

The language and social background questionnaire: Assessing degree of bilingualism in a diverse population

John A. E. Anderson¹ · Lorinda Mak¹ · Aram Keyvani Chahi¹ · Ellen Bialystok¹

Published online: 9 March 2017
© Psychonomic Society, Inc. 2017

Abstract Research examining the cognitive consequences of bilingualism has expanded rapidly in recent years and has revealed effects on aspects of cognition across the lifespan. However, these effects are difficult to find in studies investigating young adults. One problem is that there is no standard definition of bilingualism or means of evaluating degree of bilingualism in individual participants, making it difficult to directly compare the results of different studies. Here, we describe an instrument developed to assess degree of bilingualism for young adults who live in diverse communities in which English is the official language. We demonstrate the reliability and validity of the instrument in analyses based on 408 participants. The relevant factors for describing degree of bilingualism are: (1) the extent of non-English language proficiency and use at home, and (2) non-English language use socially. We then use the bilingualism scores obtained from the instrument to demonstrate their association with: (1) performance on executive function tasks, and (2) previous classifications of participants into categories of monolinguals and bilinguals.

Keywords Bilingualism · Latent-factor-analysis · LSBQ · Language and social background questionnaire

In recent years, there has been an enormous increase in the quantity of research investigating the effect of bilingual experience on cognitive and linguistic processing across the lifespan (for review, see Bialystok, Craik, Green, & Gollan,

2009). This body of research has both reversed earlier beliefs that warned of dire consequences of bilingualism for children (for review, see Hakuta, 1986) and contributed to the growing evidence for profound effects of experience on brain and cognition (Kolb et al., 2012; Pascual-Leone, Amedi, Fregni, & Merabet, 2005). With estimates that at least half of the world's population is bilingual (Grosjean, 2013), the possibility that there are implications of this normal experience for cognitive and linguistic functioning is important. However, the research is complex and the outcomes are often inconsistent, with some studies concluding that bilingualism has no effect on cognitive systems (e.g., Gathercole et al., 2014, Paap & Greenberg, 2013). Part of the inconsistency in research outcomes is because of a lack of clear methodological approaches to these issues.

A significant impediment to progress in investigations of the cognitive consequences of bilingualism is the absence of a common standard for determining how to describe individuals in terms of this complex experience. Different research groups apply different criteria to assign participants to monolingual or bilingual groups or to determine a quantitative assessment indicating the degree of bilingualism. Variations in these criteria make it difficult to compare research across groups and are in part responsible for the recent emergence of contradictory findings in some of this research (for review, see Paap, Johnson, & Sawi, 2015; for discussion, see Bialystok, 2016). Therefore, it is important to articulate criteria by which individuals are considered to be monolingual or bilingual. Most research on bilingualism uses some form of self-report questionnaire to gather information relevant to this designation, but the design and interpretation of such instruments are vague because bilingualism is a multifaceted experience shaped by social, individual, and contextual factors. Language experience lies on a continuum: Individuals are not categorically “bilingual” or “monolingual” (Luk & Bialystok, 2013). Therefore, there is a need for an

✉ Ellen Bialystok
ellenb@yorku.ca

¹ Department of Psychology, York University, 4700 Keele Street, Toronto, Ontario M3J 1P3, Canada

evidence-based instrument with high reliability and validity that captures relevant aspects of this multidimensional experience that can be implemented widely.

An obvious solution to the problem of describing an individual's bilingualism is to administer language tests in both languages and devise a score that captures both the absolute and relative proficiency levels of each. This method has been used effectively by some researchers (e.g., Bialystok & Barac, 2012; Gollan, Weissberger, Runngvist, Montoya, & Cera, 2012; Sheng, Lu, & Gollan, 2014). However, the limitation of this approach is that it must be possible to identify and to test proficiency in both languages. In much of the USA, this is feasible because the majority of bilinguals who participate in bilingualism research in that country speak Spanish and English, and many standardized tests for Spanish proficiency are available. Similarly, bilingualism research conducted in Barcelona typically relies on Spanish-Catalan bilinguals, and again, objective testing in both languages is an option. Outside of these examples, however, the situation is more complicated. In a diverse population, the range of second languages can be overwhelming and impossible to assess through formal testing. Moreover, merely assessing proficiency in both languages fails to capture the full complexity of the bilingual experiences. In short, this type of assessment cannot reveal the extent and pattern of use of each language.

Resolving the methodological problem of how to characterize participants on the multidimensional continuum of bilingual experience will contribute to theoretical discussions of the nature of experience, empirical interpretations of contradictory evidence, and measurement issues of quantifying language experience. Dong and Li (2015) showed that age of second language acquisition, proficiency, and frequency and nature of language use influenced whether monolingual and bilingual group differences emerged in cognitive performance. These are all plausible candidates for mediating the effect of bilingual experience on cognition, but there are no reliable means of quantifying these dimensions. How much second language experience is necessary? Does a late-life acquirer who practices diligently count (and how much)? How does one categorize a person who spoke two languages in childhood but has let one language lapse? How much weight should researchers put on each of these variables to determine the level of bilingualism? A sensitive instrument that provides a comprehensive description of bilingualism that applies to a broad range of languages and contexts is crucial. Such an instrument will clarify definitions of bilingualism and allow researchers to resolve conflicting results based on different interpretations of bilingualism.

Two existing instruments have been widely used to assess bilingualism – the Language Experience and Proficiency

Questionnaire (LEAP-Q; Marian, Blumenfeld, & Kaushanskaya, 2007) and the Language History Questionnaire (LHQ 2.0; Li, Sepanski, & Zhao, 2006; Li, Zhang, Tsai, & Puls, 2014). The LEAP-Q includes questions about the participants' language proficiency, language dominance, preference for each language, age of language acquisition, and current and past exposure to their languages across settings. To establish its internal validity, Marian et al. (2007) tested the LEAP-Q with 52 heterogeneous bilinguals and analyzed their responses using factor analysis with orthogonal rotations. The results revealed that levels of language proficiency, degrees of language use and exposure, age of second language acquisition, and length of second language formal education best explained these individuals' self-ratings of bilingual experience. Items related to the degree of comfort in using a language were unrelated to the assessment of bilingualism. Based on these results, the researchers shortened the questionnaire and tested the revised version for criterion-based validity by correlating self-report and standardized language proficiency measures in a new group of 50 Spanish-English bilingual college students. These analyses revealed similar factors and demonstrated strong correlations between the students' self-rated language proficiency responses and the behavioral measures of their language proficiency. The LEAP-Q is a useful tool for measuring participants' language status as it takes into account a broad range of language experience. However, there were over 70 items submitted in their analyses but only ~50 participants in each study, a cases-to-variables ratio that falls below those recommended to estimate factor analysis models, a ratio that ranges from 2:1 to 20:1 (Hair, Anderson, Tatham, & Black, 1995; Kline, 1979). With so few cases per variable, the resulting models are unstable and should be interpreted cautiously. A further limitation of the instrument is that the factors were assumed to be uncorrelated. However, factors describing most psychological traits are rarely process-pure, so solutions based on an orthogonal rotation may be misleading. Finally, while the questionnaire provided an important demonstration of the multifaceted nature of bilingualism and incorporated different aspects of language experience, the authors did not provide enough information for users of the questionnaire to definitively classify participants.

The Language History Questionnaire (LHQ) was created by Li et al. (2006) as a web-based generic language questionnaire for the research community. The researchers examined 41 published questionnaires and constructed an instrument consisting of the most commonly used questions in those studies. Items in the LHQ included questions about the participants' language history (e.g., age of second language acquisition and length of second language education), self-rated first- and second-language proficiency, and language usage in the home environment. The researchers tested the instrument with 40 English-Spanish bilingual college students and reported sound predictive validity and high

reliability (split-half coefficient at .85). Recently, Li et al. (2014) revised the web-interface of the LHQ to make the instrument more user-friendly. The latest version of the LHQ also allows researchers to select the length and the language of the questionnaire. However, a limitation of the LHQ is that it does not provide any supplementary information on how to interpret responses collected from the questionnaire. Researchers using the LHQ are therefore required to determine their own methods for participant classification.

The Language and Social Background Questionnaire (LSBQ) presented here shares many features with both the LEAP-Q and LHQ, but the differences are important. The LEAP-Q has more questions than the LSBQ, including questions that were excluded on the basis of our polychoric correlations. The means of asking the questions is also different; in the LEAP-Q language activities were judged individually and rated on a scale for their use of English and in the LHQ participants were asked to name the language they used for a particular activity. For these reasons, the LSBQ provides a more continuous assessment of bilingualism than do the other two instruments.

In addition to the instruments described above, research with homogenous groups of bilinguals has evaluated participants' language status by combining self-report, interviews, and behavioral methods of assessment. For example, Gollan et al. (2012) and Sheng et al. (2014) classified participants into language groups by administering a self-rated language proficiency questionnaire, interviews with participants, and picture naming tasks in both languages as a behavioral measure of language proficiency. This multi-measure approach of language status assessment has advantages over other systems since it incorporates both objective and subjective assessments. However, as discussed above, the approach is restricted to situations in which there is a limited range of languages in the sample that is known in advance and possible to assess through standardized measures. For studies involving a diverse group of bilinguals, such assessment procedures are unrealistic. Thus, an alternative universally applicable method is needed when working with heterogeneous groups of bilinguals who live in diverse communities where experiences vary widely.

With few studies that objectively examine how different aspects of language experience jointly constitute "bilingualism," a consensus on which questions are most informative is difficult to establish. Given the ambiguity surrounding the classifications of this crucial independent variable, it is not surprising that different research groups report different results. Indeed, the lack of a universally-applicable instrument for quantifying bilingualism has been raised as an important methodological issue (Calvo, Garcia, Manoilloff, & Ibanez, 2016; Grosjean, 1998). Resolving this classification issue is particularly important for studies involving young

adults because of the limited variance in performance on behavioral tasks for this age group and the reports of conflicting results. The goal of the present study is to present a valid and reliable measurement tool, the LSBQ, that can be used to quantify bilingualism, lead to evidence-based classifications into language groups, and be sensitive to the nature of bilingual profiles. The intended population for the LSBQ reported here is adults with varying degrees of language experience who live in diverse communities (see list of languages below). Other versions of the instrument have been designed for use with children or older adults but will not be discussed in the present report. Although all these instruments were designed for use in communities where English is the majority language, a different language could be substituted for English in a translated version of the questionnaire.

An initial report of an earlier version of the LSBQ was presented by Luk and Bialystok (2013). They conducted a factor analysis on responses to the LSBQ and scores on two standardized measures of English proficiency from 110 young adults. The authors reported that a two-factor model best described the participants' bilingual experience, namely, daily bilingual usage and language proficiency. There were, however, several limitations of that initial study. First, the analysis was based only on responses from bilinguals and so did not account for monolinguals or individuals with marginal second language experience. Second, the questions were limited in the range of language activities included; the current version of the LSBQ includes a detailed description of bilingual usage patterns with various interlocutors and across different settings and situations. Third, the researchers combined questions that they thought were conceptually similar prior to entering them into the factor analysis, possibly forcing the emergence of expected factors. For these reasons, a new study based on the current version of the instrument is required.

The purpose of the current study is to describe the LSBQ, report its validity and reliability, and provide an interpretation guide and recommended cut-off scores for the continuous outcome variable into categorical groups. To this end, the internal validity of the LSBQ was first established using exploratory factor (EFA) analysis with a large group of young adults. We hypothesized that items designed to measure a single construct would cluster together, yielding the underlying factor(s) that influence bilingualism. The goal is to derive a composite factor score that represents overall level of bilingualism. Critically, this composite score can be used as both a continuous variable and a criterion to define groups categorically. To assess construct-validity, we tested the relationship between bilingualism (as measured by the composite score) and executive control performance for a subset of our participants. We further wished to test whether the derived factor scores would reproduce our categorical assessment of bilingualism. In other words, would our factor scores similarly classify individuals we previously designated as "bilingual" or "monolingual"?

Method

Participants

Participants were recruited from multiple studies conducted by the fourth author between 2014 and 2015. All young adults who participated in a study that used the most recent revision of the Language and Social Background Questionnaire (LSBQ) were included in this study. In this way, data were obtained from 605 young adults (386 female, 213 male, six did not specify) who ranged in age from 16 to 44 years ($M = 21.00$, $SD = 3.56$). Socio-economic status (SES) was determined by parents' education on a scale from 1 to 5 ($M = 3.00$, $SD = 1.00$, range = 1–5), where 1 indicates some high school education; 2 indicates high school graduate; 3 indicates some post-secondary education; 4 indicates post-secondary degree or diploma; and 5 indicates graduate or professional degree. In our sample, 241 individuals were born in countries where English is not the dominant language, and 364 participants were born in Canada or a country where English is the dominant communicating language (e.g., UK, USA). 147 (24.3%) participants reported having no knowledge of any language other than English and 458 (75.7%) participants reported some knowledge of a non-English language. For those participants who reported having some knowledge of a language other than English, 50 languages were represented. The 28 most common languages (90% of the non-English languages in this sample) were: French (100), Farsi (41), Korean (34), Spanish (27), Russian (23), Italian (20), Arabic (17), Cantonese (16), Urdu (15), Gujarati (12), Hebrew (11), Punjabi (11), Portuguese (11), Tamil (9), Hindi (8), Vietnamese (7), Tagalog (6), Patois (5), Somali (5), Mandarin (5), Twi (4), Greek (4), Polish (4), Dari (4), American Sign Language (4), Japanese (4), Yoruba (3), and Swahili (3). Participants varied in the age at which they first learned their second language ($M = 6.23$ years, $SD = 5.16$, range = 0–37) and whether it was English or the non-English language.

Language and social background questionnaire

Supplemental information for this article including the LSBQ and a scoring spreadsheet can be found at the figshare link <https://dx.doi.org/10.6084/m9.figshare.3972486.v1>.

The Language and Social Background Questionnaire contains three sections (see <https://doi.org/10.6084/m9.figshare.3972486.v1>). The first section (pages 1–2), Social Background, gathers demographic information such as age, education, country of birth, immigration, and parents' education as a proxy for SES. The second section (pages 3–4), Language Background, assesses which language(s) the participant can understand and/or speak, where they learned the language(s) and at what age. There are also questions

assessing self-rated proficiency for speaking, understanding, reading and writing the indicated languages, where 0 indicates no ability at all and 100 indicates native fluency. Additionally, there are questions regarding the frequency of use for each language ranging from “None” (0) to “All” (4) of the time. The third section (pages 5–7), Community Language Use Behavior, covers language use in different life stages (infancy, preschool age, primary school age, and high school age), and in specific contexts, such as with different interlocutor (parents, siblings, and friends), in different situations (home, school, work, and religious activities), and for different activities (reading, social media, watching TV and browsing the internet). As well, there are questions regarding language-switching in different contexts. Participants' language usage was rated on a 5-point Likert scale where 0 represented “All English” and 4 represented “Only the other language,” 2 represented an equivalent use of English and the other language. Descriptive statistics for all 22 items of the LSBQ are reported in Table 1.

Results

Factor analysis

From an initial sample of 605 individuals, we determined that 197 did not fit the criteria for monolingualism or bilingualism because there was a significant presence of a third language that made those individuals multilingual. Inspection of the correlation matrices for these multilinguals revealed much weaker relationships between questionnaire items than was found for the rest of the sample, suggesting that different factors are needed to describe this more complex profile. A parallel analysis of these 197 participants using Factor Analysis and Principal Components to extract eigenvalues indicated that 7–10 factors were necessary to capture the variance in the multilinguals' responses (see Fig. 1) and a subsequent factor analysis for multilinguals failed to converge on a solution. As multilingualism is known to have a different impact on cognition than bilingualism (Calvo et al., 2016; Kave, Eyal, Shorek, & Cohen-Mansfield, 2008), and given that the LSBQ is designed to assess the depth of use of *only* one or two languages, multilinguals were excluded from the final dataset. Our final sample, therefore, included 408 participants. In this sample, the average age was 21.27 years ($SD = 3.55$, range = 17–39), average SES based on parents' education was 3.29 ($SD = 1.25$), 295 (64.95%) participants were born in a country where English was the majority language, and 143 (35.05%) participants were born in a country where a non-English language was dominant. Finally, 145 (35.54%) participants reported having no knowledge of any language other than English and 263 (64.46%) participants reported some knowledge of a non-English language.

Table 1 Descriptive statistics for language and social background items

Variable	Mean	Standard deviation	Observed minimum value	Observed maximum value
Language used with Grandparents	1.78	1.84	0	4
Language used in Infancy	1.7	1.72	0	4
Code switching with family	1.61	1.44	0	4
English understanding proficiency	95.06	10.24	0	100
Non-English language speaking proficiency	42.96	40.89	0	100
Language used with other relatives	1.54	1.62	0	4
Language used in preschool	1.52	1.58	0	4
Language used with parents	1.5	1.62	0	4
Non-English language listening frequency	1.3	1.36	0	4
Non-English language speaking frequency	1.13	1.27	0	4
Language used at home	1.3	1.45	0	4
Language used in primary school	1.2	1.13	0	4
Language used for religious activities	1.18	1.46	0	4
Language used with siblings	0.83	1.22	0	4
English listening frequency	3.55	0.65	1	4
Language used for praying	0.99	1.42	0	4
Language used in high school	0.86	0.86	0	4
English speaking frequency	3.57	0.67	1	4
Language used at work	0.21	0.53	0	3
Language used at school	0.18	0.46	0	3
Language used for health care, banks, government services	0.16	0.56	0	4
Language used for shopping, restaurants, commercial services	0.25	0.59	0	4
Language used for social activities	0.44	0.77	0	4
Language used for e-mailing	0.2	0.46	0	3
Language used with friends	0.5	0.81	0	4
Language used for extracurricular activities	0.33	0.67	0	4
Language used with room-mates	0.25	0.79	0	4
Language used for texting	0.43	0.74	0	3
Language used on social media	0.38	0.73	0	4
Language used for watching movies	0.46	0.74	0	3
Language used for browsing the internet	0.33	0.7	0	4
Code switching on social media	0.84	1.2	0	4
Language used with neighbors	0.24	0.69	0	4
Language used for watching TV/listening to radio	0.51	0.86	0	4
Language used for writing lists	0.27	0.73	0	4
Language used for reading	0.34	0.63	0	3
Language used with partner	0.48	0.97	0	4
Code switching with friends	1.14	1.28	0	4
Non-English language understanding proficiency	47.69	42.02	0	100
English reading proficiency	94.27	10.62	50	100
English writing proficiency	91.62	14.18	25	100
English speaking proficiency	93.59	12.46	0	100
English writing frequency	3.66	0.7	1	4

As a first step in running an EFA, a matrix of polychoric correlations (see figshare repository) was estimated between all possible pairs for 60 items of the LSBQ using 408 cases.

Polychoric correlations are an alternative to Pearson product-moment correlations that take into account relationships between continuous and discrete variables that are not normally

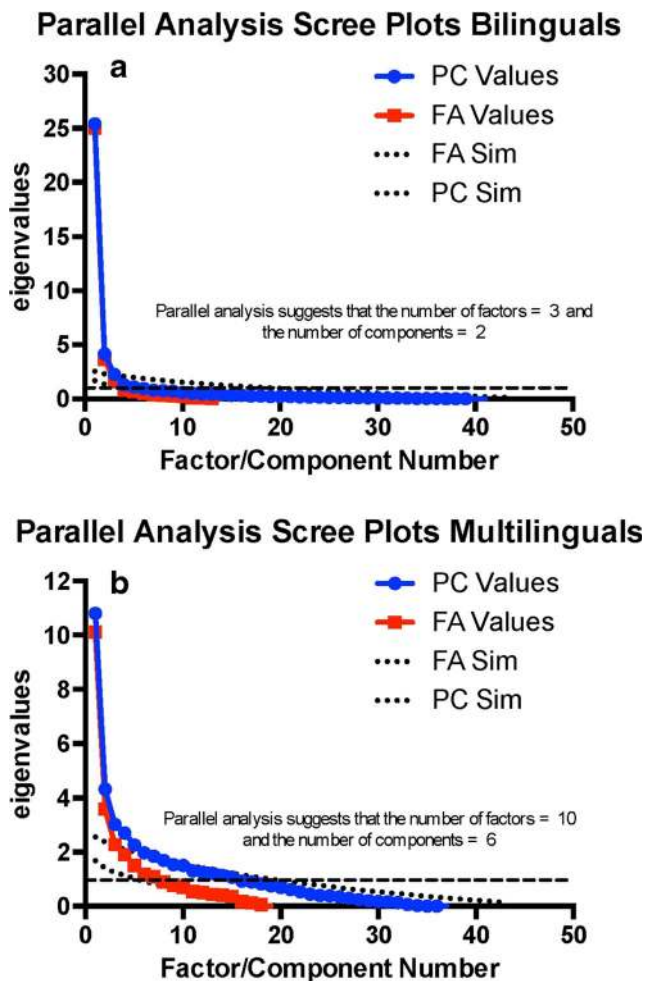


Fig. 1 Parallel analyses. (A) The three-factor solution presented in the text for bilinguals and monolinguals. (B) Greater variability in the multilinguals' responses (i.e., less variance is explained, and more factors/principal components are required)

distributed (Flora et al., 2003). An initial inspection of the resulting correlation matrix revealed that five items did not correlate well with others. For each of these items, more than 50% of their correlations with other items fell between $r = -0.3$ and $r = 0.3$. An additional 12 items were found to load equally on more than one factor following an initial analysis used to determine eigenvalues (the difference between the highest two factor loadings was $<.4$). Items in both these categories were removed before the final analysis leaving 43 items to be analyzed with an ordinary-least-squares minimum residual approach to EFA using an oblique rotation (promax). The Kaiser-Meyer-Olin measure verified the sampling adequacy for the analysis $KMO = .84$ ("meritorious" according to Kaiser, 1974), and all KMO values for the individual items were $>.72$, which is well above the acceptable limit of $.5$. Bartlett's test of sphericity $\chi^2(903) = 27775.26$, $p < 0.001$, indicated that correlations between items were sufficiently large for a factor analysis.

An initial analysis was run to obtain eigenvalues for each factor. Three factors had eigenvalues over Kaiser's criterion of 1 and in combination explained 74% of the variance. The parallel analysis and scree plot was slightly ambiguous and showed inflexions that would justify retaining either two or three factors (see Fig. 1). Given the large sample size (408), and the convergence of the scree plot and Kaiser's criterion on three factors, all three factors were retained. Table 2 shows the factor loadings after rotation for the final analysis, and a graphic representation of this information is presented in Fig. 1. Inspection of the clustering items suggests that Factor 1 represents "Non-English Home Use and Proficiency," Factor 2 represents "Non-English Social Use," and Factor 3 represents "English Proficiency." Separate reliability analyses were conducted for each factor using the raw data and the polychoric correlation matrix as input. Both sets of values are presented in Table 2.

Correlations with cognitive scores

The data reported in the present paper were collected between 2014 and 2015. Every study included the LSBQ, and most studies included measures of verbal and nonverbal intelligence through either the Shipley Institute of Living Test (Zachary & Shipley, 1986) or the Kaufman Brief Intelligence Test (KBIT, Kaufman & Kaufman, 2004). One study also included verbal and non-verbal variants of the arrow flanker task (Eriksen & Eriksen, 1974) in which participants are asked to respond to a central stimulus while ignoring irrelevant flanking stimuli. This procedure creates congruent and incongruent trials that can be assessed for accuracy and reaction time. Scores from all these measures were used in the next analyses (Fig. 2).

Following derivation of the factor structure, factor scores were calculated for each participant for each of the three factors by standardizing raw scores and multiplying these by the factor weights. The weighted standardized scores were then summed to produce a factor score for each participant on each factor. Where an item loaded on two factors, the stronger association was retained. In addition to individual factor scores, a composite score was computed by summing the factor scores weighted by each factor's variance.

The factor scores were then correlated with previously collected cognitive data for a subset of the participants (see Fig. 3B for the number of participants in each analysis). Figure 3A presents the scatterplots showing the relationship between the composite factor score and the behavioral outcome measures. Outcome measures were verbal or non-verbal, and the general pattern is that increasing bilingualism as indicated by higher composite factor scores predicted better non-verbal performance and poorer verbal performance. We then tested the difference between slopes for the verbal and non-verbal components of the Shipley and Flanker tasks using

Table 2 Factor analysis results (n=408): (A) Pattern matrix, (B) structure matrix

Item	Promax rotated factor loadings		
	Non-English Home use and proficiency	Non- English social use	English use
A			
Grandparents	1.04	-0.11	0.01
Infancy	1	-0.09	-0.04
Switching with family	1	-0.26	0.28
Non-English understanding	0.99	0.08	0.18
Non-English speaking	0.96	0.16	0.2
Relatives	0.95	0.02	0.06
Preschool	0.94	-0.07	-0.11
Parents	0.92	0.01	-0.07
Non-English listening frequency	0.89	0.03	0.1
Non-English speaking frequency	0.87	0.11	0.08
Home	0.87	0.06	-0.01
Primary	0.79	-0.03	-0.17
Religious	0.76	0.1	-0.06
Siblings	0.66	0.17	-0.16
English listening frequency	-0.6	-0.18	0.03
Praying	0.58	0.25	-0.05
High-school	0.55	0.15	-0.1
English speaking frequency	-0.51	-0.22	0.25
Work	-0.14	1.05	0.29
School	-0.19	1.02	0.19
Health care	-0.2	0.93	0.03
Shopping	-0.08	0.9	0.09
Social activities	0.08	0.85	0.08
E-mail	-0.1	0.82	-0.14
Friends	0.13	0.82	0.12
Extracurricular	-0.02	0.8	-0.09
Room-mates	0.01	0.79	-0.13
Text	0.18	0.74	0
Social media	0.08	0.71	-0.15
Movies	0.1	0.69	0.09
Internet	-0.02	0.68	-0.23
Switching with social media	0.21	0.67	0.15
Neighbors	-0.09	0.63	-0.25
TV	0.2	0.61	-0.02
Lists	0.04	0.61	-0.33
Reading	0.14	0.59	-0.19
Partner	0.38	0.51	-0.13
Switching with friends	0.36	0.49	0.11
English understanding	0.08	0.15	1.03
English reading	0.14	-0.04	0.97
English writing	0.03	0.03	0.96
English speaking	0.1	-0.03	0.95
English writing frequency	-0.39	-0.09	0.5
Eigenvalues	25.41	4.15	2.26
% of variance	33	30	11
α from raw scores	0.68	0.94	0.87
α from polychoric R matrix	1	1	0.99
B			
Grandparents	0.961	0.636	-0.411
Infancy	0.955	0.654	-0.453
Switching with family	0.687	0.292	
Non-English understanding	0.96	0.676	-0.328
Non-English speaking	0.976	0.721	-0.344
Relatives	0.942	0.674	-0.404
Preschool	0.937	0.67	-0.505
Parents	0.954	0.708	-0.504
Non-English listening frequency	0.861	0.604	-0.332
Non-English speaking frequency	0.91	0.683	-0.39
Home	0.919	0.695	-0.459
Primary	0.85	0.643	-0.52

Table 2 (continued)

Item	Promax rotated factor loadings		
	Non-English Home use and proficiency	Non- English social use	English use
Religious	0.858	0.683	-0.476
Siblings	0.852	0.736	-0.567
English listening frequency	-0.747	-0.632	0.42
Praying	0.782	0.695	-0.471
High-school	0.704	0.607	-0.453
English speaking frequency	-0.785	-0.74	0.623
Work	0.484	0.776	-0.296
School	0.454	0.764	-0.348
Health care	0.459	0.772	-0.451
Shopping	0.522	0.783	-0.424
Social activities	0.656	0.861	-0.486
E-mail	0.563	0.842	-0.605
Friends	0.669	0.845	-0.451
Extracurricular	0.593	0.836	-0.569
Room-mates	0.634	0.872	-0.617
Text	0.713	0.87	-0.538
Social media	0.663	0.863	-0.625
Movies	0.555	0.71	-0.386
Internet	0.583	0.814	-0.644
Switching with social media	0.619	0.726	-0.36
Neighbors	0.474	0.714	-0.593
TV	0.648	0.764	-0.486
Lists	0.634	0.84	-0.721
Reading	0.651	0.804	-0.614
Partner	0.804	0.86	-0.621
Switching with friends	0.658	0.68	-0.358
English understanding	-0.296	-0.426	0.9
English reading	-0.343	-0.535	0.928
English writing	-0.395	-0.537	0.927
English speaking	-0.363	-0.539	0.92
English writing frequency	-0.687	-0.676	0.737

Note. Factor loadings over .40 appear in bold

ANCOVA. Both verbal, $F(1, 245) = 6.62, p = 0.011$, and non-verbal performance, $F(1, 86) = 10.2, p = 0.002$, varied significantly with increasing bilingualism. The difference in slope appears to be primarily driven by the non-verbal condition (see Fig. 3B for associated effect sizes).

Relation between factor scores and categorical classifications

In previous research with the LSBQ, researchers were required to evaluate the responses to determine a subjective categorical assignment of participants to a monolingual or bilingual group. Generally, individuals were classified as bilingual if they reported using two or more languages on a regular basis both in the home and in their social environment. Raters familiar with the instrument emphasized the importance of oral language usage rather than literacy. To assess these subjective judgments, the composite scores derived from the present analyses were grouped into five equal bins to examine the validity of the classifications used for participants in

the study in which they participated. The distribution of the composite scores across these bins, displayed in Fig. 2, was used to assess the fit between the categorical assignment to a binary language group and the continuous factor score obtained by the individual participant. The chi-square analysis was highly significant, $\chi^2(4) = 208.05, p < 0.001$, validating the categorical classifications. The results indicate that an overall assessment of a range of information about language experience, proficiency, and use leads to reasonable judgments about a categorical decision about bilingualism.

Central to these distributions is the determination of the cutoff scores, represented as dotted lines bounding the third quantile on Fig. 4. This figure also includes minimum and maximum composite scores for each quantile, making these values flexible to accommodate different research requirements. Our approach in this sample of young adults has been to indicate that individuals with a composite score of less than -3.13 are classified as monolinguals while those scoring above 1.23 are classified as bilinguals. Those falling between these cutoff scores can be classified as not strongly

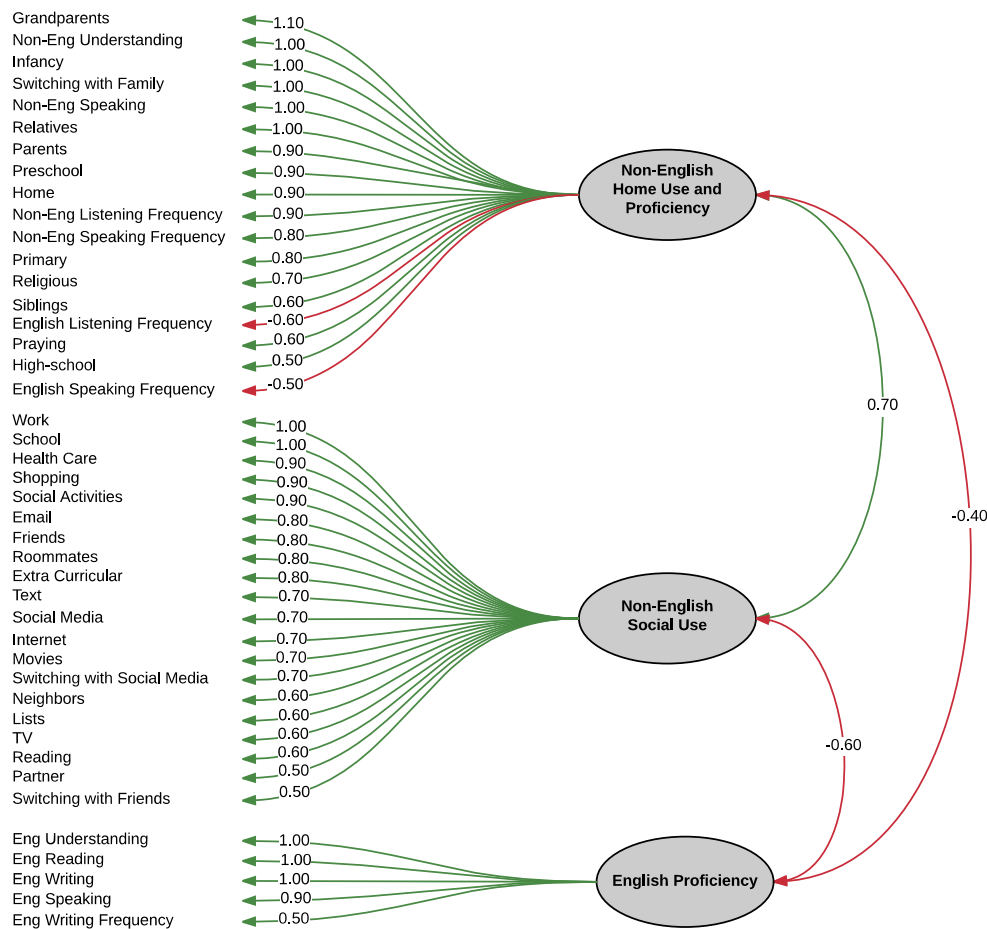


Fig. 2 Factor loadings and latent variables

differentiated or discarded, again depending on the research question. To facilitate ease of replication and use of the instrument, we include a LSBQ Administration and Scoring Manual (see <https://doi.org/10.6084/m9.figshare.3972486.v1>) and an Excel file as supplementary material which allows researchers to compute these scores independently. The cutoff values may be used to validate categorical assessments, identify outlying cases and facilitate the ease of the use of the LSBQ in labs where the instrument is less familiar, but should not substitute for good judgment and careful consideration.

Using the supplementary spreadsheet, we created two hypothetical cases. One, Monolingual Molly, was assigned the lowest scores for questions assessing knowledge of a second language and perfect knowledge of English. This case represented an individual who was extremely monolingual. A second hypothetical individual, Bilingual Betty, was assigned the highest scores for knowledge of a second language and English and represents a highly balanced and proficient bilingual. These are hypothetical cases in that an actual monolingual may have no knowledge of a second language and still not assign perfect scores to their English proficiency, resulting in a lower factor score than Monolingual Molly. Similarly, a

bilingual can be somewhat less fluent in one of the languages and still be a balanced bilingual even though the factor score is less than Bilingual Betty. The resulting factor scores and composite scores are available in the supplementary spreadsheet (Supplementary LSBQ Factor Score Calculator). Monolingual Molly had a composite score of -6.58 and Bilingual Betty had a composite score of 32.32 . These values indicate the general range of composite scores on the LSBQ.

Discussion

In most psychological research that investigates differences associated with categorical assignment to groups, the groups are easy to identify: developmental research might compare the performance of 6-year-olds to 8-year-olds, cognitive aging research might compare the performance of 20-year-olds to 60-year-olds, and educational research might compare outcomes of children in one kind of program to those of similar children in another program. Even in research based on differences in experience it can be relatively straightforward to classify participants: investigations on the effects of musical training have little difficulty in identifying individuals with

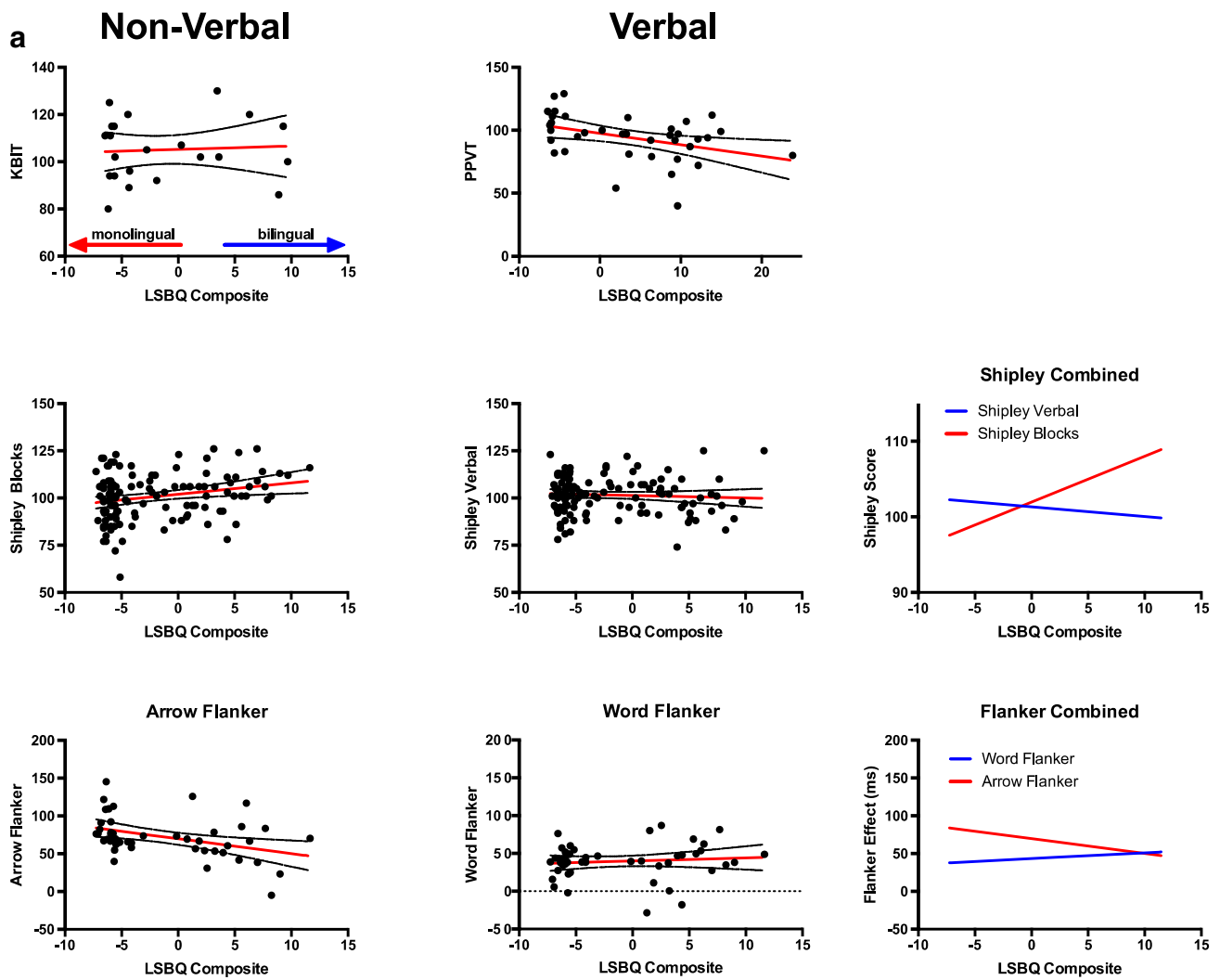


Fig. 3 Correlations of composite factor score with behavioral measures. (A) The relationship between the composite factor score and each of the behavioral measures. Dashed lines are upper and lower 95% confidence

Correlation values for each factor score with each of the behavioral measures

bounds. The sub-plots on the right side emphasize the interactions between the Shipley and Flanker word/non-word performance. (B) The correlation values and 95% confidence intervals for each factor

serious musical background and those without. Bilingualism is not like that. In the multiplicity of life experiences, we are confronted by any number of languages, even though most of those experiences leave little impression on our mental systems. The situation is even more complex for those who live in diverse, multicultural societies where multiple languages are widely represented. Add to that the exposure to languages through heritage language background, however distant or imperfect one’s proficiency, and the inevitable encounter with

foreign languages in education in all its forms. It becomes confusing at best to decide how to make a simple categorical designation to individuals in terms of this multidimensional and elusive construct.

It is important, however, that we do this. A growing body of research has shown the impact of bilingual experience for aspects of cognitive performance across the lifespan, including attention ability in infants (Bialystok, 2015; Kovács & Mehler, 2009; Pons, Bosch, & Lewkowicz, 2015), cognitive

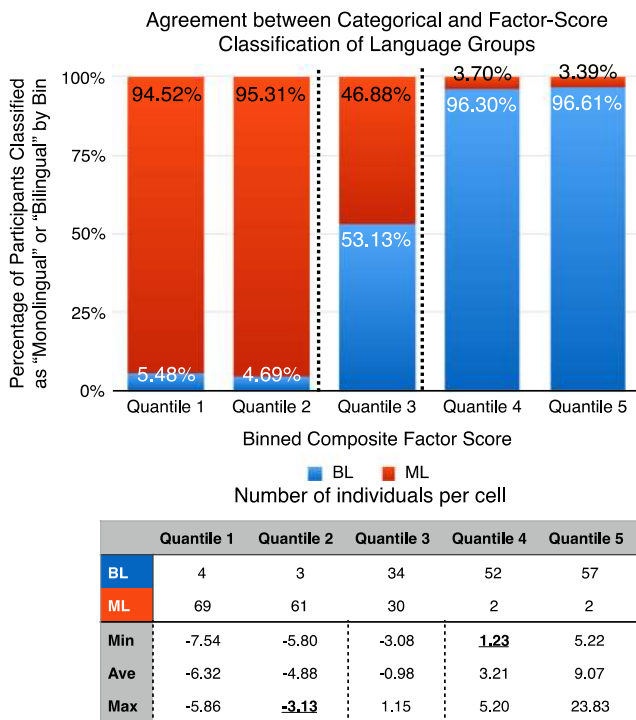


Fig. 4 Agreement between subjective and objective categorizations of bilingualism. Quantile scores are derived from the composite factor score (see text). The table below shows the count values. *BL* bilingual, *ML* monolingual, *Min and max* the minimum and maximum observed values of the composite score for each quantile. Values not contained in the dotted lines are cut-off scores

performance in childhood (Barac, Bialystok, Castro, & Sanchez, 2014), and the ability to cope with dementia in older age (Bak & Alladi, 2014). These are important findings with enormous consequences for education, public health, and social policy (for discussion, see Bialystok, Abutalebi, Bak, Burke, & Kroll, 2016). Increasingly, however, the demonstration of bilingual consequences on cognitive ability has become difficult to find in young adults. There are many reasons for this (for discussion, see Bak, 2016; Bialystok, 2016), but one strong possibility is that the definition of bilingualism used in the various studies is imprecise and incommensurate across studies.

The challenge addressed in the present paper was to present an objective and quantifiable means of establishing group membership along the monolingual-bilingual continuum, especially for individuals living in diverse societies. Previous speculation on this issue had assumed that the relevant factors would capture the extent of second language use and the degree of second language proficiency, with one of them revealing a stronger association with the cognitive outcomes associated with bilingualism (e.g., Luk & Bialystok, 2013). The results of the present study do not support that assumed dichotomy: the two factors that emerged from the questionnaire and were then associated with cognitive outcomes were proficiency and use of the non-English language at home and use

of the non-English language in social settings; in other words, second-language use is relevant for both primary factors. Moreover, the factor structure reveals an important role for the contexts in which languages are used. Different contexts of use place different demands on language control processes. For example dual-language contexts (e.g., using a non-English language for health-care services and switching from English with the receptionist to a second language with the doctor) involve more control processes than single-language contexts where one does not have to attend to cues to determine which language to use.

The emergence of context of use as an important dimension fits well with recent theorizing about the role of context in determining the nature of the bilingual effect. In the Adaptive Control Hypothesis, Green and Abutalebi (2013) identify three types of interactional contexts for bilinguals that are based on different requirements to manage, select, and switch between languages. The contexts are determined by the extent to which both languages are represented and the presumed proficiency of the interlocutor that would license a more liberal degree of language switching. These interactional contexts make different demands on attention and monitoring, and Green and Abutalebi (2013) described the cognitive and brain consequences of each. The present results support that view by demonstrating that the context in which languages are used defines the degree of bilingualism the individual possesses and that the degree of bilingualism is associated with the extent to which cognitive consequences are found. Such context effects were found in a study by Wu and Thierry (2013) in which Welsh-English bilinguals performed *better* on a non-verbal conflict resolution task when irrelevant distracters were presented in both languages than either language alone.

To date, there has been no consensus or standardized method for determining bilingualism. An important contribution to this problem was the LEAP-Q proposed by Marian et al. (2007), but this study had a small sample size and used an orthogonal model. Our study has addressed both of these limitations. Similarly, the LHQ-2.0 developed by Li et al. (2014) is notable for its ease of use, accessibility, and large sample. However, the instrument is agnostic on the essential questions concerning demarcations between monolingual and bilingual participants. The LSBQ addresses this issue by identifying suggested criteria for participant classification. For example, the LSBQ could answer whether on average the language one uses to speak to one's parents is more important than the language one uses to speak to one's neighbors.

The exploratory analysis yielded a three-factor solution (in order of variance explained): home language use and non-English language proficiency, social language use, and English proficiency. The solution in part reflects the context in which the research was conducted: a deeply diverse population in an English-speaking community in which most

bilinguals speak English and a heritage language. A different order of factor strength might emerge in different contexts; for example, in locations where two languages are equally prevalent (e.g., English and French in Montreal or Spanish and Catalan in Barcelona), proficiency in the non-dominant language could load more highly on factor 2, social language use. That is, the more one has opportunities to use a second language outside the home, the more non-majority language proficiency may co-vary with the social language use factor. This interpretation fits with the Adaptive Control Hypothesis which proposes that language control flexibly adapts to environmental demands (Green & Abutalebi, 2013). Individuals who obtain the highest bilingualism factor score on our instrument are those who have likely mastered adaptive control across each of the three interactional contexts (single-language context, dual language context, and dense code-switching contexts). These are individuals who might speak a heritage language at home, switch between languages while interacting with individuals with single-language proficiency, and interleave language use when speaking with other individuals who are highly proficient in both languages. Language control, therefore, is *dynamic* and fluctuates according to contexts and resource costs. The identification of factors specifically focused on home or social use makes the LSBQ an ideal instrument for studies investigating the Adaptive Control Hypothesis. To address these questions more broadly, the instrument can be translated into another language and the presumption of English as being the community language adjusted accordingly.

As part of assessing the ecological validity of the LSBQ, we correlated the factor scores and the composite factor score with previously collected behavioral data. In the bilingualism literature, one reliable finding is a reduction in receptive vocabulary with increasing second language proficiency, a finding we replicated using the PPVT. Notably, we demonstrated significant interactions between verbal and non-verbal homologs for two tasks: the Shipley and the Flanker. In each case, the interaction was driven by stable or decreasing performance with increasing bilingualism on the verbal component and increasing performance on the non-verbal component; Scores on the Shipley Blocks increased from ~97 for monolinguals to ~108 for bilinguals. This difference in Gf is larger than many effects found for cognitive training programs (for a recent meta-analysis, see Au et al., 2015).

The chi-square analysis confirmed that our previous categorical characterizations using the LSBQ agreed with the factor-score classifications, confirming our interpretation of our previous research results. Using five quantiles, agreement between the two methods was robust for the first and last two quantiles but split evenly at the third. Another option, however, is to use the *continuous* composite-factor score instead of a categorical assignment to two groups. Continuous measurement increases power and allows the researcher to retain the

full range of participants. A continuous measure may also be sensitive to nonlinearities in the data. For example, consider an outcome where monolinguals and highly proficient bilinguals perform similarly but intermediate bilinguals perform more poorly (perhaps struggling to manage interference). A median split or extreme groups comparison would be insensitive to this quadratic effect. The approaches of quantifying bilingualism on a scale and classifying participants into groups are not mutually exclusive but rather alternative methodologies that are best suited to different questions or different populations. This information can also be used in terms of the factor scores that focus on a particular aspect of bilingual experience. A novel strength of the LSBQ is that that same instrument can be used with all these approaches, bringing a level of coherence to the description of bilingualism that is not currently available.

Like the other available instruments, the LSBQ depends on self-report and self-assessment, but it addresses the deficiencies of self-report through multiple questions that are demonstrated through the factor analysis to be reliably related. Single-question self-assessments are notoriously unreliable (e.g., Davis et al., 2006), particularly regarding knowledge of a non-native language. For example, Delgado, Guerrero, Goggin, and Ellis (1999) reported that a group of bilingual Hispanics was able to accurately report their knowledge of Spanish but not their knowledge of English, their second language. The authors reported that despite receiving objective feedback, self-assessments improved only for the native language. This finding underscores the need for a more comprehensive and sensitive measure of bilingualism. Such a lack of precision is evident in a study by Paap and Greenberg (2013) in which they asked participants to rate their speaking and listening skills each on a single seven-point scale from "beginner" to "super-fluency." These two self-assessments were then used to classify participants as bilingual or monolingual. The use of such an approach is discouraging in the face of overwhelming evidence that coarse self-assessments of this nature are notoriously unstable. Experts are more adept at self-assessing performance than laypersons in their domain of expertise (e.g., Falchikov & Boud, 1989). More sensitive instruments capturing the same construct in multiple ways are also more effective than a single question (e.g., Dunning, Heath, & Suls, 2004). For this reason personality questionnaires, such as the Big Five Inventory, include ~50 questions to assess five characteristics (John, Donahue & Kentle, 1991). There is redundancy in well-planned questionnaires, but this redundancy helps to uncover the internal consistency and stability in a complex multidimensional construct.

One limitation of our instrument is that it is not designed to accurately assess language use beyond two languages – that is, multilingualism. The questions on the LSBQ are scaled from all English use to all non-English use, and adding a third or fourth language is not accommodated. Anecdotally,

multilingual participants appeared to be confused by those questions and responded unsystematically. Not surprisingly, therefore, the application of the current factor solution to the multilinguals failed to converge. For cases of exposure to two languages, however, the LSBQ is a reliable and valid instrument for describing bilingual experience and classifying participants.

Author note This research was supported by grant A2559 from National Sciences and Engineering Research Council of Canada (NSERC) and grants R01HD052523 and R21AG048431 from the National Institutes of Health (NIH) to EB.

References

- Au, J., Sheehan, E., Tsai, N., Duncan, G. J., Buschkuhl, M., & Jaeggi, S. M. (2015). Improving fluid intelligence with training on working memory: A meta-analysis. *Psychonomic Bulletin & Review*, 22(2), 366–377. doi:10.3758/s13423-014-0699
- Bak, T. H. (2016). The impact of bilingualism on cognitive aging and dementia. *Linguistic Approaches to Bilingualism*, 6(1), 205–226. doi: 10.1075/lab.15002.bak
- Bak, T. H., & Alladi, S. (2014). Can being bilingual affect the onset of dementia? *Future Neurology*, 9, 101–103. doi:10.2217/fnl.14.8
- Barac, R., Bialystok, E., Castro, D. C., & Sanchez, M. (2014). The cognitive development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, 29, 699–714. doi:10.1016/j.ecresq.2014.02.003
- Bialystok, E. (2015). Bilingualism and the development of executive function: The role of attention. *Child Development Perspectives*, 9(2), 117–121. doi:10.1111/cdep.12116
- Bialystok, E. (2016). Aging and bilingualism. *Linguistic Approaches to Bilingualism*, 6(1), 1–8. doi:10.1075/lab.16004.bia
- Bialystok, E., & Barac, R. (2012). Emerging bilingualism: Dissociating advantages for metalinguistic awareness and executive control. *Cognition*, 122, 67–73. doi:10.1016/j.cognition.2011.08.003
- Bialystok, E., Craik, F. I. M., Green, D. W., & Gollan, T. H. (2009). Bilingual minds. *Psychological Science in the Public Interest*, 10, 89–129. doi:10.1177/1529100610387084
- Bialystok, E., Abutalebi, J., Bak, T. H., Burke, D. M., & Kroll, J. F. (2016). Aging in two languages: Implications for public health. *Ageing Research Reviews*, 27, 56–60. doi:10.1016/j.arr.2016.03.003
- Calvo, N., García, A. M., Manóloff, L., & Ibáñez, A. (2016). Bilingualism and Cognitive Reserve: A Critical Overview and a Plea for Methodological Innovations. *Frontiers in Aging Neuroscience*, 7(12), 3–17. doi:10.3389/fnagi.2015.00249
- Davis, D. A., Mazmanian, P. E., Fordis, M., Van Harrison, R. T. K. E., Thorpe, K. E., & Perrier, L. (2006). Accuracy of physician self-assessment compared with observed measures of competence: A systematic review. *Journal of American Medical Association*, 296(9), 1094–1102. doi:10.1001/jama.296.9.1094
- Delgado, P., Guerrero, G., Goggin, J. P., & Ellis, B. B. (1999). Self-assessment of linguistic skills by bilingual Hispanics. *Hispanic Journal of Behavioral Sciences*, 21(1), 31–46. doi:10.1177/0739986399211003
- Dong, Y., & Li, P. (2015). The cognitive science of bilingualism. *Language and Linguistics Compass*, 9(1), 1–13. doi:10.1111/lnc3.12099
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5(3), 69–106. doi:10.1111/j.1529-1006.2004.00018.x
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, 16(1), 143–149.
- Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research*, 59(4), 395–430. doi:10.3102/00346543059004395
- Flora, D. B., Finkel, E., & Foshee, V. (2003). Higher order factor structure of a self-control test: Evidence from confirmatory factor analysis with polychoric correlations. *Educational and Psychological Measurement*, 63, 112–127. doi:10.1177/0013164402239320
- Gathercole, V. C., Thomas, E. M., Kennedy, I., Prys, C., Young, N., Vinas Guasch, N., ... Jones, L. (2014). Does language dominance affect cognitive performance in bilinguals? Lifespan evidence from preschoolers through older adults on card sorting, Simon, and metalinguistic tasks. *Frontiers in Psychology*, 5(11). doi: 10.3389/fpsyg.2014.00011
- Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A multilingual naming test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition*, 15(3), 594–615. doi:10.1017/S1366728911000332
- Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, 25, 515–530. doi:10.1080/20445911.2013.796377
- Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, 1(2), 131–149. doi: 10.1017/S136672899800025X
- Grosjean, F. (2013). Bilingualism: A short introduction. In F. Grosjean & P. Li (Eds.), *The psycholinguistics of bilingualism* (pp. 5–25). Malden, MA: Wiley-Blackwell.
- Hair, J. F. J., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis* (4th ed.). Saddle River, NJ: Prentice Hall.
- Hakuta, K. (1986). *Mirror of language: The debate on bilingualism*. New York: Basic Books.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The big five inventory—versions 4a and 5a*. Berkeley, CA: University of California, Berkeley, Institute of Personality and Social Research.
- Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika*, 39, 31–36.
- Kaufman, A. S., & Kaufman, N. L. (2014). *Kaufman brief intelligence test, Second Edition*. Encyclopedia of Special Education. doi:10.1002/9781118660584.ese1325
- Kave, G., Eyal, N., Shorek, A., & Cohen-Mansfield, J. (2008). Multilingualism and cognitive state in the oldest old. *Psychology and Aging*, 23, 70–78. doi:10.1037/0882-7974.23.1.70
- Kline, P. (1979). *Psychometrics and psychology*. London: Academic Press.
- Kolb, B., Mychasiuk, R., Muhammad, A., Li, Y., Frost, D. O., & Gibb, R. (2012). Experience and the developing prefrontal cortex. *Proceedings of the National Academy of Science*, 109(Suppl 2), 17186–17193. doi:10.1073/pnas.1121251109
- Kovács, Á. M., & Mehler, J. (2009). Cognitive gains in 7-month-old bilingual infants. *Proceedings of the National Academy of Sciences*, 106, 6556–6560. doi:10.1073/pnas.0811323106
- Li, P., Sepanski, S., & Zhao, X. (2006). Language history questionnaire: A web-based interface for bilingual research. *Behavior Research Methods*, 38(2), 202–210. doi:10.3758/BF03192770
- Li, P., Zhang, F., Tsai, E., & Puls, B. (2014). Language history questionnaire (LHQ 2.0): A new dynamic web-based research tool. *Bilingualism: Language and Cognition*, 17(3), 673–680. doi:10.1017/S1366728913000606
- Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal*

- of *Cognitive Psychology*, 25(5), 605–621. doi:10.1080/20445911.2013.795574
- Marian, V., Blumenfeld, H. K., & Kaushanskaya, M. (2007). The language experience and proficiency questionnaire (LEAP-Q): Assessing language profiles in bilinguals and multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4), 940–967. doi:10.1044/1092-4388(2007/067)
- Paap, K. R., & Greenberg, Z. I. (2013). There is no coherent evidence for a bilingual advantage in executive processing. *Cognitive Psychology*, 66(2), 232–258. doi:10.1016/j.cogpsych.2012.12.002
- Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, 69, 265–278. doi:10.1016/j.cortex.2015.04.014
- Pascual-Leone, A., Amedi, A., Fregni, F., & Merabet, L. B. (2005). The plastic human brain cortex. *Annual Review of Neuroscience*, 28, 377–401. doi:10.1146/annurev.neuro.27.070203.144216
- Pons, F., Bosch, L., & Lewkowicz, D. J. (2015). Bilingualism modulates infants' selective attention to the mouth of a talking face. *Psychological Science*, 26, 490–498. doi:10.1177/0956797614568320
- Sheng, L., Lu, Y., & Gollan, T. H. (2014). Assessing language dominance in Mandarin–English bilinguals: Convergence and divergence between subjective and objective measures. *Bilingualism: Language and Cognition*, 17(2), 364–383. doi:10.1017/S1366728913000424
- Wu, Y. J., & Thierry, G. (2013). Fast modulation of executive function by language context in bilinguals. *The Journal of Neuroscience*, 33(33), 13533–13537. doi:10.1523/JNEUROSCI.4760-12.2013
- Zachary, R. A., & Shipley, W. C. (1986). *Shipley institute of living scale: Revised manual*. WPS, Western Psychological Services