCIDs.

Heinrike K. Blumenfeld, 0000-0002-0538-5339

Acknowledgements.
The present project was supported by NIDCD Grant R01 DC011750 to Margarita Kaushanskaya, and Grant NICHD R01 HD059858 to Viorica Marian. The authors wish to express gratitude to all of the researchers who have adopted, adapted, and translated the LEAP-Q, and to the participants who provided the norming data. We would also like to thank everyone who has assisted with data coding, data analyses, data management, and email correspondence related to the LEAP-Q.

References


Downloaded from https://www.cambridge.org/core. Northwestern University Library, on 08 Jul 2019 at 17:46:17, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/S1366728919000038


Hespou S and Piccin TB (2009) To generalize or not to generalize: spatial categories are influenced by physical attributes and language. Developmental Science 12, 88–95.


## Appendix A. Translations of the LEAP-Q that are currently available

All versions can be accessed at [https://bilingualism.northwestern.edu/leapq/](https://bilingualism.northwestern.edu/leapq/)

<table>
<thead>
<tr>
<th>Language</th>
<th>Translators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Arabic</td>
<td>University of Maryland’s Center for Advanced Study of Language, CETRA Language Solutions, and Faiza Sultan, President of Translation4all, Inc.</td>
</tr>
<tr>
<td>Catalan</td>
<td>Eloi Puig Mayenco and Susagna Tubau Muntanillá, Universitat Autònoma de Barcelona</td>
</tr>
</tbody>
</table>
| Dutch    | *For the Netherlands*: Lisa Vandeberg, Erasmus University, adapted by Marilyn Hall, Northwestern U.  
  *For Belgium*: Freya De Keyser, Ghent University, and Marilyn Hall, Northwestern University |
| English  | Marian, Blumenfeld & Kaushanskaya, 2007  
  *Pencil-and-paper version*: Marilyn Hall, Northwestern University |
| Farsi    | Mahrokh Alamzadeh, Dr. Ali Ghanaie, & Dr. Shahla Sharifi, Mashhad Ferdowsi U. |
| Filipino/Tagalog | Maria Khristina Manuell, Theologische Hochschule Reutlingen  
  *Pencil-and-paper version*: İlker Güzelordu, Eastern Mediterranean University |
| Finnish/Suomalainen | Ari Huta and Mika Läheennäki, University of Jyväskylä |
| French   | *For France*: Anjali Bhattacharjee, Emile Michaud, and Judith Gervain, Université Paris Descartes and CNRS  
  *For Belgium*: Arnaud Szmalec, Université Catholique de Louvain  
  *For Canada*: Zvaigne, Salem, Grobleau, & Millette, McGill University  
  *For Switzerland*: Mehdi Purmohammad, Universität Bern, & Max R. Freeman, Northwestern U. |
| German   | Larissa Weigel and Monica Gonzalez-Marquez, Cornell University |
| Hebrew   | Anat Prior, University of Haifa |
| Hungarian| Timea Kutasi, University of Edinburgh, UK, & Zsuzsanna Maté |
| Italian  | Luca Bevacqua and Roberta Spelborz, University of Edinburgh |
| Japanese | Koji Miwa and Yoshino Okuma, University of Alberta, & Yu Ikemoto, Kwansei Gakuin University |
| Korean   | Ju Young Min and Luana Onnis, University of Hawaii |
| Malay    | Tze Peng Wong, University of Nottingham Malaysia, and Rosyati M. Yaakub |
| Mandarin | Mahire Yakup, Jun Wang, and Trenton Wilson, U. of Kansas |
| Portuguese| Ana Paula Scholl and Ana Beatriz Areas da Luz Fontes, Federal University of Rio Grande do Sul |
| Romanian | Ioana Tufar and Rebeca Ciupu, Babes-Bolyai University |
| Russian  | Marina Belkina, Katy Borodkin, Olga Iukalo, and Mira Goral, Lehman College, City University of New York, adapted by Marilyn Hall, Northwestern University |
| Spanish  | *For the US*: Rojas & Iglesias, Temple University  
  *For Spain (peninsular Spanish)*: Eloi Puig Mayenco, Universitat Autònoma de Barcelona, and Jason Rothman and Jorge González Alonso, University of Reading & the Arctic University of Norway |
| Thai     | Sirada Rockanivibhata, Northwestern University (including pencil-and-paper and child versions) |
| Turkish  | Özlem Yüksel-Sökmen, The Graduate Center of the City University of New York, and Sercan Şerifoğlu, Haliç University, Istanbul  
  *Pencil-and-paper version*: İlker Güzelordu, Eastern Mediterranean University |
| Vietnamese| Hien Pham, University of Alberta, Canada, & Quyen Nguyen, Vietnam National U., Hanoi |