Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005–2015

Describing bilinguals: A systematic review of labels and descriptions used in the literature between 2005-2015

SARAH SURRAIN
Harvard Graduate School of Education

GIGI LUK
Harvard Graduate School of Education

(Received: August 10, 2017; final revision received: November 11, 2017; accepted: November 20, 2017)

ACKNOWLEDGEMENTS
We acknowledge Sarah Howard, Carolynn Ianello, Analiese Reigstad and Valerie Woxholdt for their assistance in coding the database for this project. We also thank two anonymous reviewers for their constructive suggestions to an earlier version of the manuscript.

Address for correspondence:
Gigi Luk, Ph. D.
Harvard Graduate School of Education
14 Appian Way
Cambridge, MA 02138
U.S.A.
Gigi_luk@gse.harvard.edu

Abstract

Recent years have seen a surge in research comparing bilinguals to monolinguals, yet synthesizing this literature is complicated by the diversity of language and social backgrounds behind these dichotomous labels. The current study examines the labels and descriptions reported in 186 studies comparing bilinguals and monolinguals published between 2005-2015 in order to understand how bilingualism has been operationalized and describe the degree to which different facets of bilingual experience are reported. Proficiency and usage were the most frequently reported features (77% and 79%), followed by language history (67%) and the language of schooling (60%). However, less than half of the studies measured proficiency objectively or reported proportional usage, and even less – 30% – described the sociolinguistic context from which the sample was drawn. Given the increase in language contact due to globalization, more transparent and comprehensive reporting of participant characteristics is critical to building our understanding of how bilingualism affects experience.

Keywords: bilingualism, multilingualism, systematic review, heterogeneity, contextual factors

Introduction

In a world where language diversity within nations and communities is rapidly increasing due to migration and globalization, inquiry into bilingualism as a life experience is more important than ever. The recent surge in research output and annual citations reflect this need (Kroll & Bialystok, 2013). Most of these studies have compared binary groups, namely, bilinguals and monolinguals, to test whether bilingualism affects various behavioral outcomes. However, despite the volume of output generated by this interdisciplinary field, generalizable findings remain elusive and syntheses of the literature are complicated by a lack of clarity and consistency across studies in determining whether participants are bilingual or monolingual. We know that bilingualism is a dynamic and interactive experience characterized by both individual and contextual factors (Grosjean, 2013). Yet, studies differ in how individual attributes such as first (L1) or second language (L2) proficiency (Hulstijn, 2012), language acquisition history or language dominance (Birdsong, 2016; Birdsong & Vanhove, 2016), and contextual factors such as opportunities to use one's languages and the local status of one's languages (for adults, see Green & Abutalebi, 2013; for children, see Paradis & Nicoladis, 2007) are reported. If studies utilize different criteria for determining who is bilingual, comparison or synthesis across studies becomes challenging. To address these fundamental methodological issues, the present study documents labels and descriptions for bilinguals and monolinguals in studies published between 2005 and 2015 that directly compared these two groups. We aim to demonstrate the variability in reporting comprehensive participant characteristics in research involving bilinguals and monolinguals. The study is meant to describe the landscape of the research rather than to define bilingualism. Given the societal implications of research on bilingualism, it is essential to explicitly report the characteristics that qualify participants to be bilingual (and monolingual) in

order to facilitate knowledge translation.

The first methodological difficulty in comparing bilinguals and monolinguals is that bilingualism is not a categorical variable; there is no clear line between what constitutes bilingual versus monolingual experience (Luk & Bialystok, 2013). As Mackey (1962) noted more than 50 years ago, "the point at which a speaker of a second language becomes bilingual is either arbitrary or impossible to determine. … We are forced to consider [bilingualism] as something entirely relative" (p. 52). Not only is bilingual experience a matter of degree, but it also comprises multiple features, such as history and context of language acquisition as well as proficiency and current usage of each language (Fishman & Cooper, 1969; Grosjean, 2013; Kaushanskaya & Prior, 2015; Luk, 2015). Although these features collectively constitute bilingualism as a life experience, they are not always reported comprehensively in research studies.

The second methodological challenge in comparing monolinguals to bilinguals is that some studies recruit individuals whose experiences fall closer to the extremes of the bilingual-monolingual continua of age of acquisition, proficiency, or current usage, while others recruit those that inhabit the grey areas in between. For example, Peal and Lambert compared balanced, highly proficient bilinguals to "unambiguous" monolinguals in their landmark 1962 study, while more recent studies often take a broader view of bilingualism and determine the line dividing bilinguals from monolinguals based on the variation within a particular sample. This strategy is sound within an individual study, but not when studies are collectively considered. Generalizing findings from studies with different qualifying criteria to group differences between monolinguals and bilinguals may constrain our understanding of how different features of bilingualism are associated with different behavioral outcomes. To date, there is no empirical

documentation of how the broad label of "bilingual" is conferred. The current study aims to examine this issue in the recent literature. An additional goal is to document how monolinguals are described in research. Considered as "controls," monolingual participants are often assumed to be a homogeneous group. This assumption may not be accurate in studies where "monolingual" refers to participants with inconsistent second language usage or relatively lower self-reported second language proficiency (Alvarado & Jameson, 2011; Paap & Greenberg, 2013; Park & Ziegler, 2014). By documenting the labels and their descriptions in recent research, we emphasize the importance of understanding the variations within bilingual features and the pitfalls of assuming that bilingual-monolingual comparisons are constant and consistent across studies.

We expect that the way bilingualism is defined and described may differ based on the age of the participants. Clearly, features such as proficiency, current usage, history of acquisition and social context may shift over the lifespan and be more or less relevant to report for child and adult participants. For example, while sufficient proficiency in multiple languages is often a primary criterion for bilingualism, determining which skills to measure or where to set the bar is a complex question (Carroll, 2017). In fact, it may not be realistic to assume bilinguals can attain proficiency comparable to that of monolinguals (Grosjean, 1989). Measuring and reporting proficiency may be even more difficult in research involving young children, given that they are often still acquiring all of their languages (Unsworth, 2008). Partly due to the complications associated with proficiency criteria, several scholars have suggested usage as a more reliable indicator of the degree of bilingual experience (Grosjean, 2013). The amount that individuals use their languages on a daily basis is likely an important attribute at all ages, but may be particularly salient in studies of young children. This idea has been borne out in recent work by Bedore and

colleagues (2016), who found that current daily input and output in both languages was particularly important for understanding the development of language minority children in the United States (Bedore, Peña, Griffin, & Hixon, 2016).

In addition to contributing to our understanding of current language usage, the language of schooling provides pertinent information about bilingual experience, particularly for children. Students may receive instruction in their native language(s), other language(s) or both. For example, children of Portuguese immigrants in Luxembourg are immersed in Luxembourgish, German and French at school without access to school-based supports in their native language (Engel de Abreu, Cruz-Santos, Tourinho, Martin, & Bialystok, 2012). This configuration marks a key part of their foundational bilingual experience. In contrast, language minority children with access to schooling in their native language would likely achieve greater proficiency in that language and experience less native language attrition over time (Rolstad, Mahoney, & Glass, 2005). We expect that the language of schooling will be reported more frequently in studies of children, though it offers important clues to adults' language history as well.

Language history – more specifically, the age at which an individual began learning (Unsworth, 2016) or actively using a second language (Luk, de Sa, & Bialystok, 2011) – is an important feature that we expect to see reported in the literature. Given the relationship between age of acquisition and ultimate attainment (Birdsong & Vanhove, 2016), we anticipate that studies will be more likely to report the language history of adult participants. It is also acknowledged that bilingual language learning or usage changes across time. Therefore, studies might also record changes in language usage patterns across the lifespan. Finally, given that bilingual experience reflects how individuals interact with others in the community via language use, we also document the degree to which the sociolinguistic context is reported in the

descriptions of both bilinguals and monolinguals. Here, sociolinguistic context refers to characteristics of the community such as official languages and general language use, beyond the family unit, in the country or region where the study took place. Information about national and community language use is helpful for comparing studies conducted internationally. We acknowledge that the attributes reviewed here, namely, L1/L2 proficiency, functional language usage, language acquisition history, and sociocultural context, are interrelated and, in combination, they provide a comprehensive description of monolingual and bilingual samples in individual studies.

To describe the variation in the labels and descriptions used in recent research on bilingualism we compiled a database of 186 studies from 165 empirical articles comparing bilingual and monolingual groups, and asked the following two questions: 1) What labels and descriptions are used to report on participants in group comparison studies? 2) Do these labels and descriptions vary by age of the participants or the region where the study was conducted? By documenting the variability in describing participants' language characteristics between groups, we hope to emphasize the nuances and the complexity embedded in group comparisons involving monolinguals and bilinguals. Furthermore, we aim to encourage more transparent reporting of participants' language experience that would enrich our understanding of bilingualism as a life experience.

Method

Search strategy and inclusion criteria

The purpose of this review was to identify a representative subset of the literature in a replicable manner in PsycINFO. We searched the database using the keywords *biling** or *multiling** and

*monoling** and limited our search to English language, peer-reviewed, empirical studies from 2005 to 2015 (see Figure 1 for a decision flow chart for the search, inclusionary, and exclusionary actions). This search yielded 277 articles which were then screened based on the full text, with 141 of these meeting the inclusion criteria described below. An additional 25 articles from a previous, pilot search of PsycINFO using only the keywords *biling** or *multiling** were added to this search, yielding a total of 165 articles after removing duplicates. Some of these articles contained multiple studies with unique samples that met our criteria, and so the total number of studies reviewed was 186, reported in 165 articles.

<Insert Figure 1 about here>

In addition to the search limits described above, studies had to meet four criteria for inclusion in our review. First, we included only studies that compared bilingual and monolingual groups on one or more literacy, language or cognitive outcomes. To meet this inclusion criterion, we included studies that used a group approach in which comparisons were based on participants' language backgrounds, with at least one group broadly defined as "bilingual" and another as "monolingual." Second, we required that each language group have a sample size of 15 or larger to limit our review to studies with reasonable power for statistical inferences. Third, we required that study data be obtained from unique samples rather than using aggregate data from previous studies to avoid overlap among the participants being described. Finally, we included only studies conducted with healthy, typically-developing populations in order to focus on distinctions based on language experience rather than clinical conditions.

Coding the studies

Each study was coded for the age of the participants (adults, children or both), the country or countries in which the research was conducted, the main outcome of interest (literacy, language

or cognitive), and the labels and descriptions used to describe the language groups. The main outcome of interest was documented because these studies span different disciplines in the field of bilingual research. We wanted to capture this variance to understand the nuances in how bilingualism is described across an interdisciplinary field. Studies that focused on reading or writing skills such as spelling or reading comprehension were coded as "literacy". Studies focusing on outcomes such as lexical retrieval or morphological awareness were coded as "language". Finally, studies that compared groups on outcomes such as inhibitory control or working memory were coded as "cognitive". The procedure for coding the labels used for the bilingual and monolingual groups was as follows: First, the most consistent label applied to each language group across the title, abstract, tables, and figures was recorded (e.g. "successive bilingual Turkish-speaking children" as reported in Marinis & Chondrogianni, 2010). Second, the bilingual group labels were dichotomously coded as to whether they *only* used the words "bilingual" or "trilingual", or whether they used additional terms related to the bilingual experience of the participants. If so, these terms were then coded as to whether they named the specific languages spoken (Language Pair Label), whether they alluded to the acquisition history (History Label), whether they referenced ongoing acquisition of one of the languages (Learning Label), and whether they suggested the degree of dominance or proficiency (Dominance Label) (see Table 1 for examples of each).

<Insert Table 1 about here>

The qualitative descriptions of the participants in each language group were of particular interest, as these provided insight into the diverse ways that the terms bilingualism and monolingualism are operationalized in the literature. In order to capture the features of bilingualism that are reflected in each description, the text describing the participants (taken from

the methods section of each study) was qualitatively coded according to the rubric in Table 2. In line with our knowledge of the literature and the features that we expected to be of theoretical importance, we recorded whether each study reported on the proficiency, functional usage, school language, language history, and sociolinguistic context of the participants. Proficiency was coded as either subjective or objective. Subjective judgement refers to self-reports or other-reports whereas objective measurement refers to administering normed-referenced assessments and reporting standard scores against a norming table. Functional usage was coded as either categorical or gradient. Categorical refers to usage reported as a dichotomous indicator of whether both languages were used in daily life, while gradient refers to a relative quantitative report of language usage or exposure. This allowed us to look for relationships between how the language groups were described and other aspects of the studies such as age of participants and region. Each feature is reported independently to assess the consistency of reporting across studies. It is acknowledged that these features overlap and that some studies reported more than one feature.

<Insert Table 2 about here>

The studies were coded by a team of three research assistants along with the first author. Twenty-four percent of the studies were coded by two to four coders. Interrater reliability was initially lower than desired in the first round of double-coding (agreement=78%, Kappa=0.66), but after resolving all discrepancies and clarifying the coding scheme, agreement between coders in subsequent rounds ranged from 83% to 87% ($.7 < \text{Kappa} < .76$). All disagreements were discussed and resolved by consensus. The first author personally coded or audited 89 (48%) of the 186 studies. Among the 186 studies, the largest number of studies with the same first author

is 11 (6% of the total). The next two most frequent first authors appeared five times each in the database. No other first author appeared more than three times.

Results

Our goal was to describe how bilingualism has been operationalized in the literature from 2005-2015 and explore how this may differ by the age of the participants and the geographic region from which the participants were drawn. We first provide descriptive statistics on the distribution of these characteristics for the full set of studies in our database. We then examine the labels used to describe the bilingual groups and determine whether these differed by age or region. Next, we focus on the descriptions of the bilingual and monolingual groups and document the percentage of studies that reported on participants' proficiency, home usage, school language, language history, and sociolinguistic context, noting differences by age and region.

Studies in the database

The 186 studies in our final database included 9,143 bilingual participants and 7,836 monolingual participants (see the supplementary materials online for a complete list of the studies). The number of participants per language group ranged from a minimum of 15 (following our inclusion criteria) to a maximum of 1000[1], with a median bilingual group sample size of 33. A total of 111 studies (59%) included only child participants, while 76 studies included either only adults (n=72) or adults and children (n=4). Across all studies, 32 countries were represented. The majority of the study participants came from North America, with 56 studies (30%) conducted in the United States and 48 studies (25%) conducted in Canada. Forty-six studies (25%) described participants from across 14 countries in the European Union and the

United Kingdom, and 15 studies (8%) drew participants from countries in Asia, Australia, and Africa. Finally, 22 studies (12%) included participants from multiple regions, such as Canada and China or the United States and Mexico.

The studies compared bilingual and monolingual groups on a wide range of outcomes. Eighty-five studies (46%) focused on general language outcomes. A similar proportion of studies (45%, n=84) compared participants on cognitive outcomes. The smallest number of studies were concerned with literacy outcomes (9%, n=17). The underrepresentation of studies examining literacy as an outcome is due to the terms we adopted for our literature search. In literacy research with students from diverse language backgrounds, labels such as "English language learners" (e.g., Solari, Petscher, & Folsom, 2014) or "language minority learners" (e.g., McNeill & Everatt, 2013) are often used instead of "bilingual." These studies were not included in our search as we focused on studies with a monolingual-bilingual group comparison. Therefore, this search yielded a small number of studies with literacy as an outcome. While this breakdown provides a general sense of the domains in which bilinguals and monolinguals have been compared in the literature, it is important to note that the boundaries between these categories are not definitive or mutually exclusive. For example, most general language outcomes are relevant to literacy outcomes, and most studies investigating cognitive outcomes involve both verbal and nonverbal processing.

Labels Applied to the Bilingual and Monolingual Groups

The labels used to describe the bilingual groups ranged from more general to more specific (see Table 1). Fifty-eight studies (31%) referred to the bilingual group as simply “bilingual” without additional qualifiers. The largest group (n=85, 46%) used a label that indicated the languages spoken, but did not indicate the dominance, learning status or acquisition

history of the participants. A small number of studies used labels that indicated either the dominance (2%), learning status (5%), or history of acquisition only (6% each), while 10% applied more specific labels that provided some combination of the features described above. Disaggregated by age, we find that studies with child participants were more likely to use the Learning Only label (8% compared to 1% in adults) and the History Only label (7% compared to 4% in adults), while adult studies were slightly more likely to use the Language Pair only label (51% compared to 42% in children), though these differences did not reach statistical significance, X^2 (5, N = 186) = 6.63, p = .25.

While the distribution of labels did not differ significantly by region, X^2 (20, N = 186) = 27.09, p = .133, some diverging patterns were observed. For example, in Canada the most used label type was Bilingual Only (47% of studies), while the most used in the United States, the EU & UK, and Asia/Australia/Africa was Language Pair Only (54%, 57%, and 60%, respectively). Studies from multiple regions showed an even split in the frequency of using theBilingual Only label andthe Language Pair Only label (32% for each). These studies were also the most likely to use combination labels (23%). One interpretation of these regional differences could be that studies conducted in Canada often draw from a diverse population speaking multiple languages and so require more general labels to describe heterogeneous samples, while studies that recruit samples from multiple regions are likely targeting specific speech communities and thus can apply more specific labels.

The labels used to describe the monolingual groups follow a pattern similar to the bilingual group labels. A little over one third used the term Monolingual Only (n=66, 35%), more than half referred to the monolingual group by the language spoken (n=102, 55%), and 10% of the studies (n=18) included additional qualifiers such as “fully monolingual” (O’Hora,

Pelaez, & Barnes-Holmes, 2005), "functionally monolingual in English" (Bedore, Fiestas, Peña, & Nagy, 2006), or "Monolinguals in English Language Instruction" (Rolla San Francisco, Mo, Carlo, August, & Snow, 2006). Language dominance and learning is often assumed to be homogeneous in monolinguals, but these labels suggest that there may be important differences between the monolingual participants in different studies.

Descriptions of Bilingual and Monolingual Groups

Next, the descriptions of each language group provided in the methods section were examined to understand the extent to which features such as proficiency, home and school usage, history, and sociolinguistic context were reported. For each of these features, the pattern across all studies as well as differences by age and geographical region are reported.

Proficiency

Definitions of bilingualism and theories of how bilingual experience impacts development and learning often hinge on an individual's proficiency in each language. Across all studies, 141 (77%) reported the proficiency of the participants based on objective or subjective measures in at least one language.[2] Fifty-five studies (30% of the total) reported objective proficiency only. In these cases, proficiency was based on a formal assessment administered to both bilingual and monolingual groups, with more than half using the Peabody Picture Vocabulary Test (PPVT) and its adaptations in other languages. Other studies reported using measures such as the Expressive One-Word Picture Vocabulary Test (EOWPVT), the Shipley Vocabulary Test, or the MacArthur-Bates Communicative Development Inventory (MCDI).[3] Of the 55 studies that relied on objective proficiency measures, 21 studies (38%) tested the bilinguals in only one of their languages. Across these 21 studies, the language tested was always the dominant majority

language in the community – English (n=17), Dutch (n=2), French (n=1) or Korean (n=1), which was the second language of the bilinguals in at least 11 of the studies and the first language of the bilinguals in two studies.[4] The studies that tested participants in only one language were likely constrained by the variety of different first languages spoken by bilingual participants (e.g. Bialystok, Barac, Blaye, & Poulin-Dubois, 2010; Byers-Heinlein & Werker, 2009; Portocarrero, Burright, & Donovick, 2007). Fifty-four studies (29%) reported subjective proficiency only. That is, proficiency was based solely on self- or parent-reports. An additional 32 studies (17%) included both objective and subjective proficiency. Finally, 43 studies (23% of the total) did not meet our criteria for reporting either objective or subjective proficiency. While some reported using an objective measure with the bilingual participants only (but not the monolinguals), most of these 43 studies did not report the proficiency of either group. An overview of these results can be found in Table 3.

<Insert Table 3 about here>

There was a statistically significant relationship between the type of proficiency reported and the age of the participants, X^2 (3, N = 184) = 50.83, $p < 0.001$. Studies with child participants were more likely to report objective proficiency than were studies with adults (45% of child studies vs. 8% of adult studies). Conversely, adult studies were more likely to report subjective proficiency only (54% of adult studies vs. 12% of child studies). This finding is not surprising because children are still in the process of acquiring one or both of their languages, and objective assessment that is sensitive to developmental milestones is more appropriate for developing children. Furthermore, parents may be less able to judge their children's relative proficiency in each language given that parents may not be present in all social situations involving their children, such as childcare or schools. The relationship between the type of proficiency reported

and region was also statistically significant, X^2 (12, N = 184) = 39.70, p < 0.001. Most notably, studies from Asia, Australia and Africa had the highest proportion of studies that did not report proficiency at all (60%).

Functional Usage

The amount of daily usage in each language is an important metric for bilingual experience. Functional usage is most often operationalized in the literature as the languages spoken at home. We found that 79% of the studies (n=147) reported some information about home language usage. However, 74 of these studies only reported home usage categorically. That is, they reported whether a particular language was spoken in the home, but did not quantify to what degree it was used. For example, "students were termed 'bilingual' if caretakers reported Spanish spoken in the home" (Gibson, Oller, Jarmulowicz, & Ethington, 2012, p. 105). A roughly equal amount (39% of the total, n=73) also reported home language usage as a gradient. In most cases, this meant that the proportion that each language was used in the home was estimated based on self- or parent-report. For example, one study used a parent questionnaire to determine the percentage of exposure per week, and reported that "Bilingual children were determined as those with at least 30% exposure to one of two languages weekly since birth" (Yow & Markman, 2011, p. 563). Table 4 provides an overview of home usage reporting across the studies.

<Insert Table 4 about here>

As with proficiency, the reporting of home usage differed significantly between the child and adult participants, X^2 (2, N = 186) = 14.53, p = .001. Studies with only child participants were more likely to report home usage than adult studies (86% of child studies vs. 68% of adult studies), but a slightly smaller proportion of child studies reported home usage as a gradient

(36% of child studies vs. 43% of adult studies) and a larger proportion of child studies reported home usage as categorical (50% vs. 25%). Home usage reporting was also related to region, X^2 (8, N = 186) = 15.55, p = .049. Studies from the US were more likely to report home usage as a gradient (50%) but were also more likely to not report home usage at all compared to other regions (29%). For studies with samples from multiple regions, a considerable proportion of studies (32%) did not report home usage at all. Collectively, home usage was reported to some extent by most of the studies, but less than half quantified the proportion of each language used.

School Language

The languages used for instruction in school can have a powerful impact on second language acquisition and first language maintenance or attrition. We found that 60% (n=101) of the studies reported the language of schooling.[5] As expected, significantly more child-only studies reported the language used for schooling than adult studies. (72% compared to 46%, X^2 (1, N = 168) = 11.45, p = .001). The difference by region was also significant, X^2 (4, N = 168) = 15.84, p = .003. Of the studies from Asia/Australia/Africa, 100% reported the language of schooling compared to only 46% from the US. This may be because the language of schooling is assumed to be English only in the US, while it may be more varied and complex in societies with multiple official languages.

Language History

Given the distinction between simultaneous and sequential bilinguals in the literature and the strong relationship between age of acquisition and ultimate attainment (Birdsong & Vanhove, 2016), the language acquisition history is an important facet of bilingual experience. We found

that 67% (n=124) of the studies reported the language history of the participants.[6] Again, this proportion differed significantly by the age of the participants, with only 56% of the child studies reporting language history compared to 83% of the adult studies, X^2 (1, N = 184) = 14.16, p < .001. While this discrepancy is understandable given that the adults, by definition, have more history to report, the age at which participants were first exposed to a second language and particularly when (and if) they began using that language productively on a regular basis is of great theoretical importance (e.g. Bedore et al., 2016; Kovelman, Baker, & Petitto, 2008). The regional differences in reporting language history were only marginal, X^2 (4, N = 184) = 8.53, p =.074, with studies from Canada least likely to report history (51%) compared to the other regions.

Sociolinguistic Context

Differences in the status and normative usage of languages in a society may moderate the impact of bilingual experience. For this reason, we were particularly interested in whether studies routinely report on how languages are viewed and used in the larger society from which their participants were recruited. We found that the sociolinguistic context was reported in only 30% (n=55) of the studies. As with proficiency, home usage, school language and history, the difference between child and adult studies was significant, X^2 (1, N = 186) = 9.59, p =.002. Among studies with child-only participants, 38% included information about the sociolinguistic context in the description of the sample, while this was true for only 17% of the adult studies. Regional differences were also significant, X^2 (4, N = 186) = 11.99, p =.017. Studies with US samples were the least likely to report on the sociolinguistic context (14%), while Asia, Australia

and Africa were the most likely (47%). This may be due to an assumption that readers are more familiar with the US context and less familiar with that of countries in Asia and Africa.

In sum, we found that proficiency and home usage were the two most frequently reported features. Looking at how the reporting of these features differs by the age of the participants, we found that studies of children were more likely to report objective proficiency, categorical home usage, the language of schooling, and the sociolinguistic context. Studies of adults were more likely to report subjective proficiency, home usage on a gradient scale, and the history of language acquisition. Intriguing regional differences in the reported features were also observed. For example, studies from Asia, Australia and Africa were less likely to report proficiency but more likely to report the language of schooling and the sociolinguistic context. In contrast, studies from the US were more likely to report objective measures of proficiency, but rarely reported on the sociolinguistic context.

Discussion

In this study, we examined the labels applied to bilingual and monolingual groups as well as the descriptions of participant backgrounds to better understand how bilingualism has been operationalized in the literature from 2005-2015. Furthermore, we asked if these labels and descriptions differed by the age of the participants or the region from which the participants were recruited. We found that the most frequently used labels referenced the languages spoken without qualifiers related to dominance, learning, or history. Studies from Canada were most likely to use more general labels, while studies from multiple regions used more specific labels, perhaps reflecting the relative heterogeneity of the samples in Canada. In our analysis of the descriptions of participants, we discovered that while proficiency and home usage were each reported by close to 80% of the studies, studies varied in how they measured these features.

School language and history were both reported more than half the time, but information on the sociolinguistic context was included in only 30% of the descriptions. The extent to which studies reported on these features differed significantly by both age of the participant and region.

Our findings confirm the complexities inherent in bilingual experience, a fact that complicates the interpretations and syntheses of research findings comparing monolinguals and bilinguals across the globe. Three main findings emerge from our documentation of how recent research has operationalized bilingualism. First, there is inconsistency or a lack of transparency in reporting bilinguals' characteristics across studies. Second, differential bilingual characteristics were reported in adults and children. Third, there is a lack of sociolinguistic information reported in the literature, despite the importance of considering bilingualism as an interactive experience. We consider these findings in turn.

Considerations on reporting language background characteristics

We found that while most of the labels used for bilingual groups only named the language pair they spoke, the descriptions revealed considerable heterogeneity of participants under the same label both within and across studies. One reason for this was overlapping between-group inclusionary criteria across studies. For example, Nguyen and Astington's 2014 study in the majority Francophone city of Montreal specified that their 3 to 5-year-old participants be "bilingual children [who] were (a) exposed to both English and French from birth or before 8 months of age and (b) generally exposed to each of these two languages for a minimum of 30% of the time" (p. 400). In contrast, Scheele, Leseman and Mayo (2010) "were interested in immigrant children who were being raised primarily in the language of their cultural community" (p. 122) in the Netherlands, and so they recruited 3-year olds who heard primarily Tarifit-Berber or Turkish at home, and excluded those with less than 70% input in one of these

languages. While both samples are "bilingual," we might expect qualitative differences between the simultaneous learners of two official languages from the first study and the language-minority sequential bilinguals from the second. In addition, language exposure and language usage represent distinct experiences, particularly for younger children. Therefore, clarifying the context and nature of language input as well as output is necessary to increase transparency in research.

A more common issue than conflicting criteria was underspecified criteria. If a study applied a generic label and provided a sparse description of the participants, it was difficult to infer whether the participants in the study were comparable to those in other studies. Because it is not realistic to expect that all researchers agree on a universal definition of bilingualism, it is important that each study explicitly state how bilingualism is being operationally defined and how their samples were selected. This, together with a robust description of the features discussed below, will allow for more thoughtful interpretations and cross-study syntheses.

Given how bilingualism has been conceptualized in the literature, the observed under-specification was somewhat surprising, particularly in the areas of proficiency, usage and history. First, while few current conceptions of bilingualism require "native-like" control of two languages" (Bloomfield, 1935, p.56, cited in Hamers & Blanc, 2000), the level of proficiency in each language remains an important criterion and source of variability (Cummins, 1987; Hulstijn, 2012). Seventy-seven percent of the studies reported proficiency to some degree, but many of these relied on subjective ratings or measures in only one language. Second, Grosjean conceptualizes bilingualism as "the use of two or more languages (or dialects) in everyday life" (2013, p. 5), yet only 79% of the studies we reviewed reported the languages used at home. Moreover, recent conceptualizations of bilingual development have highlighted the role of

absolute and relative input in each language (De Houwer, 2011, 2014; Hoff et al., 2012), and yet only 39% of the studies reported language usage on a gradient scale. Finally, research on bilingualism often hypothesizes a qualitative difference between simultaneous and sequential bilingualism (e.g. Montrul, 2009), and Birdsong and Verhoeven (2016) have emphasized the importance of age of acquisition in bilingualism, yet the history of language learning was only reported in 67% of the studies. These different conceptualizations emphasizing features such as proficiency, usage and history are not necessarily problematic as long as the criteria are clearly described and the characteristics of participants are reported in a way that facilitates cross-study comparisons.

Considerations on reporting characteristics for adults and children

The differences in the features reported by age group make intuitive sense: the features that are more apparent or quantifiable are not always the same for children and adults. However, we should supplement practical considerations with theoretical importance in determining the essential features to report for both groups. For example, based on the current research, we found that studies involving adults relied significantly more on subjective proficiency measures (e.g., *"*The participants rated their ability to speak, read, write, and comprehend English and Korean" Baker & Trofimovich, 2005), while studies involving children often included objective proficiency measures, such as standardized vocabulary assessments. This makes sense on pragmatic grounds, since adults can more readily report their perceived proficiency in a language. However, as Thomas (1994) noted in her review of how L2 proficiency was defined across 157 studies of second language acquisition, "impressionistic judgment of proficiency [is] a dubious basis for generalizing results beyond the sample or for comparing across groups." (1994, p. 315). While the subjective judgments reported may be sufficient for interpreting the

results of an individual study, they may be insufficient for making cross-study comparisons. It is recognized that, at least for adults, subjective judgement collected from questionnaire is an efficient way to capture language usage data. However, it is noted that self-rated proficiency may need to be collected in conjunction with language usage data as well as language history data to improve the validity of subjective judgment. For children, the overwhelming use of vocabulary tests to determine proficiency may reflect cultural knowledge rather than comprehensive language proficiency. While vocabulary is a relatively direct measure of knowledge of words, other measures that provide more comprehensive assessment of language proficiency should be considered (e.g., BESA (Peña, Gutiérrez-Clellen, Iglesias, Goldstein, & Bedore, 2014) and QUILS (Golinkoff, De Villiers, Hirsh-Pasek, Iglesias, & Sweig Wilson, 2017) for young children).

We noted that several studies assessed the bilingual group's proficiency but not that of the monolingual group. Including comparable assessments will help readers to understand to what degree the bilingual and monolingual samples are homogeneous. While bilinguals often demonstrate lower verbal abilities than monolinguals (Bialystok, Luk, Peets, & Yang, 2010; Oller, Pearson, & Cobo-Lewis, 2007; Portocarrero et al., 2007), statistical dispersions of the samples can be considered as indications of sample homogeneity. This is particularly necessary when language of testing is conducted in the second language of the bilinguals but in the first (and only) language spoken by the monolinguals. Notably, language skills may differ even within speakers of a single language, and so an objective assessment of monolingual proficiency is also called for. Finally, it is worth noting that many of the standardized measures of language did not consider bilingual or linguistically diverse children in their norming process (Luk & Christodoulou, 2016). Given this observation, cautious interpretation of the standardized

measures is warranted in order to avoid framing bilingual children as "delayed" or having a "disadvantage" based on isolated language assessments. In addition, many of the standardized normed tests were developed to evaluate children and adults for atypical language behavior instead of evaluating language proficiency. Therefore, selection of tests to measure (second) language proficiency should be aligned with the appropriate uses of the tests.

We also found that 27% of the child-only studies report neither objective nor subjective proficiency. This may be due to a lack of suitable instruments for assessing very young children, in which case the amount of exposure and/or productive usage may be a suitable proxy. This brings us to our finding that while child-only studies were more likely to report home language usage, only 36% reported this on a gradient scale. Since research has shown that the proportion each language is spoken at home may be an important predictor of variation in bilingual development (Paradis, 2017; Peña, Bedore, & Kester, 2016), this may be one area where future research could take a more fine-grained approach. While parent reports generally provide a reasonably reliable estimate of home usage (Marchman & Martinez-Sussmann, 2002), direct observations of child-directed speech (i.e., speech produced by adults in the environment that is addressed to children) using all-day recordings may offer a more accurate snapshot of the home language environment that is predictive of processing skills in each language (Marchman, Martínez, Hurtado, Grüter, & Fernald, 2017). The proportion of current usage may also be important for adults, but studies must be mindful that their usage spans many spheres such as home, work, school, and social life, and may depend more on the interlocutor than the context. For both adults and children, the contexts for exposure to languages and the actual usage of languages may have differential implications on individual bilingual experience.

For children, the main context for language use outside of the home is the school, and so the language of instruction and opportunities to use or learn additional languages at school is critical to understanding a child's current bilingual experience. As expected, we found that the language of schooling was reported more for child-only studies than for adult studies. However, 28% of studies of school-aged children did not provide this information. This could be due to assumptions that the school language can be inferred based on the location of the study, but there is considerable variation within countries such as the US and Canada in the degree to which bilingual children experience English-only or bilingual instruction. While seemingly less important for adult participants, the language through which they experienced their education provides important clues as to their language history and literacy skills in one or both of their languages. Therefore, we recommend that the language of schooling be attended to for school-aged children and adult university students alike.

Language history was more likely to be reported for adult participants, which is understandable given their longer history of language learning. However, this is another feature that makes sense to report for all participants, regardless of their age. Also, since judgments about the line between simultaneous and sequential bilinguals differ, reporting the age of exposure and active usage will facilitate generalizations across studies. As bilingualism is an experience relying on an individual interacting with others through languages, language history is especially important in capturing the dynamic changes of language exposure and use over time. As Grosjean (2013) points out, bilinguals' language usage often waxes and wanes over time, and events such as moving to a new country or losing a family member can have a large impact. Thus, more detailed information on the history of both language learning as well as language loss would provide a more complete understanding of the complexity of the bilingual

experience. To capture the dynamic nature of bilingualism, it is crucial that research employing longitudinal design should be adopted. Currently, there is a scarcity of longitudinal research involving bilinguals to fully understand the dynamic changes associated with development and learning in bilinguals.

Considerations on the social contexts relevant to bilingualism

The social contexts for learning and usage, both in the immediate sense of home, school, work and community and in the larger sense of sociopolitical contexts, are important for both child and adult studies. Since bilingualism can be thought of as an individual experience (Mackey, 1962), it is easy to lose sight of how social contexts can shape what it means to be bilingual. Both the definition and the psychological experience of being bilingual may differ greatly depending on the social norms around using multiple languages as well as the status of different languages in the society (Hamers & Blanc, 2000). Surprisingly, we observed that the sociolinguistic context where the study was conducted was the least reported of the features we coded. While child-only studies were more likely to report on the sociolinguistic context, this was *not* reported in 83% of the adult studies. It is important to note that some studies may have commented on the sociolinguistic context outside of the participant description, but we were interested in whether the relative status and usage of languages in the population were considered as integral to the participants' experience and a potential factor in interpreting findings.

The scarcity of attention to this facet of the bilingual experience suggests a tendency to focus on individual language experiences and provide little to no information on the social contexts where languages were used (and perceived). Considering the geopolitical setting is important for all inquiry into human development and cognition, but it may be particularly salient for individuals whose daily communications are colored by the status of the languages

they speak. The influence of the larger social and political context on the individual is integral to theories of child development (e.g. Bronfenbrenner, 1977; Sameroff, 2010), and sociolinguists have explored how macro-level issues of language and power influence individual experiences (Cummins, 2000; Fishman & Cooper, 1969; García, 2009). Bilingualism researchers are starting to consider the importance of the sociolinguistic context of participants (Green & Abutalebi, 2013; Hartanto & Yang, 2016), but we found this feature to be rarely reported or accounted for in the literature. While a thorough discussion of the importance of sociolinguistic context is beyond the scope of this paper, we encourage future studies to explore this dimension of bilingualism more explicitly. For example, recent studies have examined how the status of having a majority home language versus a minority home language is relevant to children's language use and development (Dijkstra, Kuiken, Jorna, & Klinkenberg, 2016) and adults' motivation to use one or both languages (Gasiorek & Vincze, 2016; Prochazka & Vogl, 2017).

In addition to the obvious differences in the social context, the very criterion for determining who is "bilingual" and "monolingual" may vary relative to local norms. For example, in some regional contexts such as the US (Kaushanskaya, 2012), Serbia and China (Basnight-Brown, Chen, Hua, Kostić, & Feldman, 2007) and Saudi Arabia (Gathercole & Moawad, 2010), having learned a foreign language in school (Spanish in the US case and English in the others) was the criterion for inclusion in the bilingual group. In other contexts, particularly urban centers in East Asia, foreign language instruction is so widespread that very few people could be counted as monolingual if this same criterion were applied. This limitation is evident in a study that drew its Korean-monolingual participants from Seoul: "Although a monolingual speaker is by definition an individual who speaks only one language … a more relaxed operationalization of the term has been adopted for the current study since most Koreans

start receiving English education from a very early age" (Park & Ziegler, 2014, p. 414). This was also the case for a study in Taiwan, which reported that, "50% of the first graders and 70% of the second graders received after-school English instruction. Excluding all of them from the study would have yielded a sample that was not representative of the first- and second-grade populations in Taiwan" (Kuo & Anderson, 2010, pp. 371–372). These examples illustrate that the prevalence and quality of foreign language instruction in schools can influence who qualifies as bilingual or monolingual in different studies. In these cases, bilingualism is not only a life experience, but also a learning experience.

While many of the features we recommend reporting are relevant to an individual's experience with languages, language use is *interactive* and is influenced by the values attached to language(s) in the society (Bak, 2016). We recommend that at the very least studies report the status and function of the languages used by the participants. It may also help to know the incidence of bilingualism in the society, and whether using multiple languages is encouraged or discouraged politically and socially. Finally, studies may want to consider how the context might influence both the selection of participants and the psychological experience of participants themselves. With all the individual and social factors considered, the binary categorization of "monolingual" and "bilingual" reflects sample-level group differences that is contextualized by the social perception of using majority and minority languages in regions where the study was conducted.

Conclusion

While recent years have seen an abundance of research comparing bilinguals and monolinguals on academic, linguistic and cognitive outcomes, our systematic review suggests that our

understanding of the effects of bilingualism is limited by three key methodological challenges. First, studies define and measure bilingualism differently, making cross-study comparisons difficult. Second, because bilingualism is a dynamic life experience, the factors that are most salient to report may be different for different ages of participants. Finally, study participants are embedded in different sociolinguistic contexts, and yet studies differ in the extent to which they report the status of the participants' languages in the society and consider how this may influence outcomes. After all, language is an interactive experience and it is essential to report the social context in which bilingual (and monolingual) participants use their languages.

In conclusion, not only is the use of multiple languages across and within the spheres of home, school, community and work becoming more common, but perceptions of the value of bilingualism have shifted in recent decades. A growing understanding of the benefits of bilingualism has spurred burgeoning interest in dual language immersion schools and bolstered efforts to maintain minoritized mother tongues alongside socially dominant languages worldwide. While the social benefits of speaking multiple languages are self-evident, inconsistent results from studies comparing bilingual and monolingual groups have raised questions about the cognitive benefits of bilingualism (Costa, Hernández, Costa-Faidella, & Sebastián-Gallés, 2009; Karlsson et al., 2015; Ross & Melinger, 2017). Considering the fact that understanding bilingualism relies on collaborative efforts from researchers around the world, it is imperative that we examine the nuances of what constitutes being bilingual in different studies before making a conclusive statement about how bilinguals categorically differ from monolinguals. After all, because research on bilingualism has important and significant societal value, a clearer understanding of how bilingual experiences shape development and learning is warranted.

References

Alvarado, N., & Jameson, K. A. (2011). Shared knowledge about emotion among Vietnamese and English bilingual and monolingual speakers. *Journal of Cross-Cultural Psychology*, *42*(6), 963–982. https://doi.org/10.1177/0022022110381125

Bak, T. H. (2016). Cooking pasta in La Paz: Bilingualism, bias and the replication crisis. *Linguistic Approaches to Bilingualism*. https://doi.org/10.1075/lab.16002.bak

Baker, W., & Trofimovich, P. (2005). Interaction of native- and second-language vowel system(s) in early and late bilinguals. *Language and Speech*, *48*(1), 1–27. https://doi.org/10.1177/00238309050480010101

Basnight-Brown, D. M., Chen, L., Hua, S., Kostić, A., & Feldman, L. B. (2007). Monolingual and bilingual recognition of regular and irregular English verbs: Sensitivity to form similarity varies with first language experience. *Journal of Memory and Language*, *57*(1), 65–80. https://doi.org/10.1016/j.jml.2007.03.001

Bedore, L. M., Fiestas, C. E., Peña, E. D., & Nagy, V. J. (2006). Cross-language comparisons of maze use in Spanish and English in functionally monolingual and bilingual children. *Bilingualism: Language and Cognition*, *9*(3), 233–247. https://doi.org/10.1017/S1366728906002604

Bedore, L. M., Peña, E. D., Griffin, Z. M., & Hixon, J. G. (2016). Effects of age of English exposure, current input/output, and grade on bilingual language performance. *Journal of Child Language*, *43*(3), 687–706.

Bialystok, E., Barac, R., Blaye, A., & Poulin-Dubois, D. (2010). Word mapping and executive functioning in young monolingual and bilingual children. *Journal of Cognition and Development*, *11*(4), 485–508. https://doi.org/10.1080/15248372.2010.516420

Bialystok, E., & Luk, G. (2012). Receptive vocabulary differences in monolingual and bilingual adults. *Bilingualism: Language and Cognition, 15,* 397-401.

Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. Bilingualism: Language and Cognition, 13(4), 525–531. https://doi.org/10.1017/S1366728909990423

Birdsong, D. (2016). Dominance in bilingualism: Foundations of measurement, with insights from the study of handedness. In J. Treffers-Daller & C. Silva-Corvalán (Eds.), *Language dominance in bilinguals: Issues of measurement and operationalization*. Cambridge, United Kingdom: Cambridge University Press.

Birdsong, D., & Vanhove, J. (2016). Age of Second-Language Acquisition: Critical Periods and Social Concerns. In E. Nicoladis & S. (Linguist) Montanari (Eds.), *Bilingualism across the lifespan: factors moderating language proficiency* (pp. 163–181). Washington, DC: American Psychological Association: Walter de Gruyter.

Bloomfield, L. (1935). *Language*. London: G. Allen & Unwin, Ltd.

Bronfenbrenner, U. (1977). Toward an experimental ecology of human development. *American Psychologist*, *32*(7), 513–531. https://doi.org/10.1037/0003-066X.32.7.513

Byers-Heinlein, K., & Werker, J. F. (2009). Monolingual, bilingual, trilingual: Infants' language experience influences the development of a word-learning heuristic. *Developmental Science*, *12*(5), 815–823. https://doi.org/10.1111/j.1467-7687.2009.00902.x

Carroll, S. E. (2017). Exposure and input in bilingual development. *Bilingualism: Language and Cognition*, *20*(1), 3–16.

Costa, A., Hernández, M., Costa-Faidella, J., & Sebastián-Gallés, N. (2009). On the bilingual advantage in conflict processing: Now you see it, now you don't. *Cognition*, *113*(2), 135–149. https://doi.org/10.1016/j.cognition.2009.08.001

Cummins, J. (1987). Bilingualism, language proficiency, and metalinguistic development. In P. Howel, M. Palij, & D. Aaronson (Eds.), *Childhood bilingualism: Aspects of linguistic, cognitive, and social development* (pp. 57–74). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Cummins, J., 1949-. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire* (Vol. 23). Clevedon; Buffalo, N.Y.: Multilingual Matters.

De Houwer, A. (2011). Language input environments and language development in bilingual acquisition. *Applied Linguistics Review*, *2*, 221–240.

De Houwer, A. (2014). The absolute frequency of maternal input to bilingual and monolingual children. In *Input and Experience in Bilingual Development* (pp. 37–57). John Benjamins Publishing Company.

Dijkstra, J., Kuiken, F., Jorna, R. J., & Klinkenberg, E. L. (2016). The role of majority and minority language input in the early development of a bilingual vocabulary. *Bilingualism: Language and Cognition*, *19*(01), 191–205. https://doi.org/10.1017/S1366728915000012

Engel de Abreu, P. M. J., Cruz-Santos, A., Tourinho, C. J., Martin, R., & Bialystok, E. (2012). Bilingualism enriches the poor: Enhanced cognitive control in low-income minority children. *Psychological Science*, *23*(11), 1364–1371. https://doi.org/10.1177/0956797612443836

Fishman, J. A., & Cooper, R. L. (1969). Alternative measures of bilingualism. *Journal of Verbal Learning and Verbal Behavior*, *8*(2), 276–282. https://doi.org/10.1016/S0022-5371(69)80076-9

García, O. (2009). *Bilingual education in the 21st century: A global perspective*. Malden, MA: Wiley-Blackwell.

Gasiorek, J., & Vincze, L. (2016). Modeling motives for bilingual accommodation by minority and majority language speakers. *Journal of Language and Social Psychology*, *35*(3), 305–316. https://doi.org/10.1177/0261927X15605144

Gathercole, V. C. M., & Moawad, R. A. (2010). Semantic interaction in early and late bilinguals: All words are not created equally. *Bilingualism: Language and Cognition*, *13*(04), 385–408. https://doi.org/10.1017/S1366728909990460

Gibson, T. A., Oller, D. K., Jarmulowicz, L., & Ethington, C. A. (2012). The receptive–expressive gap in the vocabulary of young second-language learners: Robustness and possible mechanisms. *Bilingualism: Language and Cognition*, *15*(1), 102–106. https://doi.org/10.1017/S1366728910000490

Golinkoff, R. M., De Villiers, J., Hirsh-Pasek, K., Iglesias, A., & Sweig Wilson, M. (2017). *Quick Interactive Language Screener (QUILS)*. Baltimore, MD: Brookes Publishing Co. Retrieved from http://products.brookespublishing.com/Quick-Interactive-Language-Screener-QUILS-P1029.aspx

Green, D. W., & Abutalebi, J. (2013). Language control in bilinguals: The adaptive control hypothesis. *Journal of Cognitive Psychology*, *25*(5), 515–530. https://doi.org/10.1080/20445911.2013.796377

Grosjean, F. (1989). Neurolinguists, beware! The bilingual is not two monolinguals in one person. *Brain and Language*, *36*(1), 3–15. https://doi.org/10.1016/0093-934X(89)90048-5

Grosjean, F. (2013). Bilingualism: A short introduction. In F. Grosjean & P. Li (Eds.), *The psycholinguistics of bilingualism* (Vol. 2nd, p. 5). Hoboken, NJ: Wiley-Blackwell.

Hamers, J. F., & Blanc, M. (2000). *Bilinguality and bilingualism* (2nd ed.). Cambridge; New York: Cambridge University Press.

Hartanto, A., & Yang, H. (2016). Disparate bilingual experiences modulate task-switching advantages: A diffusion-model analysis of the effects of interactional context on switch costs. *Cognition*, *150*, 10–19. https://doi.org/10.1016/j.cognition.2016.01.016

Hoff, E., Core, C., Place, S., Rumiche, R., Señor, M., & Parra, M. (2012). Dual language exposure and early bilingual development. *Journal of Child Language*, *39*(01), 1–27. https://doi.org/10.1017/S0305000910000759

Hulstijn, J. H. (2012). The construct of language proficiency in the study of bilingualism from a cognitive perspective. *Bilingualism: Language and Cognition*, *15*(2), 422–432.

Karlsson, L. C., Soveri, A., Räsänen, P., Kärnä, A., Delatte, S., Lagerström, E., … Laine, M. (2015). Bilingualism and performance on two widely used developmental neuropsychological test batteries. *PLoS ONE*, *10*(4).

Kaushanskaya, M. (2012). Cognitive mechanisms of word learning in bilingual and monolingual adults: The role of phonological memory. *Bilingualism: Language and Cognition*, *15*(3), 470–489.

Kaushanskaya, M., & Prior, A. (2015). Variability in the effects of bilingualism on cognition: It is not just about cognition, it is also about bilingualism. *Bilingualism: Language and Cognition*, *18*(01), 27–28. https://doi.org/10.1017/S1366728914000510

Kormi-Nouri, R., Moradi, A.-R., Moradi, S., Akbari-Zardkhaneh, S., & Zahedian, H. (2012). The effect of bilingualism on letter and category fluency tasks in primary school children: Advantage or disadvantage? *Bilingualism: Language and Cognition*, *15*(02), 351–364. https://doi.org/10.1017/S1366728910000192

Kovelman, I., Baker, S. A., & Petitto, L.-A. (2008). Age of first bilingual language exposure as a new window into bilingual reading development. *Bilingualism: Language and Cognition*, *11*(2), 203–223. https://doi.org/10.1017/S1366728908003386

Kroll, J. F., & Bialystok, E. (2013). Understanding the consequences of bilingualism for language processing and cognition. *Journal of Cognitive Psychology*, *25*(5), 497–514. https://doi.org/10.1080/20445911.2013.799170

Kuo, L.-J., & Anderson, R. C. (2010). Beyond cross-language transfer: Reconceptualizing the impact of early bilingualism on phonological awareness. *Scientific Studies of Reading*, *14*(4), 365–385. https://doi.org/10.1080/10888431003623470

Luk, G. (2015). Who are the bilinguals (and monolinguals)? *Bilingualism: Language and Cognition*, *18*(01), 35–36. https://doi.org/10.1017/S1366728914000625

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. https://doi.org/10.1080/20445911.2013.795574

Luk, G., & Christodoulou, J. A. (2016). Assessing and understanding the needs of dual-language learners. In N. K. Lesaux & S. M. Jones (Eds.), *Leading edge of early childhood*

education: linking science to policy for a new generation (pp. 67–90). Cambridge, Massachusetts: Harvard Education Press.

Luk, G., de Sa, E., & Bialystok, E. (2011). Is there a relation between onset age of bilingualism and enhancement of cognitive control? *Bilingualism: Language and Cognition*, *14*(04), 588–595. https://doi.org/10.1017/S1366728911000010

Mackey, W. F. (1962). The description of bilingualism. *Canadian Journal of Linguistics*, *7*, 51–85.

Marchman, V. A., Martínez, L. Z., Hurtado, N., Grüter, T., & Fernald, A. (2017). Caregiver talk to young Spanish-English bilinguals: comparing direct observation and parent-report measures of dual-language exposure. *Developmental Science*, *20*(1), e12425. https://doi.org/10.1111/desc.12425

Marchman, V. A., & Martinez-Sussmann, C. (2002). Concurrent validity of caregiver/parent report measures of language for children who are learning both English and Spanish. *Journal of Speech, Language, and Hearing Research*, *45*(5), 983–997.

Marinis, T., & Chondrogianni, V. (2010). Production of tense marking in successive bilingual children: When do they converge with their monolingual peers? *International Journal of Speech-Language Pathology*, *12*(1), 19–28. https://doi.org/10.3109/17549500903434125

McNeill, B., & Everatt, J. (2013). Meta-linguistic predictors of word-level literacy skills in monolingual and language minority learners. *Speech, Language and Hearing*, *16*(2), 92–98. https://doi.org/10.1179/2050571X13Z.00000000012

Montrul, S. (2009). Reexamining the fundamental difference hypothesis: What can early bilinguals tell us? *Studies in Second Language Acquisition*, *31*, 225– 257. https://doi.org/10.1017/S0272263109090299

Nguyen, T.-K., & Astington, J. W. (2014). Reassessing the bilingual advantage in theory of mind and its cognitive underpinnings. *Bilingualism: Language and Cognition*, *17*(2), 396–409. https://doi.org/10.1017/S1366728913000394

O'Hora, D., Pelaez, M., & Barnes-Holmes, D. (2005). Derived relational responding and performance on verbal subtests of the WAIS-III. *The Psychological Record*, *55*(1), 155–175.

Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007). Profile effects in early bilingual language and literacy. Applied Psycholinguistics, 28(2), 191–230. https://doi.org/10.1017/S0142716407070117

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. https://doi.org/10.1016/j.cogpsych.2012.12.002

Paradis, J. (2017). Parent report data on input and experience reliably predict bilingual development and this is not trivial. *Bilingualism: Language and Cognition*, *20*(1), 27–28. https://doi.org/10.1017/S136672891600033X

Paradis, J., & Nicoladis, E. (2007). The influence of dominance and sociolinguistic context on bilingual preschoolers' language choice. *International Journal of Bilingual Education and Bilingualism*, *10*(3), 277–297. https://doi.org/10.2167/beb444.0

Park, H. I., & Ziegler, N. (2014). Cognitive shift in the bilingual mind: Spatial concepts in Korean–English bilinguals. *Bilingualism: Language and Cognition*, *17*(02), 410–430. https://doi.org/10.1017/S1366728913000400

Peal, E., & Lambert, W. E. (1962). The relation of bilingualism to intelligence. *Psychological Monographs: General and Applied*, *76*(27), 1–23. https://doi.org/10.1037/h0093840

Peña, E. D., Bedore, L. M., & Kester, E. S. (2016). Assessment of language impairment in bilingual children using semantic tasks: two languages classify better than one. *International Journal of Language & Communication Disorders*, *51*(2), 192–202. https://doi.org/10.1111/1460-6984.12199

Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B., & Bedore, L. M. (2014). *Bilingual English Spanish Assessment*. San Rafael, CA: A-R Clinical Publ.

Portocarrero, J. S., Burright, R. G., & Donovick, P. J. (2007). Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*, *22*(3), 415–422. https://doi.org/10.1016/j.acn.2007.01.015

Prochazka, K., & Vogl, G. (2017). Quantifying the driving factors for language shift in a bilingual region. *Proceedings of the National Academy of Sciences*, *114*(17), 4365–4369. https://doi.org/10.1073/pnas.1617252114

Rolla San Francisco, A., Mo, E., Carlo, M., August, D., & Snow, C. (2006). The influences of language of literacy instruction and vocabulary on the spelling of Spanish-English bilinguals. *Reading and Writing*, *19*(6), 627–642. https://doi.org/10.1007/s11145-006-9012-3

Rolstad, K., Mahoney, K. S., & Glass, G. V. (2005). Weighing the evidence: A meta-analysis of bilingual education in Arizona. *Bilingual Research Journal*, *29*(1), 43–67.

Ross, J., & Melinger, A. (2017). Bilingual advantage, bidialectal advantage or neither? Comparing performance across three tests of executive function in middle childhood. *Developmental Science*, *20*(4), e12405. https://doi.org/10.1111/desc.12405

Sameroff, A. (2010). A unified theory of development: A dialectic integration of nature and nurture. *Child Development*, *81*(1), 6–22. https://doi.org/10.1111/j.1467-8624.2009.01378.x

Scheele, A. F., Leseman, P. P. M., & Mayo, A. Y. (2010). The home language environment of monolingual and bilingual children and their language proficiency. *Applied Psycholinguistics*, *31*(01), 117–140. https://doi.org/10.1017/S0142716409990191

Solari, E. J., Petscher, Y., & Folsom, J. S. (2014). Differentiating literacy growth of ELL students with LD from other high-risk subgroups and general education peers: Evidence from grades 3–10. *Journal of Learning Disabilities*, *47*(4), 329–348. https://doi.org/10.1177/0022219412463435

Thomas, M. (1994). Assessment of L2 proficiency in second language acquisition research. *Language Learning*, *44*(2), 307–336. https://doi.org/10.1111/j.1467-1770.1994.tb01104.x

Unsworth, S. (2008). Comparing child L2 development with adult L2 development: How to measure L2 proficiency. In E. Gavruseva & B. Haznedar (Eds.), *Current trends in child second language acquisition*. Amsterdam: John Benjamins.

Unsworth, S. (2016). Early child L2 acquisition: Age or input effects? Neither, or both? *Journal of Child Language*, *43*(3), 608–634. https://doi.org/10.1017/S030500091500080X

Yow, W. Q., & Markman, E. M. (2011). Bilingualism and children's use of paralinguistic cues to interpret emotion in speech. *Bilingualism: Language and Cognition*, *14*(04), 562–569. https://doi.org/10.1017/S1366728910000404

---

[1] The largest bilingual sample was from one multi-site study in Iran (Kormi-Nouri, Moradi, Moradi, Akbari-Zardkhaneh, & Zahedian, 2012); the next largest had a sample size of 240.

[2] Two studies (embedded in a single paper) were not coded for this feature, as the participants were newborn infants.

[3] Despite relying on parent report, we coded the MCDI as an objective proficiency measure because it compares the words known and produced by young children to a norming sample in order to assess relative language proficiency, instead of eliciting a global, subjective judgment.

[4] For the remaining eight studies that relied on objective proficiency in only one language, the language tested is one of the bilinguals' two first languages (in the case of simultaneous bilinguals, n=2), or the bilingual group includes both first and second-language speakers of the tested language (n=6).

[5] Eighteen studies were not coded for this feature, as the children were too young to be of school age.

[6] The same two infant studies that were not coded for proficiency were also not coded for this feature.

Table 1.
Labels used to describe bilingual groups

Label	Examples	n	%
Bilingual Only	*Bilingual*	58	31%
Language Pair Only	*Tagalog–English Bilingual, Mandarin-English bilingual, French-English Bilingual*	85	46%
Dominance Only	*Nearly Balanced Bilingual, Fully Bilingual*	3	2%
Learning Only	*Bilinguals in English language instruction, English Language Learners*	10	5%
History Only	*12-Month-Old Bilingual, Early Bilinguals*	11	6%
Combination of two or more	*Emergent bilingual Zulu–English, French-English Simultaneous Bilingual, Successive bilingual Turkish-speaking children*	19	10%

Table 2.
Description of the coding scheme

Feature	Qualifying Questions	Coding
Proficiency	Do I know the proficiency of the participants based on a **subjective** self, parent, or researcher rating? Do I know the proficiency of the monolingual and bilingual groups (in at least one language) based on an **objective** assessment?	0 = neither 1 = subjective only 2 = objective only 3 = both subjective & objective
Home Usage	Do I know which language(s) are spoken at home? Do I know what proportion each language is used at home, or in the participant's personal life?	0 = not reported 1 = categorical 2 = gradient
School Language	Do I know what the language of instruction is, based on what was reported in the study?	0 = not reported 1 = reported (coded as N/A if children were younger than school aged)
Language History	Do I know in which order the bilinguals learned their languages, and around what age they learned their second language?	0 = not reported 1 = reported
Sociolinguistic context	Do I know about the general status and usage of languages in the population where this study took place, based on what was reported in the study?	0 = not reported 1 = reported

Table 3.
Percentage of proficiency types reported in the literature, by age of participants and region

		Age		Region				
	Overall n=184	child n=108	adult n=76	US n=56	Canada n=47	EU/UK n=44	Asia/Aus/ Africa n=15	Mult. n=22
Objective Only	**30%**	45%	8%	23%	49%	23%	33%	23%
Subjective Only	**29%**	12%	54%	39%	15%	34%	0%	45%
Obj. & Subj.	**17%**	15%	21%	23%	23%	9%	7%	14%
Neither	**23%**	28%	17%	16%	13%	34%	60%	18%

Note: EU = European Union, UK = United Kingdom, Aus = Australia, Mult. = Multiple Regions.

Table 4.
Percentage of home usage types reported in the literature, by age of participants and region

		Age		Region				
	Overall n=186	child n=110	adult n=76	US n=56	Canada n=47	EU & UK n=46	Asia/Aus/ Africa n=15	Mult. n=22
Categorical	**40%**	50%	25%	21%	51%	46%	53%	41%
Gradient	**39%**	36%	43%	50%	32%	41%	33%	27%
Not reported	**21%**	14%	32%	29%	17%	13%	13%	32%

Note: EU = European Union, UK = United Kingdom, Aus = Australia, Mult. = Multiple Regions.