

**The influence of Education Policies for Bilingualism in Stakeholder's Language  
Assessment Practices and views in West-Center Colombia Language Education:  
A Multicase Study**

Luis Felipe Jaramillo-Calderón

Advisor: Prof. Daniel Murcia

Universidad Tecnológica de Pereira  
Facultad de Bellas Artes y Humanidades  
Maestría en Educación Bilingüe

2023

**The influence of Education Policies for Bilingualism in Stakeholder's Language  
Assessment Practices and views in West-Center Colombia Language Education:  
A Multicase Study**

Luis Felipe Jaramillo Calderón

Trabajo de grado como requisito para optar por el título de Magister en Educación Bilingüe

Universidad Tecnológica de Pereira  
Facultad de Bellas Artes y Humanidades  
Maestría en Educación Bilingüe

2023

### **Acknowledgements**

A mis padres y hermana, cuyo apoyo incondicional permitió la culminación de este proyecto dentro de los tiempos fijados. Su respaldo ha sido clave en momentos difíciles, y su entrega ha sido la inspiración para afrontar los retos de esta aventura.

Estoy convencido de que la calidad de una persona y profesional se debe en gran medida a la influencia de aquellos con quienes se rodea. Por esto, me considero muy afortunado por toparme en el camino mentores y colegas de excepcional calidad humana y profesional, especialmente el profe Daniel, mi asesor, de quien guardo profunda admiración e innumerables aprendizajes. También quiero expresar mi profundo agradecimiento a grandes docentes y profesionales como las profes Ángela María López Velásquez, Rosa María Guilleumas e Isabel Sánchez, quienes me brindaron su confianza, consejos y apoyo en el proceso. Este recorrido fue mucho más ameno gracias a mis compañeros de la maestría, principalmente el *pana* Cuervo y la Rox, con quienes abundó el café, las risas y las noches en vela. Por último, agradezco a todos mis amigos y seres queridos, tanto los cercanos como aquellos que ya no lo son tanto, quienes en algún momento me ofrecieron palabras de aliento y una mano amiga en el camino.

Con gratitud y nostalgia

**-Luis Felipe Jaramillo**

### **Abstract**

This multi-case study explores the influence of recent language policies on stakeholder's language assessment and evaluation practices in Colombian language education from the perspective of different stakeholders. Additionally, the study explores the language and education ideologies of language policies and Colombian policy stakeholders, and the way stakeholders exercise their agency in the formulation and application of the policies. Data includes interviews with stakeholders (policy makers, scholars, test designers, administrators of second-language teacher preparation programs and language centers, and teachers) from different cities of West-Center Colombia and recent national policy documents. The study shows an emerging shift in stakeholder's ideologies, discourses, and practices towards alternative, heteroglossic bilingual oriented views that are. Nonetheless, it also shows that practices are still constrained by traditional (de-facto) policies such as high-stakes tests and transnational policy trends and economic interests pushing towards accountability, standard-oriented monolingual practices. Additionally, it shows the active and relevant role of stakeholders in policy enactment processes and suggests, as well, the influence they have exercised on written policies and official discourses. The findings imply the need for participatory, context-sensitive policies prioritizing learner's needs, broadening language curricula and promoting heteroglossic teaching and assessment practices.

*Key words:* Language policy, Language Assessment/testing, Language Education

### **Resumen**

Este estudio de casos múltiples explora la influencia de las políticas lingüísticas recientes en las prácticas de evaluación de los involucrados en la educación lingüística de Colombia a partir de los casos estudiados. El estudio también busca captar las ideologías lingüísticas y educativas detrás de las políticas lingüísticas, de los involucrados en estas políticas y cómo estos últimos ejercen su agencia en la formulación y aplicación de las políticas. Los datos

incluyen entrevistas con los involucrados (ej., formuladores de políticas, académicos, diseñadores de pruebas, administradores de programas de preparación de docentes de segunda lengua y centros de idiomas, y docentes) y documentos de políticas nacionales recientes. El estudio muestra un cambio emergente en las ideologías, discursos y prácticas de las partes interesadas hacia puntos de vista bilingües heteroglosos alternativos que, sin embargo, todavía están restringidos por políticas tradicionales (de facto) como pruebas de alto impacto, tendencias de políticas transnacionales e intereses económicos que empujan los procesos hacia la rendición de cuentas o prácticas monolingües orientadas a estándares. Además, muestra el papel activo y relevante de los actores en los procesos de ejecución de las políticas y sugiere, también, la influencia que han ejercido en la formulación de las políticas escritas y los discursos oficiales. Los resultados implican la necesidad de políticas participativas y sensitivas al contexto, priorizando las necesidades de los aprendices, ampliando los currículos y promoviendo la enseñanza y evaluación heteroglosa.

*Palabras clave:* Política lingüística, Evaluaciones y examinación de lenguas, Educación en lenguas.

## Contents

Statement of the Problem .....	11
Purpose of this Study.....	16
Research Questions .....	17
Objectives .....	18
Theoretical Framework .....	19
Conceptual Background .....	19
The Bilingualism Complexities that Colombian Language Policies Have Failed to Grasp .....	19
From a Monoglossic to a Heteroglossic Approach to Bilingual Education.....	22
Language Assessment and Evaluation in Bilingual Education Settings .....	25
Exploring Language Policy on Bilingualism and Bilingual Education Through Comparative, Sociocultural and Critical Lenses .....	32
Language Assessment Policy.....	37
Literature Review.....	41
Official Language Assessment and Evaluation Practices under Scrutiny: Validity, Accountability, Washback and Impact.....	42
Language Policy: approaches and underpinning ideologies.....	45
Language Assessment and Language Policy interplay: traits towards standardization, outcome-oriented, measurement and accountability.....	48
Methodology .....	52
Research Approach.....	52
Positionality Statement .....	53
Context and Participants.....	54

Data collection Methods.....	57
Semi-Structured Interview.....	57
Documents.....	58
Data Analysis Method.....	58
Data Analysis Process.....	59
Findings and Discussion .....	63
Complexities of Bilingualism and Bilingual Education in Colombia: Divergent Positions; Changing Paradigms; Written, Enacted, and Perceived Policy Disparities.....	63
Divergent View of Bilingualism and Bilingual Education: the Case in Labeling Language Education in Colombia .....	63
Stakeholder’s Changing Paradigms on Bilingualism and Assessment.....	66
Written Policy Discourse vs Implemented Policy and Stakeholder’s Perspective .....	68
Influence of Language Policies in Assessment Practices: Guiding and Limiting Practices.	73
Influence of Transnational Policy Trends and Market Forces on Language Assessment Views in Colombia .....	78
Standards and Outcome-Oriented Views.....	78
Market Forces Influencing Language Policies and Assessment Practices .....	81
CEFR Power and Influence on Language Policies and Stakeholder’s Discourses.....	85
Emerging Practices on Language Assessment and Language Assessment Policies.....	88
Relevance of Formative and Alternative Assessment on Language Policies and Stakeholders’ Discourses and Practices. ....	88
Emerging Heteroglossic Bilingual Assessment Practices .....	91
Increasing Tech Integration to Language Assessment Practices.....	93
Stakeholders’ Active Engagement in Policy Enactment.....	97

Limitations .....	109
References .....	112
Appendices .....	126
Appendix 1: Informed Consent .....	126
Appendix 2: Interview Questions Examples .....	127
Appendix 3: Codebook .....	130
Appendix 4: Extracts in Original Language .....	131



**Tables**

**Table 1** *A priori Codes from research questions*.....60

**Figures**

<b>Figure 1.</b> tests, measurement, assessment, teaching and evaluation .....	26
<b>Figure 2</b> <i>Framework for Comparative Education Analysis</i> .....	35

### Statement of the Problem

*Colombia Bilingüe* has become not only a national project but a motto in educational and political discourses across the country. For several years now, bilingualism has gained a primary role in the language and educational policies of the country, especially since the National Bilingualism Program (*Programa Nacional de Bilingüismo, 2004-2019*) established guidelines, actions and resources to reach (Spanish-English) bilingualism in the country within fifteen years. The initial reaction of scholars and institutions led policymakers to revise the program and caused an awakening of debates on bilingual education in the country. At the time, several further actions and policies were introduced: *Programa de Fortalecimiento de Competencias en Lengua Extranjera 2010-2014, Programa Nacional de Inglés 2015-2025, Ley 1651, 2013: Ley de Bilingüismo, Colombia Bilingüe 2015-2018*). Even though the initial goal of bilingualism seems far from being reached (Bonilla Carvajal & Tejada-Sanchez, 2016; Alonso Cifuentes et al., 2018), many of the decisions that were questioned seem to be still in place: 1. the view of bilingualism as being able to speak English and Spanish (disregarding other forms of bilingualism) (Mackenzie, 2020); 2. the adoption of the Common European Framework of Reference (henceforth CEFR) as the guiding principles for curricular planning, language learning and teaching, and assessment; 3. the incorporation of a series of competency standards taken from the CEFR; and, 4. the increase of “English competency assessment procedures for both teachers and students around the country” (Correa and Usma, 2013, p. 227). These decisions suggest that standardized testing has come to play a predominant role in the evaluation of policy implementations and teacher results, using the CEFR competency labels as the metric to do so.

In the goal to become a bilingual country in English and Spanish (Graddol, 2006; Cenoz & Gorter, 2012) the Colombian Ministry of Education decided to establish in 2005 proficiency goals in English for students, expecting them to reach an A2 (or lower-intermediate) level in primary education, a B1 (intermediate) level in secondary education, a B2 (upper-intermediate) level by the end of undergraduate studies, and a C1 (advanced) level

in case of pre-service or in service language teachers (Jimenez et al., 2017). These proficiency expectations and the way they are measured by standardized high-stakes tests reduce the conception of bilingualism to the measurement of English proficiency in regard to some linguistic skills such as grammar, vocabulary, or reading comprehension disregarding, for example, the other competences bilinguals develop or even students' knowledge of the first language as valuable linguistic resources.

Additionally, bilinguals tend to be misconceived as the sum of two monolinguals with separate language systems (Hamers and Blanc, 2004; Escobar & Dillard-Paltrineri, 2015; Grosjean, 1989). Therefore, it seems that these policies fail to acknowledge that bilinguals, by virtue of being exposed and using to two linguistic codes, experience unique psycholinguistic and sociolinguistic dynamics. Even in contexts like Colombia, where Spanish is the majority language, for English language learners there is a dynamic relationship in terms of cognition between Spanish and English as their bilingualism develops. In our globalized world, people have more and easier contact with speakers of other languages. Therefore, it is fair to say that exposure to any language requires exposure to its culture, where learners not only have to learn the linguistic aspects of the language, but also the sociolinguistic components. Even further, bilinguals develop unique sociolinguistic features of communication when communicating in bilingual settings, making use of all semiotic resources at hand, including code-switching and code-mixing (García, 2009).

Authors have criticized the adoption of such language policies because of their limited view of language learning and bilingualism, as well as the apparent dissonance between the sociocultural and political conditions of the country and the objectives, procedures, funding and timelines of such policies. For instance, Cadavid et al. (2004) warn that “educational policies should not be implemented blindly without a careful analysis of our context and without knowing the real needs of our teachers and students” (p. 44). Language policies in Colombia have adopted foreign standards such as the CEFR that are not necessarily aligned to the realities of the intended population, that is, the wide array of

students of all contexts, socioeconomic conditions and backgrounds in the country, most of which, do not resemble the realities of the original populations of the CEFR.

Ayala and Álvarez (2005) also state that “imported standards not only deal with language policies of foreign countries but also with foreign curricula, syllabi, teaching-learning methodologies, testing, assessment, evaluation, and instructional materials” (p. 12). This is why the development of bilingual curricula ought to follow contextual features instead of adopting imported bilingual models (Peñafort, 2002). To sum, these authors agree that copying foreign standards or methods is a practice that must be avoided as it is a complicated process that has to bear in mind the conditions of the context of implementation.

As a reiteration of this simplification of the socio- and psycholinguistic bilingual realities, it is evident that, even though some guides such as Guía 22 (2006) (a set of basic English standards for the planning of the different grade levels of the school system) talk about sociolinguistic competence, national tests like SABER 11 and SABER PRO, official measuring tools of education outcomes at the school and university level respectively and gatekeeping tools in the case of Saber 11, focus on vocabulary, grammar, and reading, ignoring other language skills such as the productive ones and social and cognitive competences of bilinguals, which results in a “incomplete language profile of students” (Jimenez, et al., 2017, p. 134). These test decisions affect the education of students as teachers tend to teach to the test (Pan, 2009). As the washback hypothesis (Alderson & Wall, 1993) states, “tests that have important consequences will have (high) washback (effects)” (pp.20-21). For example, the SABER 11 test is a gatekeeper in the access to higher education. As a result, this test can have high washback effects such as teaching to the test and the perception of students as broken language learners (in case of poor results) rather than emerging bilinguals developing complex psycholinguistic and sociolinguistic skills.

The policy language ideology behind the tests reflects an incomplete view of bilingualism (Gómez, 2017; Valencia, 2005), which not only may present negative or unintended washback effects, but also suggest the poor validity of tests. Secondly, as Pan

(2009) argues, test validity has increased its scope and complexity beyond construct validity in the last two decades. Authors such as Messick (1989) introduce the term “consequential validity” in which addresses the social value and consequences of tests. In this sense, if the language tests that are used to measure bilingualism provide an incomplete bilingual profile of language users, and they have high stakes in a society for students, teachers, and institutions, then, it might be the case that these tests evidence a lack of (consequential) test validity.

As a response to the criticisms (e.g. Herazo et al., 2012; Mackenzie, 2020), several new actions have been taken since the National Bilingualism program. For example, the creation of the Suggested English Language Curriculum and the Basic Learning Rights of English (Colombia, MEN, 2016) and some of the conceptions on bilingualism seem to have been updated to some degree, for instance, the national bilingualism program for the years 2018 to 2022 talks about “functional bilingualism” (Colombia, MEN, 2021), a term we will explore in the data analysis. Therefore, it is important to explore the new language policies documents as well as exploring the different stakeholders’ views and practices so as to get a better sense of the current state of bilingualism, how it is or should be perceived and assessed in the country.

My own experience as a student, teacher, and administrative staff of a language center has been impacted by the issues presented above. As a student at a public school, I not only experienced poor English instruction, but also poor development of literacy skills in Spanish; additionally, both languages were kept apart. When the national exams were approaching, teachers focused their instruction only on the test. Teachers gave students practice tests with questions similar to the actual exam and gave tips to get higher scores such as guessing the response even though we did not fully understand the questions. I took a national test where the English language section only included questions about vocabulary, language use and reading comprehension. I had a very similar experience with the *Exámenes de Calidad de Educación Superior* (ECAES), the national exams for undergraduate students. As a teacher and administrator, I have also noticed how the first language is disregarded as

an important factor in learning English. Foreign and decontextualized policies and assessment/evaluation instruments are blindly implemented as the foreign and Eurocentric productions are acknowledged as the valid voices in the English as a second language (ESL) domain.

This research explores the relatively unexplored terrain of language assessment policies in bilingual education and fosters a deeper debate in the field aiming to better understand it. Additionally, while several research focuses on either the macro level of governance or the microlevel of the classroom itself, this study focuses on the in-between level of policymakers, administrators, and academics which is less explored in these types of research. More importantly, this study will be conducted in a conjunctural moment when the deadlines imposed by the government to improve bilingualism rates (or English language learning) in the country are concluding, such as *Programa Nacional de Inglés* (PNI 2015-2025) and *Programa Nacional de Bilingüismo 2018-2022* and new policies will most likely be introduced. These future policies must be better informed by the academic community and should be based on research. This study is a research effort to address that need.

### **Purpose of this Study**

This research project intends to explore and describe, from the different study cases, nine stakeholders from West-Center Colombia and two national policy documents, the language assessment and evaluation practices in Colombia second language education and the interplay of practices and language policies in the country. This can help to determine the influence of the language policies in the assessment and evaluation practices in Colombian bilingual education as well as other possible influencing factors that come into play. Acknowledging that policies are a conflicting terrain, this research also intends to characterize how the stakeholders of the English language policies exercise their agency in the formulation and/or enactment of the policy. Finally, this research explores the ideologies about language [bilingualism], bilingual education and assessment behind relevant language policies that have attempted to influence the assessment and evaluation practices of the country. To do so, a multicase study was conducted, in which nine stakeholders of language policies and language education system across the western-center country such as policymakers, administrators, teachers, test-designers and scholars were interviewed, and two recent policy documents guiding the policy actions of the Programa Nacional de Bilingüismo were analyzed.



### **Research Questions**

This study intends to answer the following questions:

- What is the influence of Colombian Language policies on the assessment and evaluation practices of different stakeholders from the bilingual educational system in the west-center side of the country?
- How do the stakeholders of the English language policies in west-center Colombian education exercise their agency in their practices and influence in policy making and enactment?
- What are the language and education ideologies behind the selected language assessment policies of the Colombian Educational system and stakeholders from the west-center side of the country?

### **Objectives**

This study is guided by the following objectives:

- Explore the influence of the Colombian Language policies on assessment and evaluation practices from the cases of different stakeholders of the west-center side of the country.
- Explore the way west-center Colombian stakeholders' agency is restricted or manifested, how they are influenced by language policies, and they influence language policy making and assessment practices.
- Explore the language and education ideologies behind selected policies and stakeholders.

## **Theoretical Framework**

### **Conceptual Background**

The construct language education emerged as the discussion of the project allowed for an in-depth reflection on the divergence and relation of bilingual education and language education in Colombia. This examination focused on the interaction of languages within educational environments of varying natures and the development of (multiple) languages in students' overall language development. Therefore, for the purposes of this project, it is relevant to problematize the concepts of bilingualism and bilingual education in the Colombian context, especially since they have been established as the goals of education policies for bilingualism in the country.

### ***The Bilingualism Complexities that Colombian Language Policies Have Failed to Grasp***

Bilingualism refers to the state of languages in contact (two or more) at the individual, interpersonal or global level. In other words, it refers either to “the use of two or more codes in interpersonal and intergroup relations [also called social bilingualism] as well as the psychological state of an individual who uses more than one language [also called bilinguality or individual bilingualism]” (Hamers & Blanc, 2004, p.6). Even though, it is used to refer to the state of two languages in contact and there was a distinction between bilingualism and multilingualism (more than two languages), some authors now use bilingualism as an umbrella term for the learning and use of two or more languages and language varieties (De Houwer & Ortega, 2019).

Traditional views on individual bilingualism held the notion that a bilingual individual was expected to perform as a native speaker in both languages (Bloomfield, 1935). However, those views have been questioned and reconsidered. Not only was the notion of *native-like proficiency* a vague standard as native speakers of all languages have varied degrees of proficiency, but it also stigmatized foreign accents and it ignored the fact that competence in two or more languages is rarely equivalent (Birdsong, 2016; Grosjean, 1996; Romaine, 1989; Treffers- Daller, 2019; and Silva- Corvalán & Treffers-Daller, 2015).

More recent conceptions about bilingualism identify the relativity of bilinguals' competence and state that bilingualism is rather a matter of degree in a continuum between bilingualism and monolingualism (Grosjean, 1989). The frequent use of more than one language, rather than the degree of competence in each language, is what determines bilingualism. These later definitions also tend to prioritize *use* over *competence*, stating that bilingualism is the frequent use of two or more languages or language varieties. Finally, they highlight the fact that bilinguals are not the sum of two or more monolinguals as they have developed a unique language behavior (Grosjean, 1985a).

As there are so many ways of being bilingual, there are categories that help to better understand this social phenomenon. For the present study, three aspects of bilingualism will be used: competence in both languages, skills, and social status of the languages. In respect of competence, Lambert (1955) differentiated between balanced and dominant bilinguals. Balanced bilinguals have a roughly equivalent competence in both languages while dominant bilinguals have a stronger language, normally the mother tongue. The term *emergent bilinguals* is used to refer to individuals who are in the initial steps of learning a second language. Based on official reports by government institutions such as ICFES (e.g. 2021), it could be said that most English learners in Colombia are either emergent bilinguals or dominant bilinguals, with English being the weaker language.

The second relevant category has to do with the language skills bilinguals have or use. People tend to associate bilingualism with being able to *speak* a second language, however this is not the only case. Some authors (Döpke, 1992; Bialystok, 2001; Schüppert & Gooskens, 2012) distinguish between receptive and productive bilinguals, the former being able to understand a language while not being able to produce a lot in it. This is the case, for example, of some emergent or early bilinguals as well as bilinguals that have been exposed to lots of input in the second language but had not produced a lot in it. Additionally, it has to do also with learning experiences and actual communicative needs (Bialystok, 2001), that is, while some bilinguals may dispose of all four skills (reading, listening, writing and speaking) for their daily base communication, some bilinguals may either need receptive skills, the

ability to read and write (bilitracy) (for example a literature or philosophy professor or the ability to listen and speak (for example some immigrant workers).

Finally, some authors notice that not all languages or bilinguals are treated the same, but it depends on the social position of the learners or language users and the dominance and market value of the languages that make up their bilingualism (García, 2019) so these authors distinguish between elite bilingualism and folk bilingualism (Fishman, 1977). Elite bilingualism is the presence of at least two languages one of which is an European language, normally it is the bilingualism of Elite speakers, individuals belonging to the dominant social group who learn a second dominant language either for prestige or economic advantage. On the other hand, the bilingualism of Indigenous, colonized, refugee or immigrant people, people with unofficial non-standardized languages or the bilingualism of the deaf people is treated as problematic by government and institutions and systematic actions are normally in place to eradicate this type of bilingualism (García, 2009). The latter type of bilingualism is what Fishman (1977) called *Folk bilingualism*. Several authors in Colombia have identified and denounced that the Bilingualism promoted by language policies such as Colombia Bilingüe is a simplified bilingualism that only favors a powerful language, that is, they promote Elite bilingualism.

In this sense, it could be said that most people in the world are bilingual in some way and there are plenty of bilingual societies, such is the case of Colombia, a country with not only Spanish presence but also more than 60 native languages, a creole language, an official sign language and the presence of some prestigious languages such as English or French to some level in the educational and other work and social domains. Nonetheless, most people do not perceive Colombia as a bilingual society, and do not know of or value the existence of such variety of languages; non-prestigious languages like Colombian Sign Language and indigenous languages are underrepresented and undervalued in the Colombian societies and most of their policies (Gómez, 2017; Guerrero, 2008; Mackenzie, 2020; Mora et al., 2019; Quintero Polo, 2009; Usma, 2009).

For many years, bilinguals have been treated as two monolinguals and their languages have been separated in educational contexts (García, 2009) while several researchers and studies have shown strong reasons and evidence on why bilingual people are more than two monolinguals and their language processes and behavior have developed unique characteristics (Hamers & Blanc, 2004, García 2014, 2019). As García and Lasagabaster and García (2014) say, “the process by which a student develops their bilingualism is dynamic due to the different proficiency levels the learner has of their two languages” (p.3). Bilingualism is also dynamic in the sense that bilingual people in bilingual contexts may take advantage of all semiotic resources available, including code-switching or code-mixing, as for achieving a communicative purpose (García, 2019). In a similar sense, Cummins has argued for the importance of developing both languages in bilinguals and taking advantage of each language, especially language one, for the development of the other as they are interrelated and follow similar cognitive processes, what he called common underlying proficiency (Cummins, 1984a). He highlights the importance of acknowledging or the linguistic baggage in L1 of language learners.

### ***From a Monoglossic to a Heteroglossic Approach to Bilingual Education***

Just as bilingualism tends to be related to the good command of two languages, bilingual education is commonly thought of as the use of two languages in education; however, likewise bilingualism [and bilingual education are] simple label[s] for very complex phenomena (Cazden & Snow, 1990). Baker (2001) reports that bilingual education is sometimes used to refer to the education of already bilingual people and other times used for the education of people learning an additional language. That is, the education could be either for speakers of the majority language of their society learning an additional foreign or second language or for speakers of a minority/minoritized language learning the dominant language. Authors such as García (2009), one of the most prominent authors on bilingual education, differentiate between “traditional language education programs” and “bilingual education” (p.17). For the author, traditional language education programs normally teach

the additional language as a subject while bilingual education draws on both languages as a medium of instruction; in other words, the additional language is used to some higher or lower level as the language for instruction in content subject(s) in bilingual education.

García (2009), nonetheless, emphasizes that in the 21st century the boundaries between traditional language education programs and bilingual education are becoming blurred as the former programs are increasingly integrating content and language while the latter, bilingual education programs, are paying more attention to the explicit language instruction besides the content subjects. Furthermore, while aiming to teach a second language in an isolated way, language education programs tend to evidence bilingual practices by the contact of languages either in the materials or the communicative practices of teachers and students. Similarly, some so-called bilingual programs are aiming at monolingual practices of only using a single language in instruction in a given subject, thus, making the boundary blurrier. At the end, what keeps these two types of programs apart from each other are their end-goals, bilingual education aiming for the use of two languages to “educate generally, meaningfully, equitably, and for tolerance and appreciation of diversity” and traditional language programs simply aiming at learning a second language (p. 18).

Even though traditional language programs, programs in which education is in a student's first language and the second or foreign language is studied as an isolated subject, tend not to be considered as bilingual education programs, it is nonetheless a type of education for bilingualism. This type of education, what Baker (2001) called “MAINSTREAM [education] with Foreign Language Teaching” (p. 194) is considered to have a limited enrichment in terms of societal end education goals and a limited performance in case of language outcomes; it is a “weak form of education for bilingualism” (p. 194) according to the author.

Another way to classify bilingual education and other language education programs is based on their aims. Ferguson et al. (1977) listed ten examples of bilingual education aims, some of which can sound familiar to the Colombian context, such as to facilitate

communication with the broader world, to equip citizens with marketable language skills that can support employment and status and to enrich the comprehension of language and culture. Some authors may argue that in terms of language policies, the end goal of bilingualism in the country has been mainly to provide language skills that are marketable and raise status (Usma, 2009; Mackenzie, 2020).

Similarly, García (2009) differentiates between subtractive, additive, recursive and dynamic models in bilingual education. Subtractive bilingual education seeks to have students transition from their home language to the majority language. It may offer either monolingual instruction in the majority language or some form of transitional education in which students' home language is used only while the student gets proficient enough in the majority language. Additive bilingual programs, on the other hand, aim at adding a second language to the learner's repertoire and to maintain both languages. Even though this type of program aims at developing bilingualism, it still follows (as subtractive programs) a monolingual ideology as it keeps both languages apart and learners are expected to behave as two monolinguals.

The former models belong to the more traditional ways to perceive bilingual education, however, the following two have emerged as an alternative to the aforementioned monoglossic ideology of bilingualism. On one hand, recursive bilingualism emerges in contexts where "language practices of a community have been suppressed" (García, 2009, p 47). In such contexts, as the author says, it is not as simple as adding another language, bilingualism must be recursive, individuals move back and forth along the bilingual continuum as bits of the ancestral language are reconstructed for new functions. This would be a type of bilingualism to be implemented with some of the endangered languages in the country. Finally, dynamic models of bilingual education acknowledge that bilingualism is not a linear phenomenon but a dynamic one. They encompass a heteroglossic ideology and start to disregard traditional categories such as first language and second language as the world starts to interact and communicate in new dynamics. As García (2009) puts it, "bilingualism



involves a much more dynamic cycle where language practices are multiple and ever adjusting to the multilingual multimodal terrain of the communicative act". (p. 48).

### ***Language Assessment and Evaluation in Bilingual Education Settings***

The Cambridge Dictionary defines assessment as 1) "the act of judging or deciding the amount, value, quality, or importance of something, or the judgment or decision that is made" or 2) "the process of testing, and making judgment about, someone's knowledge, ability, skills, etc. or the judgment that is made". Mousavi (2009) defines assessment as "appraising or estimating the level or magnitude of some attribute of a person" (p. 36). In educational settings, assessment is a continuous process that includes an ample range of techniques (Brown, 2018). Therefore, language assessment has to do with "collecting information (in a variety of formal and informal ways) and making judgements about language learners' knowledge of a language and ability to use it" (Chapelle et al., 2020).

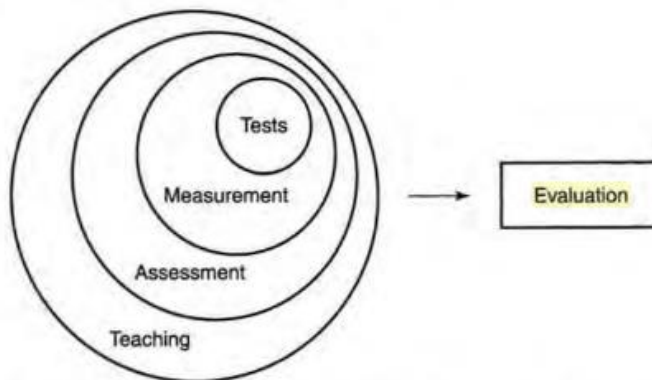
Even though testing, assessment and evaluation are terms that people may use interchangeably, many authors agree that they are in fact different (e.g., Brown, 2018). Testing is one of the techniques of assessment; that is, language assessment encompasses but is not limited to language testing. Although language assessment tends to be used to refer to assessment practices within the classroom (Chapelle & Plakans, 2012), in this study it is also conceived as transcending the classroom and entailing the overall *appraisal* of second language learners' competence and bilingualism in the educational system, high-stakes tests being one important practice to consider.

Similarly, evaluation goes beyond the act of testing or assessing individuals' language abilities and extends to the broader scope of appraising the overall quality and outcomes of language learning and other educational processes. Evaluation takes place when results are valued or judged, and the worth of performance is conveyed with reference to the consequences of such performance. (Brown & Abeywickrama, 2018). Similarly, Bachman (1990), defines evaluation as "the systematic gathering of information for the purpose of making decisions". Norris (2009) states that evaluation allows for the understanding,

improvement, and judging the way elements are working towards a target goal or outcome. Evaluation tends to go beyond student learning to judge and make decisions about educational aspects such as educational programs or curricula (Chapelle, 2009), instructional interventions, testing practices (López, 2009), language policies or teacher performance (e.g. Bailey, 2009). To clarify the confusing overarching nature of these terms and their hierarchy, figure 1 suggests a graphic representation.

**Figure 1. tests, measurement, assessment, teaching and evaluation.**

*Framework for Comparative Education Analysis*



*Note:* illustration taken from the work of Brown and Abeywickrama (2018)

Language assessment (and evaluation) can be informal or formal, formative or summative, and norm-referenced or Criterion-referenced, internal or external. Informal assessment manifests in the classroom whenever teachers promote tasks that elicit students' performance "without recording results and making fixed judgments about [students'] competence" (Brown & Abeywickrama, 2018, p. 7). On the other hand, formal assessment refers to systematic and planned measurement techniques that can offer teachers and learners an appraisal of learner's achievement or progress (Brown & Abeywickrama, 2018). Formative assessment, the author says, takes place during the learning process, with the goal of helping learners reach the goals, through the "appropriate feedback on performance", while summative assessment pretends to qualify/quantify or summarize what a learner has

grasped; it tends to take place by the end of the course or unit. Another important continuum in assessment is the one about norm-referenced vs Criterion-referenced tests. In norm-referenced tests, test-takers' results are interpreted in relation to an average score, middle score, standard deviation and/or percentile rank. The purpose is to rank learners in relation to one another (Brown & Abeywickrama, 2018). This is typical of standardized tests such as the TOEFL or IELTS, or the Saber 11 and Saber Pro. On the contrary, criterion-referenced tests are designed to provide feedback to test-takers on given course or lesson objectives, they are limited to the classroom and directly connected to the curriculum. While norm-referenced test tend to spread students throughout a continuum with most students in the middle, low performance student on an end and high-performance students on the other end (like in a bell-curve graph), criterion-referenced tests are concerned with how much of the learning objective learners have got, so, there is no problem if all students reach 100% of the objective as long as the construct and items are well designed.

Assessment and evaluation can take on both internal and external natures, constraining to the classroom or extending beyond the boundaries of the classroom. An example of an external assessment and evaluation tool is standardized tests which may both inform, to some extent, on student's language knowledge or be used to determine certification processes, admission to education and job opportunities, or accountability processes.

A number of factors come into play when it comes to language assessment, such as: the examinee's language capabilities and performance in the examination, the testing methods, the scores and decisions based on them and the assessment context. Assessments are designed based on a construct, in other words, A construct is an abstract, theoretical concept, such as knowledge or proficiency that has to be explicitly described and specified in test design (Chapelle et al., 2020). As the construct is concerned with language and language ability, the ideology or bias towards languages and bilinguals is important since it can shape the assessment design from there on. To define the construct, there are different approaches test designers can take, for example, ability-based constructs or performance-based

constructs. Ability approaches “define the construct as an unobservable trait that is not tied to any particular context of language use ... performance on a test is the result of some underlying capacities, which are also responsible for performance in non-test settings” (Chapelle et al, 2020, p.297). On the contrary, performance approaches aim at making more direct inferences from test performance to non-test performance by simulating real-life-like tasks and conditions. In that sense, grammar and vocabulary knowledge or reading comprehension (Such as the ICFES ones) are ability constructs, while speaking or written tests with simulated purposes are regarded as having performance-based constructs.

Broadfoot (1987) identified a range of purposes for educational assessment (and evaluation): for curriculum (which can provide diagnostic information), for communication (for certification and selection processes), and for accountability (demonstration of outcome achievement). Assessments could seek to measure *proficiency*, a person’s general ability, or they could seek to measure the *achievement* of a student in regard to a specific program. Chapelle et al. (2020) explain that the assessment purpose is closely related to the stakes attached to it, so it determines the assessment tool and resources invested in its development.

The purpose behind assessment and evaluation defines the type of tool to be used. Commercially designed and administered tests, for example, tend to be used to measure language proficiency, to place students in different levels of a program, or to diagnose students’ strengths and weaknesses in some linguistic categories, while classroom-based tests may be used for diagnostic purposes or achievement measurement (Brown & Abeywickrama, 2018). Tests may fall into one of these types: achievement tests, diagnostic tests, placement tests, proficiency tests and aptitude tests. The first four types of tests are present one way or another in the assessment and evaluation practices within the Colombian educational system, which is why a brief conceptualization (based on Brown & Abeywickrama, 2018) is needed:

- *Achievement tests*: These tests are normally used to determine if a course content has been learnt and skills promoted by the end of a process of

instruction. They tend to be summative and could be as short as a 5-minute quiz or long as a 3- or 4-hour final test.

- *Diagnostic tests:* These tests intend to diagnose aspects of language students need to develop or the course should include. They elicit information on what learners need to learn or work on in the future.
- *Placement tests:* These tests are used to place learners into a given level or course within a language program. Depending on the design, it can be diagnostic as well, offering some insights on student performance which can be useful for decisions on instruction.
- *Proficiency Tests:* These tests aim at measuring global competence or proficiency; it tests overall ability. There are national standardized tests such as the Saber 11/Saber 11 that consist of multiple-choice items on grammar, vocabulary and reading comprehension or many commercial tests some of which include written and spoken tasks such as the Test of English as a Foreign Language (TOEFL Test). Proficiency tests are normally summative and norm-referenced, they provide a score that tends to be used as the *gate-keeping* for the access of some form of education or other opportunity.

Assessment can take many forms besides test-related ones. Some teachers may use oral presentations, roleplays, hands-on activities, portfolios, self-assessment, pair-assessment, etc.; nonetheless, having into account that all assessment is situated in a social context (Chapelle et al. 2020) assessment choices and practices not only influences society but is influenced by society (and consequently, policy) as well; for example, by directly or indirectly having teachers teach to the test because of the high-stakes of a national test (Hopfenbeck et al., 2015; Jones et al., 1999; Smith, 1991).

A final general aspect to highlight about assessment is the principles they should follow, especially formal assessments. Assessments should be practical, reliable, authentic, valid and positive washback and impact effects should outweigh negative ones (Brown & Abeywickrama, 2018). Practicality has to do with "costs, the amount of time it takes to

construct and to administer, ease of scoring, and ease of interpreting/reporting the results" (Mousavi, 2009, p. 518). In brief, reliable tests should be consistent across applications; they have uniform rubrics, and they are unambiguous (Brown & Abeywickrama, 2018). They should be authentic in the sense that they resemble real-life-like language. Regarding validity, which is the most important criterion of any assessment tool, it has to do with "the extent to which inferences made from assessment results are appropriate, meaningful, and useful in terms of the purpose of the assessment" (Gronlund, 1998, p. 226). Some of the validity evidence an assessment tool should have are content-related validity (it contains what it is to measure), construct-validity (the tool relates to the construct that was defined), and consequential validity or impact (intended or unintended consequences of assessment). Finally, Washback has to do with the positive or negative effects of assessment on teaching and learning. Consequential validity, washback and consequences result of special interest for this study, so it is important to explore them in more depth.

**Consequential Validity in Language Testing.** Consequential validity (Messick, 1996) tends to be referred to as *impact* (Wall, 1997). The author says that *impact* is sometimes used as a synonym of washback, but she differentiates between the two concepts. For the author, impact is "any of the effects that a test may have on individuals, policies, or practices, within the classroom, the school, the educational system or society as a whole" while washback tends to be more related to the effects of test (or other forms of assessment and evaluation) in learning and teaching (Wall, 1997). Messick (1980, 1981, 1989, 1996) has extensively explored the aspects of ethics in language testing, he has called for a careful look at the (potential) consequences or side-effects of tests and has claimed that the potential impact or consequences of tests are an important matter of its validity. As the author says, tests are not a neutral practice inasmuch as values pervade their development and interpretation (Messick, 1981). Stobart also assures that "testing is never a neutral process and always has consequences" (2003, p. 140).

Tests can have an impact before and after their implementation, at the macrolevel and microlevel. Regarding the macro level, high-stakes standardized tests, for example, when used for gatekeeping, can have a negative impact on society as they “deprive[s] students of crucial opportunities to learn and acquire productive language skills,” (Choi, 2008, p. 58) besides the other learning opportunities and multidimensional benefits of accessing higher-education. These tests can also have effect before their implementation in the sense that teachers and students may focus their efforts in getting prepared to take the tests rather than focusing on the curriculum itself. In large-scale testing this can be especially problematic in the sense that not every test-taker has the same opportunity to prepare for a test as not everyone can access preparation courses or have equal overall education because of socioeconomic factors (McNamara, 2000).

A part of the consequential validity is the effect that tests have on teaching and learning (Hughes, 2003) also known as washback (or backwash). As Messick (1996) highlights, tests can either promote learning or inhibit it, then, washback can be either beneficial or negative. When it positively influences teaching and learning, offers positive feedback, is more formative than summative and allows conditions for peak performance by test-takers, it is said to have beneficial washback (Brown & Abeywickrama, 2018). On high-tests, washback refers to the effects it has on instruction and learning for example in phenomena such as *teaching to the test*.

**Measurement of Bilingualism/Assessing bilingually.** Since ideologies towards language learning have been historically monolingual, conceiving languages as separate units both cognitively and socially, assessment practices have also been monolingual, assessing or testing languages independently and frequently aiming at native-like competence in both of them (Stavans & Hoffman, 2015). This is problematic since so long as bilinguals are measured based on monolingual standards, they will “appear to be greatly disadvantaged both in linguistic and cognitive terms” (Herdina & Jessner, 2002, p. 7). Several authors (e.g., García, 2009; Grosjean, 1989; and Lopez et al., 2017) insist that

bilinguals use their different languages for different purposes, in different settings, and with different people, frequently drawing from their different semiotic resources available.

For this reason, several scholars have called for a “more holistic view of language acquisition, multilingualism, and multilingual development” (Block, 2003; Lafford, 2007; Lopez et al., 2017,). They suggest a heteroglossic approach towards assessment that allows for the presence of multilingual practices including “language choice, translanguaging, code-switching, and code-mixing” (López et al, 2017). These types of tests, of course, take different forms depending on the contexts, purposes and populations where they are to be applied, yet, in general, they are a more flexible approach to bilingual assessment that allows for bilinguals of varying language developing levels to take advantage of their different languages and other semiotic resources and focusing more actually on bilinguals’ communicative strategies, knowledge of the languages and the world.

This approach to language assessment poses a number of problems and difficulties, though (López et al, 2017). One of them is that language assessment and evaluation practices tend to be influenced by economic and political forces instead of sociolinguistic reasons (Stavans & Hoffmann, 2015). Such forces embrace monoglossic ideologies, thus, assessment designers and implementers do likewise. It could be implied, then, that a shift in language policies paradigms is needed to foster actual heteroglossic bilingual assessment practices.

### ***Exploring Language Policy on Bilingualism and Bilingual Education Through Comparative, Sociocultural and Critical Lenses***

Language policy and planning was configured as a distinct field of research in the 1960s. Language planning refers to deliberate efforts to affect the structure, function, and acquisition of languages. Then, this planning is undertaken by official institutions such as ministries of education. They become language policies in education, that being statements of goals and means for achieving them, which serve as guidelines or rules governing language structure, language usage, and language learning within educational institutions (Tollefson, 2008). Furthermore, Spolsky (2004) defines the language policy field as made of



three interrelated yet independent components. The first component is the real language practices of a speech community- the conventions on language usage within that community, also known as the ecology of speech; the second component has to do with the values assigned to the different language varieties by the members of a speech community, the ideologies behind these values. Finally, the last component is the management efforts of some people within the community (or outside) who have or believe to have power over the community and try to modify their language practices, for example by reinforcing a language variety or variant or weaken another one.

Throughout the history of language policy studies, policies have been perceived in a number of ways depending on the way policymakers and researchers view the world, language, education and policy itself, as explained by Usma (2015) and Heck (2004). Some theoretical perspectives towards policy have been rational (Metz, 2003), institutional (Meyer, 2008), comparative (Bray & Thomas, 1995), critical (Ball, 2007), or sociocultural (Levinson & Sutton, 2001; Steiner-Khamsi, 2000, 2004).

Rational views of policy and policy analysis perceive educational institutions such as schools as rational and bureaucratic organizations that rely on proficiency, abilities, benchmarks, record-keeping, protocols, mandates, employment oversight, and responsibilities applicable to all employees” (Perrow, 1986). As any other company, efficiency is a big factor and control is exercised in order to assure the most effective way to achieve goals (Perrow, 1986). This approach to exercising policy in educational institutions is reinforced by standardizing teaching and assessing procedures, formalizing curricula, and exercising top-down decision-making and accountability processes (Bidwell & Quiroz, 1991). In this classical view of policy-making and analysis, planning is conceived as politically authorized activities such as fact-finding, planning goals, strategies, and outcomes, implementation and feedback (Jernudd & Nekvapil, 2012). As Usma (2015) explains, this view of policymaking resembles a good deal of the logic followed by the Colombian

government in the formulation and implementation of current national language and education policies. In his words (p.25):

By using this approach, we may identify the main policy discourses and tools employed by the government in the last years, how these models have reinforced the use of standards and tests, and why the main focus of reform has been placed on teachers as the most determinant factor for education failure.

Inasmuch as the government logic may follow this approach, understanding policy processes from this point of view results very limiting. This is why Usma (2015) calls for a “comparative, critical and sociocultural approach for the study of policy in Colombia” as a possible solution to the limitations of the aforementioned rational view of language policy understanding.

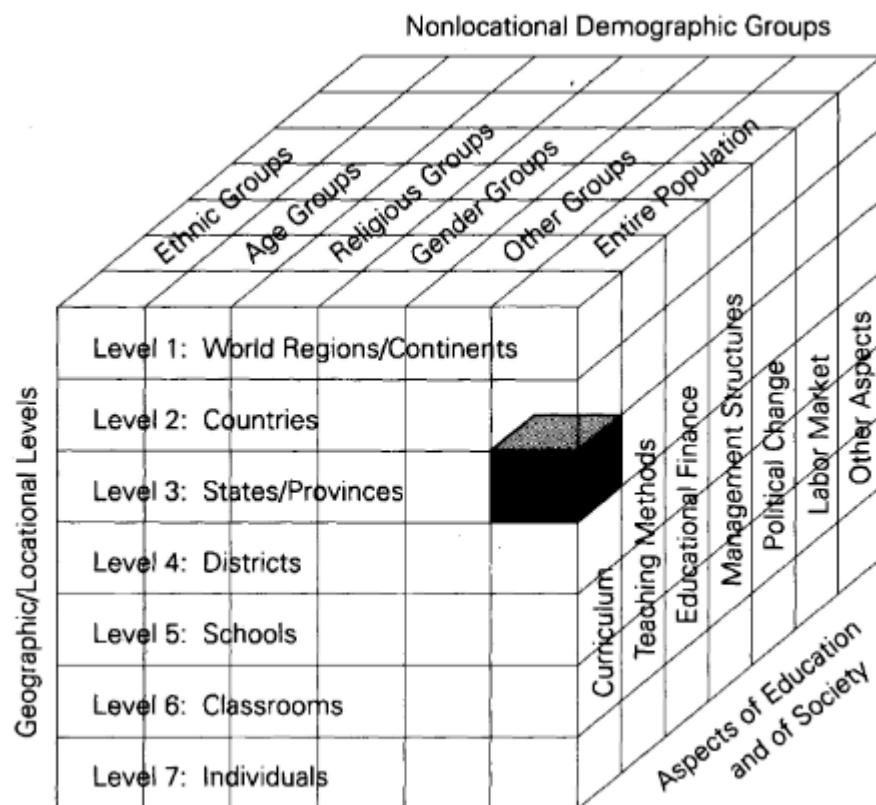
**Comparative Approach.** Bray and Thomas (1995) define comparative education as “all studies that inspect similarities and/or differences between two or more phenomena relating to the transmission of knowledge, skills, or attitudes from one person or group to another” (p. 473). The authors propose a three-dimensional way of understanding comparative studies: locational, non-locational, and educational/societal aspects of the groups. Locational aspects encompass “world regions/continents, countries, states/provinces, districts, schools, classrooms, and finally individuals” while non locational aspects include: “ethnicity, religion, age, and gender” and finally, educational/societal aspects such as “curriculum, teaching methods, finance, management structures, political change, and labor markets” (p. 473-474). This three-dimensional approach was illustrated in a diagram (see figure 1.); according to the authors, all comparative studies involve one or more aspects of each dimension and therefore they can be located in one or more cells of the diagram.

As it is evidenced in the figure and the authors make clear, comparative studies can occur at very large levels such as cross-national levels or so specific as individuals or classrooms, there are significant social and methodological considerations depending on the level of the study, however. They suggest that studies at the cross-national or national levels

could benefit from insights on more local contexts while studies on local context should bear in mind the transnational or national dynamics that may influence the phenomenon under study.

**Figure 2**

*Framework for Comparative Education Analysis*



*Note.* Illustration taken from the work of Bray & Thomas (1995).

**Critical Approach:** In research on language-policy, “critical” as a concept has three possible and connected meanings: (1) It refers to work that is critical of traditional, main-stream approaches to language policy research; (2) it includes research that is aimed at social change; and (3) it refers to research that is influenced by critical theory” (Toffelson, 2006, p.42). The first way to conceive critical in critical language policy is the questioning of the way traditional research has depoliticized policies and their analysis, focusing on

technicalities and perceived them as solvers of communication problems while ignoring social and political aspects that influence language policy (Tollefson, 2006).

As the author explains, the second meaning of critical in critical language policy has to do with a commitment to social change and social justice in research; it examines the role of language policies in economic, political or social inequality and aims at developing more socially-committed policies that improve the social conditions of people and their languages. This goes in the opposite direction of positivist “objective” approaches to research, as it takes a position, the position of the oppressed. It is concerned with the ethical questions of policy and research methodology.

Finally, the third meaning of critical refers to work influenced by Critical Theory, a field of work that explores and examines the way in which systems of inequality are created and maintained, specially, the almost invisible inequality that, through ideological mechanism, seems as “the natural condition of human social systems” (Tollefson, 2006, p. 43”).

**Sociocultural Approach.** A last way to perceive language policies in this study has to do with acknowledging the complexity of them as social and cultural phenomena, in which, even when imposed in a top-down way, several actors are involved in the policy enactment and they can actively engage in appropriation, interpretation or resisting processes from their contexts and possibilities. This approach recognizes how initially imposed policies can be appropriated and reconfigured at the local levels based on the real conditions, needs, and interests of the communities and school stakeholders (Levinson & Sutton, 2001; Steiner- Khamisi, 2000; Hart, 2002; Usma, 2015). The policy making process is a highly conflicting and contested terrain (Ozga, 2000) in which not only the policy writers and top state officials determine the language policies to be implemented, but all high-medium- and lower-level stakeholders, including education and language professionals, program developers, test designers administrators, academics/researchers, and, specially, teachers play their role in the final manifestation of the policy (Hornberger, 1996), a dynamic

unwritten unauthorized policy. It is evident, then, how policy making, and implementation is at the end a political, dynamic and unpredictable process (Honing, 2006).

Levinson and Sutton clarify that:

Policy as normative discourse may be what we call officially authorized, that is, backed by enforcement mechanisms of government or corporate charter. On the other hand, policy may also develop in more spontaneous and informal fashion, outside the agencies or offices that are constitutionally charged with making policy (Levinson & Sutton, 2001, p.770).

The latter way of making policies is sometimes referred to as unauthorized policy, a type of study that the traditional approach to language policy study does not recognize. Therefore, this study intends to adopt a broader and much more complex perspective of policies, even though the written documents are important for the analysis, for this study, language policies are all those sets of intentional and unintentional, formal and informal, explicit and implicit, evident and hidden interventions and mechanisms that arrange and manipulate language beliefs and language usage in diverse sociocultural settings and situations (Shohamy, 2006; Spolsky, 2004; Usma, 2015).

Similar to Usma's (2015) study, this project pretends to get a more profound understanding on how discourses on language policies are similar or different to the real actions and how the official-written documents may be resignified by the different stakeholders. In the author words, this comparative, critical and sociocultural approach serves to identify "gaps, strengths, contradictions, interconnections, and possibilities of the current discourses that accompany education and language reform in Colombia, and based on that understanding, illuminate future alternatives" (Usma, 2015, p. 36).

### ***Language Assessment Policy***

Taking into account that language policies may directly or indirectly influence assessment and evaluation practices, assessment tools such as tests may turn into de-facto policies, themselves, or they may influence language policy processes as well, it is important

to explore the nature of tests and how they can be used as tools of power. Similarly, the ways in which the Common European Framework have in a way become intended or unintended policy, even beyond its initial scope.

**Tests as tools of power.** Tests are a particular form of assessment that serve as power tools, depending on the decisions that may derive from the scores. This power is socially granted, however, in a global society that increasingly values standardization and accountability processes in education (Molstad et al., 2020), tests keep gaining power and the stakes of the tests keep growing. This influences even more the educational behaviors and strategies of test takers and educational systems in a struggle to succeed in such assessment practices (Shohamy, 2017). Therefore, tests are not neutral; they are socially and politically constructed, reflecting specific language ideologies, usually favoring monolingualism. Their impact extends beyond classrooms and schools, shaping teaching methods, materials, and even the curriculum. Moreover, tests can influence society at large, favoring particular languages and limiting migration processes, as well as impacting employment and higher-education prospects (Shohamy, 2001).

To illustrate the nature of tests as tools of power, Foucault (1979) explored the use of tests in his famous work “Discipline and Punish: The birth of the Prison”. He argued that tests are one of the tools through which power and control are exercised and maintained, helping to sustain existing hierarchies. According to the philosopher, tests can serve for surveillance, quantification, classification, judgment and punishment. Examination is a ritual that exists for “the deployment of force and the establishment of truth. ... it manifests the subjection of those who are perceived as objects and the objectification of those who are subjected.” (p. 184).

Hanson (1993) also questions the great power tests and testing institutions such as testing companies and government agencies have gained, having the right to define and predict people’s ability as well as maintaining power and surveillance over them. The fact that these institutions and the tests they produce are useful for the status quo has led to the lack of oversight and accountability for their designing processes and decisions, taking them

for granted, not challenging them. Similarly, McNamara (1998) argued that tests are an arm for policy reform (or policy enforcement) in education and migration policies. High-stakes tests may be useful tools to impose policies since, because of their stakes, test-takers have little choice but to strive to succeed at them.

Even though tests in general have some significant power over students and/or teachers, for example a summative final-term test that may define whether a student is promoted to the next year or course, standardized large-scale tests tend to have higher stakes to society, as they can define the future of thousands of students at once and hold schools and teachers accountable by measuring their “performance” through their students’ results. These large-scale and high-stakes tests are especially good at guiding “a wide range of education choices, including curricula, textbooks, and materials, pedagogy, teacher preparation, programming, and language medium of instruction” (Menken, 2017, p. 387). However, the author also argues that policymakers rarely pay attention to the byproducts of standardized tests, which end up being “de facto” policies that can be harmful for emergent bilinguals, their teachers and schools. On the relationship of language policies and high-stakes language tests, Menken (2017, pp. 393-394) concludes that:

Language policies are created by high-stakes testing at every level of educational systems around the world in ad hoc, uncoordinated, and often competing ways. More often than not, this is done implicitly, with the language policy implications of tests rarely being discussed openly or explained from the outset. Yet tests wield enormous power over the lives of students and educators, and shape how testing policy is exercised in schools and societies.

**CEFR as intended and unintended policy.** As it was evidenced in the literature review, education policies around the world have moved towards standardization and outcome-oriented processes (e.g., Molstad et al., 2020; Tsaia & Tsou, 2009). The CEFR is a good example of such an attempt to establish standards towards language, learning, teaching

and assessment in several European countries. This language policy has gained such popularity around the world that several countries have adopted it or at least adapted its standards in their language education system. However, for several years the CEFR has been explored and reviewed by several scholars (Byram and Parmenter, 2012; and Morrow, 2004) who have come to suggest issues both about the design, and specially the effects, either intended or unintended of the policy, both in Europe and the rest of the world.

Regarding the design of the standards, Barni and Salvati (2017) note that, even though the CEFR advocates for multilingualism throughout the document as a social good for a globalized multilingual world, when it comes to the standards, they are described from a monolingual perspective, expecting language users, to perform individually in each of their languages since “only standard language is supposed to be used, ... language is still considered a bounded system linked with a bounded community and a plurilingual repertoire is just considered as the juxtaposition of different monolingualisms” (p. 420). Additionally, authors such as McNamara (2009) question the way these standards were defined, as they were the result of a general agreement between all parties rather than the result of empirical data. This is especially problematic since, even though there are many other aspects in the CEFR, the standards are the most referenced and instrumentalized aspect of it.

Another aspect that Barni and Salvati (2012) question are the intended or unintended political uses of the CEFR and its direct link with high-stakes and gatekeeping standardized testing, most of which ground their construct validity on the aforementioned descriptors. The framework has served several institutions to establish or justify benchmarks that keep some people from getting work, education, and migration opportunities. The authors conclude that “the power of tests becomes even stronger when test criteria affect language policy and the definitions of ‘what it means to know a language’ answer generic descriptions which are far from any context and from the contextualized nature of language and language performance in multilingual scenarios” (p.422).



A final general observation of the CEFR standards to be mentioned here is that its language proficiency descriptors for the lower levels imply some previous basic knowledge and literacy, which could be problematic for groups that are largely “functionally illiterate or have low literacy skills, [similarly] the CEFR descriptors at higher levels presuppose higher levels of education” (Extra et al. 2009) which is also problematic for groups that have not accessed higher education but are expected to perform according to the “B2” or “C1” descriptors. This ignores the fact that language users that have not accessed higher education can showcase a rich and complex linguistic repertoire as well, or the fact that some second language users may not require these high general descriptors to perform in their contexts but a rather lower and specific set of skills and competences.

It is evident how the CEFR has become an operational tool serving to justify choices (Barni & Salvati, 2017; Byram & Parmenter, 2012) by policymakers, both at an educational and a social level. Additionally, Byram and Parmenter (2012) agree that the CEFR is evidently a written policy containing values and intentions. Yet, like any policy text, the intentions of its authors may not be read by its users and be taken in entirety, but only used in part for the purposes of the users.” (p.4). Finally, the different observations by scholars mentioned earlier evidence an “ideological-linguistic [bias] hiding behind the CEFR as well” (Barni & Salvati, 2017, p.424). Then, it can be concluded that awareness of the underlying ideologies of the framework, as well as the missuses in policymaking, is important for better addressing the impact of language education and taking actions in that regard if necessary.

## **Literature Review**

The literature review is presented in three sub-sections outlining important issues in the field of language assessment, evaluation and its relation to language policies. The first sub-section encompasses studies that explore the validity and impact of official assessment and evaluation practices, such as high-stakes tests. The second subsection explores language policy making and enactment processes, the role of all stakeholders, factors influencing policymaking, top-down approaches to policies and the gap between policies and social

realities. Finally, this review also includes studies that pay a closer look at the relationship between assessment and language policies and how they interact with each other; the studies suggest traits in language assessment policies towards standardization, outcome-oriented processes, measurement and accountability. A closer look at these studies is offered in the following subsections.

***Official Language Assessment and Evaluation Practices under Scrutiny: Validity, Accountability, Washback and Impact.***

Scholars have divided opinions regarding official assessment and evaluation practices such as high-stakes standardized tests. Some see them as an opportunity for the improvement of English programs (e.g., Jiménez et al, 2017), while others find them to be imposed and counterproductive (e.g. Hurie et al., 2018; López, 2009 & 2010). However, there is a general agreement that certain standardized tests do not provide a valid and complete language profile of Colombian test-takers. As an illustration, Lopez and Janssen (2010) sought to study the validity of the construct, tasks and impact of the English ECAES test, a policy-driven exam that all undergraduate students should take before finishing their studies. The data for the study was gathered through content evaluation sessions by teachers and think-aloud protocols with students. Although the researchers found some positive evidence for the validity of the test –such as relatively low student and teacher washback and teachers not teaching to the test and some discourse-based lexical and grammatical tasks– the overall English test seems to lack validity as the construct does not include important aspects of general English language ability such as listening, speaking and writing. This study found three additional problems with the test: It is not fully aligned to the CEFR as it claims to be, nor is it aligned to the contents of the intended programs in Colombia, and it does not offer authentic language, and it fails to challenge students' higher thinking skills.

The previous study is consistent with Jimenez et al. (2017) claims about the incomplete language profile of test-takers in the ECAES test and echoes a similar study by Lopez et al (2011) about the validity of the *Saber 11* test—a test that is applied to 11th graders

across the country. It also relates to other studies on standardized English Language Proficiency tests used around the world as benchmarks (e.g., Brown et al., 1997; Tsaia & Tsou, 2009) in that they tend to fail at embracing the classroom realities and learning processes, leaving a discrepancy between what is taught and what is assessed.

Lopez (2010) concludes that “1) general English language proficiency cannot be accurately judged from this test; 2) we cannot make responsible generalizations about the test takers’ English language ability beyond the testing situation; and 3) we cannot make responsible predictions about the test takers’ ability to use the English language in real-life situations” (p. 443), then, it follows that no responsible decisions in policy making can be made from those results.

Regarding the impact of the ECAES test, López (2009), through a multiple case study, studied the impact of the ECAES, a policy-driven test, on foreign language programs in Colombian Universities. The author interviewed five coordinators from five language programs to determine the impact of Colombia Bilingüe 2005, an early policy effort, at their institutions. His findings suggest that at the time little impact was perceived by the positive or negative results of their students in the ECAES test mainly because of two reasons: The language programs did not find appropriate such an incomplete test, and the universities resisted the implementation of Colombia Bilingüe, a policy that was imposed to them with no previous consulting.. This study shows how university language programs have questioned the importation of decontextualized standards and also the active role of stakeholders in the dynamics of policy implementation, as they did not passively receive the policy, but rather analyzed it, came to conclusions and reacted. The fact that the ECAES exam does not serve as gatekeeping for students but rather as an evaluation or accountability process of the higher education programs, and that the exam, as stated before, does not offer a complete and reliable description of students bilingual profile could also help to explain the reasons for the findings of López (2009).

Although assessment, evaluation and standard-oriented education have gained importance since the implementation of the *Colombia Bilingüe* program, for example with

the previously mentioned tests, there is a lack of sensitivity on what they really mean and imply. Lopez and Arendina (2009) presented the results of a study on English teacher's perception towards assessment and the way they use assessment in their classes. The researchers used a qualitative online interview to collect the perception of eighty-two teachers with different experience, training and work settings. They found that, depending mostly on the level of training teachers had, their perception would rather favor or go against language assessment. The most trained teachers talked more positively about assessment, mentioning its formative potential, while the teachers with less training perceived assessment only as a summative process, a mandate and an instrument to exercise power and control. These types of studies suggest that the assessment practices of Colombia not only at the macro level of high stake tests but also at the classroom level may lack a formative approach intended to guide students towards learning and mainly measures the *successes* or *failures* of the students (and other stakeholders such as teachers and school administrators) holding them accountable for the results.

These accountability efforts by central governments have received a number of critiques by Colombian scholars and other stakeholders (e.g. Hurie, 2018), yet other bottom-up alternatives have emerged and received better comments. As Janssen et al. (2013) affirm, accountability projects for locally-contextualized and locally-driven purposes efforts could be more empowering and productive. The previously cited authors measured the reliability and validity of a locally-driven effort to place PhD candidates in a level appropriate course for their studies. Besides wanting to understand the locally used test since the testing context had changed, the researchers wanted to gain insights on general recommendations concerning the value of locally-driven accountability projects. Researchers used descriptive statistics (both statistics of centrality and statistics of dispersion) to process the results of the piloting process. They found that each of Cronbach's alpha values (a statistical indicator of a data set's consistency) are appropriate for a high-stakes testing situation; they all were .93 or higher. Based on the research, authors suggest that not only should tests be evaluated due to external pressures but as a "continuous monitoring and development of test items serves to

ensure that local test administrators are aware of how well their instruments fit their test-taker population and are functioning the way they are intended to” (pp. 110-111). The authors also foster “self-reflected test development” as it fosters continuous action-research and education for local educators. Additionally, the authors suggest that their participation leads to broader knowledge about effective assessment practices and better understanding about “the local test items and how these intersect with locally held curriculum goals” (p.11) which could benefit the issues earlier presented by Lopez et al., (2009) about assessment illiteracy. This research is relevant as it provides an interesting alternative to top-down imposed assessment and accountability practices and it also urges for more locally-relevant tests as “[they] are, or should be, situation-specific” (Brown, 2005, p. 30).

### ***Language Policy: approaches and underpinning ideologies***

Language policy research is a field that has gained significant importance over the last years in the country and in the region. Several studies, some of which will be explored below, acknowledge the international and power-oriented nature of language policies, how international agendas affect national policies and appropriation processes by stakeholders. Such is the case of the following study. In a thorough inquiry, Usma (2015) investigated foreign language education policies adopted by the Colombian government, and the interpretation and appropriation by local stakeholders in Medellín, Colombia. The author employed a vertical case study in which he drew from globalized and historical trends in language policy; and its relation with national policy decisions. Additionally, he analyzed policy documents, semi-structured interviews, field notes and participant observations. Based on all this, the author elaborated how English is used as synonym of education quality and competitiveness, the way language reforms are adopted for the country, how the reforms are connected to transnational policy making, the role of different stakeholders at the macro and micro level, and how teachers interpret these discourses and agendas by adopting an academic or nurturing approach in their final appropriation of the policy. Additionally, the study highlighted the unpredictable nature of policymaking processes and

the difficulties of importing foreign standards, frameworks and discourses. This study represents a significant point of reference as it offers several insights on policy trends such as standardization, English-oriented policies, teacher and student testing, and adoption of foreign standards and discourses. Also, it illustrates the disconnection between national mandates and local realities, and the complex interaction of stakeholders throughout the education system in relation to the policy and each other.

The complex nature of policies and disconnection between national mandates and local realities is also illustrated in other studies from different countries. Such is the case of Chiapas, illustrated by García and Velasco (2012), in which the government has tried to acknowledge the unique educational needs of the local indigenous population through intercultural bilingual education programs. However, the article questions the appropriateness of top-down language education policies as reactions of bottom-up revolutionary dynamics, using intercultural bilingual education as a *palliative*, excluding the indigenous population from the policy building as well as the structural participation in economic and political life of the country. The value of García and Velasco (2012) relies on the fact that the authors illustrate the consequences of top-down policies that do not systematically involve other stakeholders, and specially the communities they intend to impact, and also the rather *symbolic* nature of several policies that do not achieve many of their goals because of poor design and implementation practices.

Peláez and Usma (2017) also acknowledge the paramount role of all stakeholders in the appropriation of foreign language education policies. The authors adopt the concept of policy appropriation as “moments of the policy process, when the formulated charter, temporarily reified as text, is circulated across the various institutional contexts, where it may be applied, interpreted, and/or contested by a multiplicity of local actors” (Levinson and Sutton, 2001, p. 2). Based on this definition, the authors study how varied stakeholders in a rural region of Colombia perceive foreign language policies; they also explore how their perceptions about the policy and their role in the policy design and implementation process affects the way they exercise their agency at the ground level. This study sheds light on how

rural communities experience policymaking processes. It also illustrates the active and passive role of different stakeholders in the “continuous recreation and appropriation of language education reforms” (p.121). Peláez and Usma’s work serves to provide evidence on how policies influence stakeholders and stakeholders reappropriate them and also shows hidden ideologies/perceptions of stakeholders and policies that ultimately determine the way they exercise their role in the educational system.

Despite the high importance of all-level stakeholder participation in policymaking processes, it has been evident that many of the medium and low-level stakeholders tend to be excluded from the process; even further, many of their voices are silenced when it comes to language policies. Lovón-Cueva and Quispe-Lacma (2020) critically analyze the discourses of two congresswomen aiming to evidence their hidden ideologies and how through such ideologies certain groups are excluded from the debate on language policies. The authors evidence that through their discourses, the congresswomen (specially the one with the highest position of power) reproduce power hierarchies and discriminatory social practices. This analysis is important as it shows how power relations affect the plural participation of stakeholders in policy building processes. It shows that the process of language policy is strongly influenced by policymakers’ language, education and social ideologies which they reproduce and enforce through the policy formulation and implementation; hence the importance of also exploring the ideologies behind stakeholders, policymakers and language policies.

Even while focusing on different aspects of the language policy spectrum, previous studies seem to arrive at similar postulates and conclusions. Firstly, most of these acknowledge the role and importance of the different stakeholders in the policy writing and implementation dynamics; they question top-down or imposed policy practices, and find significant disconnection between national or government-led mandates and local realities; and they illustrate how stakeholders interpret and appropriate language policies either by (partially) complying with them or resisting them (García & Velasco, 2012; Hopfenbeck et al, 2015; Peláez & Usma, 2017; Usma, 2015;). Additionally, Lovón-Cueva and Quispe-Lacma

(2020) evidenced how some policy-makers in a situation of power try to monopolize language policy discourses excluding others from the debate. Finally, not only do these studies offer relevant information from their findings, but they also evidence the importance of having different stakeholders as participants of the study and varied contexts so as to try to compensate for the great complexity of policy dynamics.

***Language Assessment, evaluation and Language Policy interplay: traits towards standardization, outcome-oriented, measurement and accountability***

Bearing in mind that tests, and other assessment and evaluation tools, are also a political act that can be used to exercise power and can have a significant impact on students and society, some researchers have studied them through the lens of policy analysis and the way assessment, evaluation and policies interact with each other and the stakeholders. To exemplify, Molstad et al. (2020) show how policies and national tests influence teaching practices in the classroom. They illustrate the historical shift in Norwegian education from content-oriented to outcome-oriented curricula and teaching in accordance with international favoring of education standardization. The researchers intend to determine how assessment and goals are constructed in between policies and practice. Their qualitative study employed comprehensive content analysis of data from key policy documents and teacher interviews to process the data. The findings indicate that “teachers are finding ways to negotiate and adjust to the language in the policies [and they] have developed their professional language according to and further beyond the policies” (p. 308). Their study also illustrates how teachers try to prioritize content but must give more importance to tests and outcomes as the final years and examinations approach. The importance of this study relies on the fact that it evidences how standard- and outcome-oriented policies and evaluation tools influence the teaching and assessment practices of teachers and how they, as stakeholders, try to appropriate the policies and negotiate with them so that their voice is also present in their teaching practices.



It is evident, as explored in the statement of the problem, the aforementioned study and some other authors mentioned before, that standardization, outcome-oriented learning and evaluation for measuring purposes have been a trait in the last decades in language assessment policies (Jiménez, Rodríguez & Rey Paba, 2017; Lopez, 2009; Lopez & Janssen, 2010; Usma, 2015). Secondly, a trait towards accountability in language assessment and evaluation policies have been identified (Copp, 2016; Hopfenbeck et al, 2015; Hurie, 2018; Janssens & Meier, 2013; Lopez, 2009; Usma, 2015;). The accountability issue can be illustrated with the following study.

Hopfenbeck, et al., (2015) conducted a insightful study about the implementation of a language assessment policy on assessment for learning and how different stakeholders in Norway experienced it. The policy aimed to support schools in adopting formative assessment practices, raising awareness of assessment as a tool for learning, and enhancing assessment skills levels and awareness. The researchers interviewed stakeholders in charge of the policy such as Ministers of Education, members of the Directorate of Education and Training, and other actors such as administrators (municipality leaders, school leaders), teachers and students. Some factors determined the successful implementation of the policy. In the contexts where there was mutual trust among the local stakeholders and where the policy was tailored to the local needs, it showed good results and positive appropriation. On the contrary, in the contexts where the policy was perceived as a controlling device, the implementation was more challenging. This study highlights the importance of accountability and trust in language assessment policy processes.

In a similar way, Copp (2016) quantified the effect large-scale tests have on teaching professionals across Canada. Two of the main factors influencing the results were the presence of high-stakes exit examinations and the pressures perceived by teachers from the testing process. In the study, teaching to the test was highly evident and strongly related to aspects controlled by the assessment policies. The author calls policy-makers and test designers for a redesign of the large-scale tests based on the findings, a better use of the results of such tests (so as not to exert an excessive pressure on teachers, students or

schools). To add up to the previous aspects, Barrance and Elwood (2018) address the issue of student voice. In a mixed-methods research that explored views and experiences of students in three countries of the UK about the evaluation and reform of GCSEs (the main examinations taken by students at age 16 in these countries), findings showed that students supported at some level the reform, they were concerned whether they would be in the student's best interest and the consequences of some of these tests. The findings evidenced an interest of students to be actively involved in decisions concerning national assessment policies. Recognizing the impact of such decisions and the perspectives of students (which tends to be ignored), the author argues for the broadening of the student voice concept so that they be also integrated to national assessment, evaluation and education decisions.

The studies of this subsection have expanded the debate on language assessment policies highlighting traits towards standardization, outcome-oriented, measurement, accountability in the different countries where they took place: Norway, Canada, UK, (and Taiwan in Tsaia & Tsou, 2009). These transnational traits in policy have not exempted Colombia. Some of the researchers in this subsection also call for a 1) reduction on the pressures by assessment policies on schools, teachers and students, as they could be counterproductive to the teaching and learning process (for example resulting on more teaching to the test rather than to the curriculum); 2) more context-relevant language policies; and the inclusion of all stakeholders and language assessment policies, including the voice of students.

This literature review shed light on the state of the art regarding language assessment and evaluation issues, language policy dynamics and approaches, and the meeting point between language policies and evaluation . The studies exploring official language assessment and evaluations as well as the ones on language assessment policies showed concern towards large-scale assessments and their capacity to address local needs or even provide for a complete assessment of students' language knowledge (especially in the case of Colombian tests). The researchers also expressed concern towards accountability and pressure issues in top-down policies and called for a more active participation of *all*

stakeholders in the formulation of these policies and tests. Likewise, studies on language policies demanded more participation of stakeholders and highlighted the complex dynamics of language policies during the implementation, interpretation, assimilation, appropriation and resistance processes.

Besides finding several insightful aspects in all the studies, some research gaps were also found. For instance, even though Lopez's (2009) study on the impact of ECAES test on teacher training programs evidenced some resistance towards the policy and little impact, it might be the case that after more than 12 years, the impact of Colombian language policies and assessment tools have a bigger impact on language training programs. Also, in regard to Colombian research, it was identified that studies mostly focus on teachers and students as stakeholders, but little attention is placed on other stakeholders such as test designers, administrators, policy-makers and academics. Finally, no studies exploring the influence of language policies on Colombian assessment and evaluation practices were found, nor an exploration of certain stakeholder's ideologies towards language, education or society in their relationship with the policies and assessments.

## **Methodology**

### **Research Approach**

In order to unveil the influence that language policies have on assessment and evaluation practices in the education system and the ideologies behind specific stakeholders and national policies, this descriptive study is framed under a qualitative and critical paradigm. The study is descriptive in nature as it attempts to describe “given state of affairs as fully and carefully as possible” (Frankel, et al., 2012). The study follows a qualitative approach since it seeks an in-depth understanding of a social phenomenon through the analysis of non-numerical data. Finally, the study also perceives the phenomenon through advocacy-lenses embracing critical theory (Frankel, et al., 2012) as an attempt to question existing unfair structures and contribute to the changing of them. In this sense, the study perceives language policy processes from a comparative (Bray & Thomas, 1995), critical (Toffelson, 2006) and sociocultural (Levinson & Sutton, 2001) point of view. The study deploys a multi-case study (Stake, 1995, 2013). A multi-case study is “a special effort to examine something having lots of cases, parts, or members” (Stake, 1995, p. vi). The parts are studied in diverse settings. As he claims, with this approach, a small group of people, activities, policies, or relationships are studied in depth. It intends to help understand how the whole operates in different situations. The study uses a multicase study as it analyzes the case of different stakeholders including policymakers, test designers or administrators, institution administrators, scholars and teachers, their perspective, ideologies and relation to the language policies, and assessment and evaluation practices. The different cases also varied in location/setting intending to achieve a broader understanding of the state of West-Center Colombian bilingualism in relation to language assessment policies.

While the qualitative nature of the study allows for in-depth exploration of language policy influence in assessment and evaluation practices, it limits the generalizability of findings to other contexts. To mitigate this limitation, a purposive sampling strategy was employed to ensure diversity in the selection of bilingual education settings, thereby

enhancing the transferability of the study's findings. Even though the study intends to accomplish an ample perspective through the different cases that were analyzed, being qualitative, it does not aim for generalization. Rather, it acknowledges the great complexity of the system with the active interaction of their members. However, the study does offer analytical insights from the theoretical interpretations of the collected data which allow for implications and possible predictions of what may happen in similar contexts.

### **Positionality Statement**

This positionality statement is included in order to present to the reader my background and interests as the researcher in regards to the topic at hand so as to illustrate any position or bias I may have on the topic; especially, since the research study has a criticality component.

I hold a B. A. in Bilingualism with an Emphasis on English Language, and, at the time of writing, I am studying an M. A. in Bilingual Education, both at Universidad Tecnológica de Pereira. Therefore, my views on bilingualism and language assessment have been nurtured by formal training in these fields. I have also worked in a binational center for three years as an English language teacher and administrative staff of the academic team. My research interests, beside bilingual education, assessment and policy, include decolonial studies and critical interculturality, biliteracy, language teaching and STEM, and language teaching with emerging technologies. As part of my intellectual identity, these views were important in the analysis of the data and the findings. In fact, these views frame the purpose of this research and are embedded throughout this research (i.e., the theoretical framework, the data collection methods, and the analysis). I am committed to the promotion and interaction of different languages under equal conditions, questioning the power and influence of elite languages (like English in the Colombian context). I also acknowledge the complex nature of policy dynamics and uses a critical-sociological perspective to consider these phenomena. Finally, I question the usage of assessment tools such as tests as power mechanisms rather than learning tools.

### **Context and Participants**

Due to the nature of this study and the intended population, the sampling purpose was to have participants that did not belong to a single context, but from various contexts and diverse positions within the language educational system in the west-center side of Colombia. However, a common characteristic that all cases have is their involvement in assessment and evaluation processes/practices and language policies being part of the language education system.

The participants in this study include nine stakeholders of different instances of the bilingual education system in main cities in West-Center Colombia: (1) policymakers, (2) test-designers, (3) scholars, (4), language institutions and language teacher training program administrators and (5) teachers. It is worth noting that some participants held two different instances, e.g. being an administrator and a scholar or test-designer and teacher at the same time. Besides, Law 1651 from 2013 (Colombia) and the *Plan Sectorial de Educación 2018-2022* (MEN, 2021) were selected as policy documents of interest. Since the purpose of this study was to understand different cases within the system, the number of participants was small, however, carefully selected. Stake (2013) explains that “even in the larger [multiple] case studies, the sample size is often much too small to warrant random selection” (p.24); therefore, as the author suggests, it is recommended to use a purposive sample of cases tailored to the needs of the study. Fraenkel et al. (2011) describes purposive sampling as a nonrandomized sampling method based on previous knowledge of the population or specific purposes of the study when the researcher believes that the specific population will be representative enough. In this sense, the participants were selected in order to assure this characteristics in the overall population of the study: 1) different levels and roles in the bilingual education system (from national-level policymakers to the classroom teachers), 2) representation of main cities of the west-center Colombia, 3) representation of the public and private sector, and 4) more than one year of experience in their field, 5) participation in policy-planning and/or enactment, 6) willingness and availability to participate in the study. A more detailed description of each case will be provided in the case descriptions section.

### Case Descriptions

Below I provide a brief description of the different context and participants considering that identities should be kept confidential.

**Int #1: PM-PB-DC.** This participant is a policymaker who a relevant project in the country aiming at promoting bilingualism and foreign languages. He holds a B.A. in foreign language teaching from a public university in Bogotá. He did his masters in TIC applied to Education in a French University. He has worked for over 15 years in the public/official sector, in leadership positions in roles related to innovation in education and bilingualism. He has worked in the national institute for the deaf and in the British Council. Finally, he has also taught in language teacher training programs. As he is in charge of a national program that aims at designing and promoting policy strategies to promote language learning and address education issues in regard to language learning, he mainly works in Bogotá D.C., but he is constantly in contact with local education secretary offices and local education institutions across the country.

**Int#2: TD/Ter-PB-RIS.** This participant is a test-designer and language teacher at the foreign languages institute of a public university. He has been a language teacher for over 13 years, and he is a highly skilled and knowledgeable teacher who is responsible for designing and administering language proficiency tests for university students. He has a deep understanding of language assessment and professional development matters as he also works as an academic advisor for the institute. In addition, he works collaboratively with other language teachers and administrators to ensure that language testing aligns with the university's goals and standards. Finally, he is a well-informed teacher about language policies in the country and he has participated in the implementation of some important local initiatives.

**Int #3: Sch/TD-PV-DC.** This participant is one of the most acknowledged scholars on language assessment in the country. During his professional career, he worked in some bilingual schools in Bucaramanga. After that, he did a Masters in TESOL and a PHD in education with an emphasis in bilingualism and assessment. After that, he worked in a

testing services company, and he started working at a private university in Bogotá. In that university, he created an assessment center. He has analyzed high-stakes tests and participated in the design of several assessment tools. He is now part of one of the biggest testing services companies in the world. His late research work has been focused on bilingual and multilingual assessment.

**Int #4: Ter-PV-RIS.** This participant is a highly-dedicated language teacher at a private school in Risaralda who guides the middle- and high-school grades. She is the leader of the foreign languages team. She has been an English teacher for over four years and has taught at that school for over two years. The teacher and the institution to which she belongs are both committed not only to promoting English in the school but to getting high results at the national tests.

**Int #5: Sch/Adm-PB-ANT.** This participant is a renowned scholar in the field of professional development and language policies in the country. He is also the administrator of the language institute at a public university. Both the participant and the institution he belongs to have held critical positions towards many language policy approaches and strategies the government has adopted, and they have analyzed them in depth. A number of languages besides English, including local languages, are promoted in the institute.

**Int # 6: Sch/Adm-PB-CAL.** This participant is also a highly renowned scholar on professional development and language assessment literacy in the country. He is also the director of a Masters in language education program at a public university in Caldas and a professor of the pre-graduate program on second language teaching. The stakeholder holds a B.E. in English teaching and a Masters in ELT with an emphasis on assessment. He has extensively trained pre-service and in-service teachers on language assessment literacy.

**Int #7: Ter-PB-RIS.** This participant is a provisional<sup>1</sup> language teacher at a public institution (*megacalorie*) on the road that connects Armenia and Pereira. She works with students coming from rural areas of Risaralda in middle- and high-school. He not only

---

<sup>1</sup> Not having an indefinite-term contract with the government



guides the English class itself, but also a dedicated hour to the national tests preparation in all the grade levels she goes to. As reported by the teacher, students come to sixth grade with a very poor English level; additionally, the lack of resources the institution faces makes it extra hard for teachers to promote second language development.

**Int #8: ADM-PV-VAC**

This participant is the director of the language institute at a private university in Valle del Cauca. Before that, she was also the director of a language institute at a smaller private university in the Colombian Coffee Axis. The institution she belongs to is typically attended by the middle- and higher-class students of the country. As stated by herself, the institution is highly committed towards internationalization processes, English being a key part in that process.

**Int #9: PM/Adm-PB-QUI**

This last stakeholder is a local policymaker from Quindío. He has also held administrative positions at a public university from that department, and he has worked as a professor in the language teacher training program. He is the communication person between the undergraduate program, the local and national government and even private institutions such as the British Council.

## **Data collection Methods**

### ***Semi-Structured Interview***

Semi-structured interviews were used to collect information from the stakeholders. Semi-structured interviews are formal verbal questionnaires designed to obtain specific information from participants (Fraenkel, 2011); even though in semi-structured interviews researchers prepare some strategic questions beforehand, the room is also open for further questions based on real-time respondents' answers. This method was selected since it has the potential to unveil ample and deep data from stakeholders to answer the research questions; additionally, since the setting and experience of participating stakeholder is so different, the adaptability of this method to the development of the interview results handy. Interview

questions (see two examples of the interview in Appendix 2) focused on the understanding of their institutions and their work, how it interplays with language assessment and language policies, and their experience, attitudes, and ideologies towards both language assessment and Colombian language policies. Interviews lasted approximately 40 minutes to one hour (Interview #1, however, lasted almost two hours). All interviews, but one, were conducted in Spanish, the first language of participants. There was one interview in English because the interviewee started from the very beginning in that language code, and I adapted to his language choice. Interviews were audio recorded to facilitate the transcription and analysis.

### ***Documents***

Two official educational language policy documents were analyzed: law 1651 from 2013 (Colombia), which specifies some mandates about bilingualism in the country, and the Plan sectorial 2018-2022 (MEN, 2021). The latter one was selected since it portrays the policy actions of the government that was elected from 2018-2022 including the ones related to language development such as those under the national bilingualism program. This document was analyzed with a content-analysis technique, contrasted and complemented with the interview to one of the policymakers from the aforementioned program. As Bowen (2009) states, the analysis of documents can point out situations that need special attention; it can help to focus the research and to make it more critical and comprehensive.

### **Data Analysis Method**

To carry out the analysis of data, both interviews and policy documents, this study uses Frankel et al. (2011) frame of content analysis since this is a “technique that enables researchers to study human behavior in an indirect way, through an analysis of their communications.” (p. 478). As the name implies, it is the analysis of the contents of a given form of communication, such as textbooks, songs, interviews, articles, etc. As the authors state, attitudes, values or ideas are often revealed through their communication. In the same line, Patton (2002) refers to content analysis as a sense-making effort through the

processing of qualitative materials to identify main consistencies and meanings. It concerns a “second-level, interpretative analysis of the underlying deeper meaning of the data” (Dörnyei, 2007). A protocol that is commonly followed by this approach to data analysis is coding for themes, looking for patterns, making interpretations and building theory (Ellis and Barkhuizen, 2005). The data analysis process is described in detail below.

### **Data Analysis Process**

The data analysis process followed several steps to ensure rigor and transparency in the analysis of the data collected. The process began with the collection of data which included the interview of selected stakeholders in center-west Colombia and the selected policy documents, which outlined the *written* view of languages, bilingualism and bilingual education and the strategies to promote bilingualism during that period of 2018-2022 by the MEN and the “*Programa Nacional de Bilingüismo*” (Plan sectorial 2018-2022, MEN, 2021). The data of the interviews were then transcribed in their original language to facilitate analysis; they were uploaded to a qualitative data analysis web-based software called Atlas.ti, and interviewee tags were generated to keep the confidentiality of participants. The tagging system works as follows:

Int #?: Role(s)-Sector-Region:

- # of interview (1-9)
- Role(s)
  - Policy Maker (PM)
  - Scholar (Sch)
  - Test Designer (TD)
  - Administrator (Adm)
  - Teacher (Ter)
- Sector
  - Public (PB)
  - Private (PV)
- Cities and Region

- Bogotá DC (DC)
- Antioquia (ANT)
- Valle del Cauca (VAC)
- Risaralda (RIS)
- Quindío (QUI)
- Caldas (CAL)

E.g., Int #3: Sch/TD-PV-DC

Once data was organized, a-priori codes were established based on the research questions to guide the analysis. In table #1, a priori codes can be evidenced. After the a-priori codes were established, the data were analyzed in the original language using Atlas.ti, and, during the analysis process, ad-hoc codes were generated to capture emergent aspects of interest that were not captured by the a-priori codes. All codes were listed in a codebook (see Appendix 3) to ensure consistency and transparency in the coding process.

**Table 1.**

*A priori Codes derived from research questions*

Question	A priori Codes
What is the influence of Colombian Language policies on the assessment and evaluation practices of different stakeholders from the bilingual educational	Political action <ul style="list-style-type: none"> <li>- Plan language policies</li> <li>- Assume a language policy</li> <li>- Participation in discussion about policies (such as challenge)</li> <li>- Reinforcement-imposition</li> </ul> - Contextualizing-translating policy <ul style="list-style-type: none"> <li>- Assessment and evaluation practices:               <ul style="list-style-type: none"> <li>- Teaching to the test</li> <li>- Alignment to high-stakes tests</li> <li>- Monolingual-oriented practice</li> </ul> </li> </ul> - Marketing influence

---

system in the west-center

side of the country?

---

How do the stakeholders of  
the English language policies  
in west-center Colombian  
education exercise their  
agency in their practices and  
influence in policy making  
and enactment?

- Stakeholders:
  - Policymaker
  - Test designer
  - Administrator
  - Scholar
  - Teacher
- Agency
- Agency restriction

---

What are the language and  
education ideologies behind  
the selected language  
assessment policies of the  
Colombian Educational  
system and stakeholders  
from the west-center side of  
the country?

- Monoglossic
  - Heteroglossic
  - Standard-oriented
  - Proficiency-based
  - Native-speakerism
-

It is important to note that the analysis was conducted in the original language of the interviews and documents (eight interviews in Spanish, one in English, and policy documents in Spanish). Only those fragments in Spanish that were selected for the presentation of the findings were translated into English using, to the extent possible, literal translations to attempt to preserve style and meaning as much as possible. Fragments in the original language are available in Appendix 4.

Finally, the data were analyzed in search of patterns and themes based on the codes identified. The data analysis process resulted in the identification of five key themes related to the influence of language policy on assessment and evaluation practices in second language education. These themes made up the findings that are presented below.

## **Findings and Discussion**

### **Complexities of Bilingualism and Bilingual Education in Colombia: Divergent Positions; Changing Paradigms; Written, Enacted, and Perceived Policy Disparities**

Before exploring the influence that language policies have in language assessment and evaluation practices, it is important to address, from the data itself, the way stakeholders and policies perceive bilingualism and bilingual education since it can offer insights about the reasons and context behind some of the following findings. In this sense, three main aspects are to be highlighted: 1) The diverging ways in which stakeholders perceive bilingualism and bilingual education; 2) the tendency towards heteroglossic ideologies that is starting to permeate the system; and 3) the disparities between written bilingual education policies, the implementation of such policies, and the way they are perceived by medium- and low-level stakeholders. Regarding the latter, it is important to highlight the little impact that most recent policies have had on the overall ethos that the government projects upon the second language education system participants.

### ***Divergent View of Bilingualism and Bilingual Education: the Case in Labeling Language Education in Colombia***

Acknowledging the social, conceptual and ideological complexities of approaching language education in the country and the several views that there could be among stakeholders, this section explored the different perspectives and beliefs that stakeholders bring to bilingual education, and how these perspectives may shape assessment and evaluation practices. As might be expected, several views on bilingualism and bilingual education were found in the data, some of them perceiving bilingualism as the equal knowledge and capabilities of both the first and the target language, others perceiving first language as an obstacle in second language learning, others perceiving bilingualism as relevant social and cultural phenomenon and a source for learning and meaning-making,

among other views. The following Extracts serve as illustrations of some of the different perspectives.

**Extract 1:** Int #6: Sch/Adm-PB-CAL

“From the classical [vision], what is expected is that the person can be equally competent in both languages, that is the classical and traditional interpretation of bilingualism, which I personally do not find dissonant, because it seems to me that it is very clear, if someone is bilingual it is because he has the competence to function ... for me a bilingual person is a person who is capable of assimilating and reacting and interacting with the world in more than one language, not necessarily English in an equivalent way.”<sup>2</sup>

**Extract 2:** Int #9: PM/Adm-PB-QUI

“Most professors [in the program I belong to] say “No, just teach the class in English”; I teach my classes in English but as soon as they (students) ask me something in French I answer. I do it across different languages at the same time because cognitively we are ready to face them.”

**Extract 3:** Int #6: Sch/Adm-PB-CAL

"I am also familiar with and believe in the issue of bilingualism as a social phenomenon of named languages in contact that lead to communication processes between people from different countries that is obviously tied to immigration issues ... So languages are living entities that are part of a society where humans interact, so learning a language is also learning that cultural part"

In extract one, the stakeholder seems to perceive bilingualism as the sum of two monolinguals, and bilingual education as the effort to make students equally competent in

---

<sup>2</sup> All fragments showcased in this section are translated from Spanish, the original language of the interviews, except for the interview Int #9: PM/Adm-PB-QUI which was conducted in English. Original language can be evidenced in appendix 4



both the home and the target language. The teacher is aware of other, more recent views, but he does not identify with them; he seems to have a strong belief in the classical balanced view of bilingualism and sees it as a desirable goal for bilingual education. In extract two, the stakeholder, a practitioner, identifies a tendency in his program, a language teacher training program, in which several professors find it unnecessary to involve student's home language in their learning process; however, he recognizes the value of allowing for the coexistence of different languages in the classroom and acknowledges the cognitive capacities bilinguals have to move from one language to another and make use of their whole linguistic repertoire. Similarly, extract three appears to show how the scholar and administrator perceives bilingualism as a phenomenon of languages in contact at the social dimension, which involves intercultural communication because of circumstances such as immigration. The interviewee also recognizes languages as always changing entities belonging and shaping larger social contexts in which people communicate and make collective meaning.

The presented data depict that there is a diversity of views and ideologies among stakeholders regarding bilingualism and bilingual education. Some stakeholders view the achievement of a roughly equal knowledge of first language and target language as the definition of bilingualism, corroborating what Cummins (1981) and Hamers and Blanc (2004) say about the common misperception of bilinguals as the sum of two monolinguals. Similarly, as García (2009) evidenced in other contexts, a number of stakeholders of the Colombian education system perceive L1, mainly Spanish in this context, as a barrier for the promotion of English language learning. Nonetheless, more inclusive views of bilingualism appear to be evident in the data, especially among scholars and some administrators, who find bilingualism as a dynamic social and psychological phenomenon of languages as living entities in contact, and the fact that bilingualism can be manifested in several ways. Some of these views also acknowledge the particularities of the linguistic repertoire of bilinguals and bilingual ability to "optimize his communication efficiency ... by calling upon the whole range of his repertoire" (Hammers & Blanc, 2004, p. 272). These findings are consistent with the idea that defining bilingualism is a complex and contested issue (Cazden & Snow, 1990).

There are many factors potentially influencing the divergent views on bilingualism and bilingual education, some of which may be: 1) bilingualism being a big and abstract concept that can be interpreted in different ways; 2) the diverse range of schools and traditions on the matter of bilingualism and second language education, 3) the conflicting social and economic interests behind bilingual education in the country and 4) the lack of an official national bilingualism policy in Colombia as manifested by some of the stakeholders themselves, such as INT #8: ADM-PV-VAC, who reported that “we do not have a public [bilingualism] policy [document] at the national level, but a program that has been implemented for several years.”

### ***Stakeholder’s Changing Paradigms on Bilingualism and Assessment***

Even though there are different views about bilingualism and bilingual education as mentioned earlier, it was also found that these views seem to be slowly shifting towards more inclusive heteroglossic ideologies of language and language education. This shift towards heteroglossic ideologies of language and language education reflects a growing awareness of the complexities of bilingualism and the need to embrace the diversity of linguistic and cultural backgrounds that learners bring to the classroom. There is a gradual move towards more inclusive bilingual education discourses that emphasize the value of multilingualism and cultural diversity. Some data samples suggesting this shift are presented in extract four, five and six:

#### **Extract 4:** Int #6: Sch/Adm-PB-CAL

“ We are also children of that paradigm where we thought that in our context student’s Spanish was an impediment to learning English. This is already changing a little and we are thinking about how to devise, how to implement pedagogical strategies in which students use their Spanish to learn English effectively ... How can I, as a teacher, make use of those genuine, legitimate and real linguistic resources that students have to assess English or help them to learn it”

The following extract emphasizes a holistic approach to English education in schools that goes beyond a single subject.

**Extract 5:** *Int #1: PM-PB-DC*

"But in order to achieve this use of the foreign language, we must understand that English is not only a competence that is worked on in a subject called English, but that it must be something more transversal where more institutional work is done in general."

**Extract 6:** *Int #1: PM-PB-DC*

*"You need spaces where you need to use the language to interact, to learn something, to fulfill a function, to fulfill a task. Well, that's what we have been trying to do with the training processes, with the materials we have produced, that English, more than an object of study, becomes an instrument to learn other things."*

In extract 1, the stakeholder recognizes that he was initially educated in a monoglossic language education paradigm that would consider student's first language as an obstacle for the learning process, but he also notices how the paradigm is starting to change and now stakeholders in Colombia are not only acknowledging the language baggage of students but also starting to consider ways to use students' first language to boost learning processes. Similarly, it can be noticed in the other extract that even the discourses of policy makers representing the government start to show more integrated, functional views of language in which they are viewed as resources that may help to access and make sense of other contents, fostering, thus, language ecologies (Fettes, 2003) that generate a need of real foreign language usage. This aligns with current trends in language education that prioritize communicative competence and real-life language use over mere memorization and grammar-based teaching.

These data suggests that scholars' work problematizing bilingualism and bilingual education and suggesting new and fairer ways to understand it are slowly starting to have an

effect in stakeholder's view of languages and language learning, such as in the view of a learner first language role in the development of the target language. Additionally, it seems that the extensive analysis and criticisms that scholars have done to previous policies such as Plan Nacional de Bilingüismo 2004-2019 (e.g. Correa and Usma, 2013; de Mejía, 2011; Guerrero, 2008) has made policymakers beware of the way they present official discourses on bilingualism and language education, or refer to inner and foreign languages, and actually start changing some of the policy strategies, aligning to more recent trends in academia. This suggests that stakeholders, mainly medium-level ones, do have some power and influence on exercise even under top-down language policy practices (Levinson & Sutton, 2001; Usma 2015).

### ***Written Policy Discourse vs Implemented Policy and Stakeholder's Perspective***

The written policies have been changing their discourses in comparison to earlier policy analysis that have been done in the country; yet, to some extent, some policy strategies or specific programs promoted by the ministry of education keep promoting a traditional view of languages. This view includes unbalanced relationships of power, promotion of only official varieties of target languages, and disregarding other forms of bilingualism or student's existing language baggage. Besides, medium- and low-level participating stakeholders keep perceiving policies as monoglossic and non-inclusive, and they tend to refer to more traditional policies such as *Guía 22* or the suggested curriculum of English. This suggests that later efforts by the government have not had such a great impact on the perception participating stakeholders have since traditional policies keep having a lot of attributed power. The following data extract #7 is taken from a written policy plan.

#### **Extract 7:** Plan Sectorial 2018-2022 Min Educación

“The Ministry is advancing in the establishment of a functional multilingualism approach that recognizes and promotes the cultural, ethnic and linguistic richness of the country with its more than 65 native languages, two Creole languages, ROM language and Colombian Sign Language, to foster the development of global

citizenship and the cultural exchange that learning foreign languages, such as English, allows”

The extract from this policy portrays the concept “functional multilingualism” moving away from the traditional use of the term bilingualism in Colombian policy which is already perceived as meaning English-Spanish bilingualism by several stakeholders (Guerrero, 2008; Mackenzie, 2020). In this case, the policy makes more emphasis on the already existing linguistic diversity of the country, and the importance of valuing and preserving linguistic diversity as a key component of Colombian identity. The focus on fostering the development of global citizenship and cultural exchange also suggests an emphasis on the social and cultural benefits of learning languages beyond just practical considerations. This extract reflects a shift in language policy discourse towards a more inclusive and diverse approach to language and bilingual education, aligning with the trends towards heteroglossic ideologies discussed earlier. Nonetheless, as likely evidenced in extract eight from the interview of the Policy Maker from Programa Nacional de Bilingüismo, in practice not everything follows the same direction.

**Extract 8:** Int #1: PM-PB-DC

"We must be more precise on the subject of bilingual education since bilingual education is a wide range. I only manage a part of that wide range, which is English as a foreign language in official educational institutions. The rest of it, what is understood as bilingual education in ethnographic terms and in terms of second languages, creole languages, sign language, is not within my reach because there are some teams and entities specifically in charge of native languages, sign language. And there is an area also here in the Ministry that is in charge of bilingual education in private schools, it is different ... we are only in charge of the National Bilingualism Program which I would call multilingualism but we are in charge, specifically of foreign languages in the official educational system."

Contrasting the extract eight with the extract seven by *Plan Sectorial Mineducación 2018-2022*, it can be noticed that, even though functional plurilingualism suggests an integrated approach towards language development, in practice, policy strategies are fragmented and the National bilingualism Program only addresses foreign languages, largely English. By limiting the focus of this policy to foreign languages, it reinforces a language hierarchy that prioritizes certain foreign languages over other local ones, ignoring the linguistic diversity of the country, or the potential they could even have in the achievement of the foreign language learning goals. This narrow view of bilingual education overlooks the value of local languages and their contributions to the country's cultural and linguistic richness, and pedagogic potential.

**Extract 9:** Int #5: Sch/Adm-PB-ANT

“One thing is bilingualism and another one is the discourse, or the concept, or vision of bilingualism that the national government, the British Council and other institutions have been proposing in recent years, right? It is different, that is just the discourse that they have been promoting. But that does not mean that bilingualism itself is something bad, no, it is something desirable and something that we should promote.”

**Extract 10:** Int #6: Sch/Adm-PB-CAL

"Now, this Colombian policy does not explore the issue of bilingualism as a personal or even social resource ... so there is no exploration of how we can say that the focus of language policy in Colombia is not bilingualism but learning English as a foreign language. Because if it were a bilingualism approach, then in that bilingualism approach, within the same policy, learning standards for Indigenous languages in Colombia of Colombian sign language and Spanish would be made explicit. " Int #6: Sch/Adm-PB-CAL

In these extracts, there appears to be a difference between the language projected by later language policies and the perception of stakeholders about such language policies. On one hand, the first stakeholder (extract 9) differentiates bilingualism as a social phenomenon and something desirable to the type of bilingualism that has historically been promoted by the government and some transnational institutions that have held a great influence in language policies in the country. In a similar line, and corroborating the argument on the lack of articulation on language development efforts in bilingual policies, the second stakeholder (Extract 10) notes how language policies do not actually promote bilingualism but rather the learning of English as a foreign language disregarding other languages and even student's first language. The stakeholder suggests that if the policy were genuinely committed to bilingualism, it would include learning standards for Indigenous languages in Colombia, Colombian Sign Language, and Spanish. This seems to reveal the unbalanced focus of the policy towards powerful foreign languages and neglect of the existing country's rich linguistic diversity. Both stakeholders hold a critical view towards Colombian language policies and their implementation, and they agree that the policy, at least in practice, is not inclusive enough and does not adequately address the linguistic and cultural diversity present in Colombia, promoting a narrowed view of bilingualism (Treffers-Daller, 2019).

The above observations seem to indicate that there could be discrepancies between recent written policies, their implementation, and the views stakeholders have about them. The data also reinforces the importance of an integrated approach to bilingual education in language policies that takes into account both local and global concerns as to make the written policy and the implemented one more congruent with each other. As De Mejía (2011) argues, a language policy for a multicultural and plurilingual nation should consider both exolingual and endolingual concerns, that is, promoting foreign languages, but also preserving and promoting local languages and cultures. However, the fact that local languages are excluded from the *Programa Nacional de Bilingüismo* suggests an instrumental notion of bilingualism, which focuses mainly on English and excludes other

local languages, disregarding the linguistic and cultural diversity of the country (Usma, 2009).

Moreover, in regard to stakeholders' position towards bilingual policies in the country, other authors have found similar reluctance among stakeholders. For instance, Pelaez and Usma (2017) found that school actors at a municipality in Antioquia perceived current English language education policies as defunded and centralized, leading to resistance within the school community, and municipal administrators also perceived language policies as an insubstantial discourse with insufficient resources and unreach objectives, leading to apathy and doubt about the policy.

This suggests that the government has failed to approach bilingual policies in such a way that stakeholders perceive them as more democratic, inclusive, theory-based and contextualized. Also, it suggests that, even though the government approach to bilingualism is starting to shift, traditional and powerful policy attempts keep having a great influence in stakeholders' collective image of what Colombian bilingual policy is. As one stakeholder manifested, language policy in Colombia means the suggested curriculum and the standard of *La Guía 22*, etc. (INT #6: SCH/ADM-PB-CAL<sup>3</sup>).

The three sub-findings previously presented suggest that there is a diverse way in which stakeholders perceive both bilingualism and bilingual education, yet it also seems to evidence how these views and discourses (even official ones) are starting to move towards more heteroglossic conceptions of these terms, evidencing the effectiveness and influence of scholars' work on the topic both internationally and nationally. Finally, it illustrates the discrepancies that there are between written discourses on bilingualism and B.E. (more heteroglossic and inclusive), the enacted policies or official strategies to address bilingualism and B.E. (monolingual) and the perception mid-level and lower-level stakeholders hold

---

<sup>3</sup> Quote in original language: “*para esa respuesta yo me voy a basar en la política lingüística que es el el Currículo Sugerido, los estándares de la Guía, 22, etcétera, etcétera, eso lo concibo yo como la política lingüística en Colombia de bilingüismo*”



towards the government's approach to these aspects, which perceive them as highly monolingual, market-driven and English-oriented.

### **Influence of Language Policies in Assessment Practices: Guiding and Limiting Practices**

English language national tests have become increasingly important in Colombian education as a result of previous language policies that have used these tests to enforce an emphasis on English in the education system. As gate-keeping and accountability tools, these tests have gained significant power and have become, in a sense, *de-facto policies* (Menken, 2018) that guide and limit teaching and assessment practices. The data collected seems to indicate that these tests have also had the potential to perpetuate inequalities, overlook other languages and cultural practices, and narrow curriculums in an effort to get better results in the tests. In the following extracts, some data illustrating the influence of language policies, and high-stakes tests as de-facto policies will be presented.

#### **Extract 11:** Int #3: Sch/TD-PV-DC

“The test is placed to reinforce the policy and what many people do is modify what they are doing in the classroom practice so that they are aligned with the test. ... So, unfortunately as I told you, evaluation is used as a tool to force everyone to react to those policies instead of working with users, teachers, students, on how to create those policies and determine which policies would be ideal, so unfortunately the practice the government resorted to, which has been criticized for many years.”

#### **Extract 12:** Int #5: Sch/Adm-PB-ANT

"In schools [the influence of the policies] is so strong that, for example, you tell teachers, come on, let's teach in a different way. Let 's integrate all four skills. ... Let's make this a little more communicative, ... and the teachers tell you: “but if that's what they're going to evaluate, what for am I going to teach them other things? and if I don't teach them what appears in the test and students get bad results for me

promoting conversations with them. If the results in the test go bad, then the person they are going to blame here at school is me."

In these extracts, both stakeholders seem to acknowledge the great influence and power that policies and tests exercise, especially at the school level. The extract eleven suggests that typically, policies are accompanied by tests that help to enforce a policy, since the stakes that they hold, make schools and teachers adapt their practices to align to the test, and therefore, the policy. It also depicts some uneasiness in the way that national testing practices are being used as tools to force compliance with policies rather than working collaboratively with teachers and other stakeholders to develop bottom-up policies that meet actual contextual needs. Similarly, the stakeholder in extract twelve highlights how gate-keeping and accountability usages of national tests force teachers and institutions to narrow teaching and assessment practices to conform with the test format and limit practices towards it instead of broadening the scope of language development. In the following extracts, both private and public teachers talk about this influence themselves.

**Extract 13:** Int #4: Ter-PV-RIS

"The school, as I told you, was number one in the city of Dosquebradas [in the Saber 11 test], and that is one of the objectives that the administrators and the principal always have. So, we, at least since I have been in the institution, have tried to start a process at least from an early age with the children so that they recognize what the format of this type of test would be like. The influence is total, that is, everything that I was mentioning to you about the bimonthly tests, although they are designed individually, we have to base our tests on the ICFES format, yes. We strive so that children from a very young age have that [test-taking] knowledge and know what to do when they have to present the test."

**Extract 14:** Int #7: Ter-PB-RIS

“At school we offer four hours of English and since it is a school focused on the national tests, students have to take an extra hour of English oriented towards Saber 11 Tests. ... Regarding the evaluation, obviously everything is process-oriented, however, due to the focus of the school on national tests, it is mandatory to apply a *Saber-like* test by the end of each period, that is, a multiple choice test-true test, all children from sixth to eleventh grade have that type of test”

These two teachers seem to agree that the national tests, specifically the *Pruebas Saber Once*<sup>4</sup>, hold such power that curriculum, teaching and assessment practices are affected and directed towards the tests. In the first extract (13), it is noticeable how the results of the Pruebas Saber appear to be one of the highest priorities of school administrators and principals in such a way that children are trained to take such tests from very early in their learning process. Similarly, in the public school, not only are final tests aligned to the national tests, but also there is a weekly hour devoted to the preparation for that national test since sixth grade (the test being administered in 11th grade). These data suggest that not only tests are gate-keeping and accountability tools, but also a factor of school prestige and a reason for school competition, since test results seem to represent an indicator of schooling quality in society. This extract further illustrates the issue at hand.

**Extract 15:** Int #6: Sch/Adm-PB-CAL

“Well, one is the effects that the tests have on the curriculum of public institutions in the country is that there are institutions that direct or focus their curriculum so that students get prepared to take a test and raise indicators in the institution; a part of those consequences is that there is more prestige for the institutions, which is a bit ironic since that prestige rises based on the results of a test that is by nature limited”

---

<sup>4</sup> National tests high-school students take by the end of their schooling process; it may determine aspects such as access to higher education.

The previous extract seems to corroborate not only the curriculum and assessment narrowing effect of national tests, but also the sense of competition that is promoted by those tests, or their misuse. Additionally, the stakeholder highlights the irony of such bases for school prestige since tests, by nature, are limited into what they can portray of students' knowledge; and, as has been argued in this project, it is especially the case for the Pruebas Saber since the skills that are tested are limited as well. Finally, the following extract illustrates how the influence of Saber 11 may go beyond classroom teaching and assessment practices to even the formulation of national policy strategies.

**Extract 16:** Int #1: PM-PB-DC

“Be the One Challenge, which I would like to show you, is very important for us because at this moment more than 400,000 students have downloaded and used it. We are at version 3.0, it is a version of an application that is gamified, and if you see the app and open it, it is basically the English section of the *Saber 11* Test. There are four levels, the student chooses which level they are in, the four levels are pre-A1, A1, A2, and B1; when the children choose the mission, within each mission there are seven sub-missions corresponding to the English section of the Saber 11 test.”

In this last extract, the stakeholder, a policy-maker working for the national government, highlights the importance of Be the One Challenge, a language policy effort to gamify assessment tools. This by itself, offers an interesting emerging approach to assessment in language policies, which will be discussed in a later finding, however, for the purposes of this section, it is important to notice that this policy effort is oriented, nonetheless, to students' training of *Pruebas Saber Once*, since all the app, according to the stakeholders, resembles the national test itself. This portrays the significant influence that the Saber 11 Test has, not only in the classroom but also in shaping government language development strategies.

The aforementioned extracts have potentially evidenced how high-stakes standardized tests, which may emerge as policy reinforcement tools, influence teaching and

assessment practices; even further, how, when they gain power thanks to their stakes and accountability misuses, they become policies that influence high-level stakeholders actions towards language development in the country. This influence is not completely surprising, as research has shown that as tests gain power and the stakes of the tests keep growing, they have a strong influence on educational behaviors and strategies of test-takers and educational systems, in a struggle to succeed in such a practices (Shohamy, 2017).

The language policy implications of tests are rarely discussed openly or explained from the outset, yet they wield enormous power over the lives of students and educators and shape how testing policy is exercised in schools and societies (Menken, 2017) and the author also highlights how the byproducts of such practices may influence teaching and learning practices, specially, in a test such as Pruebas Saber 11 which significantly guides student's English learning experience since very early in the schooling process (the beginning of mid-term education or even primary education according to statements by the interviewed teachers).

In a similar line of thought, it is important to also consider the consequential validity of these national tests, particularly, the Saber 11 test. As Messick (1989) assured, a test validity is appropriate only as long as it serves the intended purposes and produces positive outcomes. However, as the extract seventeen suggests, these tests are being misused:

**Extract 17:** INT #6: Sch/Adm-PB-CAL

“The results of the (National) tests are being used to evaluate the quality of the teachers. The tests are designed for a specific purpose ... standardized tests in Colombia should measure a part of the knowledge of students in the country, not measure teaching skills, which is not the purpose. The fact that people and institutions or people in institutions use it for this purpose, denotes lack of knowledge to understand it, but this is a serious mistake.”

The aforementioned, and the fact that these tests fail to display part of a student's knowledge in all the language skills or their bilingual abilities, calls into question the consequential validity of the test. The power and lack of consequential validity of these tests may lead to practices such as mere memorization of grammar rules and vocabulary lists, as well as to the disregard of other important aspects of language learning, such as critical thinking and problem solving, and more authentic tasks for bilingual language learners. Therefore, this leads to the overemphasis on test preparation at the expense of meaningful language learning experiences (Wall, 2013).

### **Influence of Transnational Policy Trends and Market Forces on Language Assessment Views in Colombia**

Besides the narrowing of curriculum, teaching and assessment practices, there are some other aspects that were identified as influencing assessment views and therefore practices in bilingual education. These aspects were: transnational policy trends and market forces. Three main aspects will be analyzed from the data: the focus on standard and outcome-oriented approaches to teaching and evaluation rather than processes, the role of private economic interests in shaping bilingualism and assessment discourses and practices, and the impact of the Common European Framework of Reference (CEFR) on local policy and teaching and assessment practice.

#### ***Standards and Outcome-Oriented Views***

Transnational policy trends, such as standard and outcome-oriented approaches to language education policies, as identified by Molstad et al. (2020) and Usma (2015), are resembled in the data as characterizing the discourses and practices of stakeholders. This view may result in the reduction of attention towards the process of language learning and prioritize measurable outcomes over contextualized and diverse learning experiences. Falling under these trends could explain the narrow understanding of language learning and

assessment by language policy and some evaluation practices, which may not fully address the needs and realities of actual learners and communities across the country.

**Extract 18:** Int #4: Ter-PV-RIS

“Last year, we came in first place in the Saber 11 tests in Dosquebradas, and third place at the Risaralda level, So, the school has a very good academic level.”

This extract is a sample of the standard and outcome-oriented views that are prevalent in the Colombian Education system. Firstly, it suggests a stakeholder trust in standardized tests as indicators of quality education. By affirming the school's high quality based on the Saber 11 results, the teacher seems to showcase a belief in the national test's ability to measure education quality despite the numerous criticisms surrounding the test in question and its (consequential) validity. Secondly, the extract suggests a standard and outcome-oriented view of education, in which academic/education success is measured mainly through tests and rankings. Finally, the teacher also seems to reveal a sense of competition among schools, with schools struggling for higher rankings and therefore higher status. The following extract may evidence how transnational policies hold influence in national institutions.

**Extract 19:** Int #8: ADM-PV-VAC

“What we try to do is that the curricula are updated and aligned to different international standards.”

Similarly, the next extract seems to show how language education policies have adopted transnational trends such as outcome-oriented approaches to language education.

**Extract 20:** Int #5: Sch/Adm-PB-ANT

“At the end, these data [Saber Pro results] become important when the logic is to talk about learning results. So, what is going to happen? If one does badly the next time, they will not give the university the high-quality certification. What is going to happen? That the university administration will demand that everyone pay attention

to these tests, that they prepare students for those tests, that they have to improve on those tests. Do I make myself clear? So, the test ultimately does have an effect.”

These two extracts from administrators at the higher-education level, both in public and private domains, unveil further suggestions of how standard and outcome-oriented ideologies have permeated Colombian language education, guided by the alignment to transnational policies and standards and the accountability pressures to force that. The first administrator (extract 19) emphasizes the efforts to ensure the alignment of curricula with international standards, suggesting a strong influence of transnational policies in local practices. Similarly, the second extract (20) acknowledges the emphasis on learning results in policy discourses, which end up fueling standardized test with accountability powers that drive higher institutions, sooner or later, to align with the standard and outcome-oriented discourse narratives that are present in Colombian language policies as well as several other countries that have joined this policy-trend.

These extracts illustrate how, in fact, transnational policies such as those oriented towards standardization and outcomes have permeated the discourses and ideologies of the educational system and some of the stakeholders. As Usma (2015) identified, the introduction of foreign and prepackaged models of professional development, along with standardization and adoption of international models of quality, certification, accreditation and credentials. He elaborates that the public education system (also the private one), is being influenced by a standard-based school reform model, which emphasizes control, common parameters, centralization, permanent evaluation, measurable performance, and rewards and sanctions.

Ball (2003) coined the term *performativity* as a “culture and mode of regulation that employs judgements, comparisons, and displays as means of incentive, control, attrition, and change-based on rewards and sanctions (both material and symbolic)” (p.216). This concept can be related to the outcome and standard oriented views of education, since, as the extracts suggest, stakeholders are significantly focused on standardized tests, rankings, and



curriculum alignment with international standards. As Stainer-Khamsi (2012) identifies, policy makers (and some administrators) tend to refer to international standards instead of concrete lessons from specific educational systems to justify policy, and in this case assessment practices.

### ***Market Forces Influencing Language Policies and Assessment Practices***

Another relevant factor that shapes language policies and in turn assessment and evaluation practices are the market forces or private economic interests such as those of transnational institutions involved in second language teaching and assessment. As previous authors have identified (e.g. Gómez, 2017; Usma, 2015; Valencia, 2005), these transnational institutions have gained such power and influence on policy writing and policy exercising that they have influenced policies and practices favoring specific language varieties, teaching and assessment practices, often related to profit-oriented goals rather than genuine actions to foster language development and inclusive and democratic teaching and assessment practices in the local contexts. The following extract shed light on this.

**Extract 21:** *Int #1: PM-PB-DC*

“Some educational institutions for a particular context are targeting a B2, but they are institutions that are sometimes more technical, in issues related to tourism or suddenly have a customer service technician and then they have an inter-institutional alliance in a call center and that call center supports them, finances them, provides scholarships, well, opportunities for high school students, mainly, so there are some schools nationwide that are targeting them at the B2 level.”

This extract appears to highlight the connection between transnational economic interests, specifically those of call centers, and the shaping of educational policies and practices in the stakeholder’s context. It proposes that call centers, often foreign companies seeking cheap labor, target his city schools aiming at preparing students to work in customer service right after completing their schooling process. The influence of private markets, such

as transnational call centers, may show how market-driven forces have the potential to shape policies and practices, which is problematic in as much as it suggests a prioritization of private economic interests over broader educational needs and student aspirations and it may end up in the limitation of students' opportunities for higher education and personal development as they are coopted by these companies at an early age. This goes in accordance with what authors such as Bonilla Carvajal & Tejada-Sánchez (2016) Reyes (2015) and Gómez (2017) have recorded in previous policies such as the PNI 2015-2025 (*Plan Nacional de Inglés*), focusing on the promotion of English student employability by these transnational companies rather than social development.

The extract below focuses on another aspect of Colombian market influence on language policies and, in turn, assessment practices, the hiring of transnational institutions for the building of the materials and teaching and assessment resources:

**Extract 22:** *Int #1: PM-PB-DC*

We wanted to generate another National Bank of evaluations, which you can find on the eco web, where all the competences are evaluated, all the communication skills of the students. There are entrance and exit tests, they are for levels A1 A2 and B1. The teachers can download it, and in fact, since it is one of the most downloaded resources nationwide. Because we wanted to also have standardized tests, we developed these with an app through the British Council. They are tests that are calibrated to be able to check the level of the students, according to the four basic skills of communication.

**Extract 23:** *Int #3: Sch/TD-PV-DC*

“The first thing the Ministry of National Education did was pay the British Council a lot of money to evaluate the teachers, when everyone without even taking the exam knew the answer: the teachers do not have the level. And an investment was made in purchasing and administering the exams, it was done and "uh, surprise, everyone has a low level" instead of investing in other things..”

These two extracts also provide insights into the influence of market, transnational institutions such as the British Council, on Colombian language policy making (besides showing further indications on standardization and outcome-oriented views). In the extract twenty-two, the policymaker mentions the collaboration (one of many) of the Colombian government with the British Council to develop a national bank of standardized English tests. This partnership illustrates a way in which assessment and evaluation practices are shaped in the country, most probably inclining them towards language, education and assessment ideologies that are aligned with the British Council and similar institutions. The second scholar (extract 23) seems to raise concerns about the granting of resources to influential institutions like the British Council for pointless practices, such as investing lots of resources on standardized tests for getting an expected outcome. He argues that resources could have had a better end. Both Extracts illustrate the influences of market, often transnational market interests, on language policy and ultimately teaching and assessment practices, favoring private interests over contextually-based investments addressing actual local needs. The following extract may show how some stakeholders are not aware of such interests.

**Extract 24:** Int #5: Sch/Adm-PB-ANT

“We understand their partialized views (of bilingualism), highly influenced by economic agendas and colonial policies that we do not identify with”.

Furthermore, this stakeholder claims that official bilingualism views, and therefore policies on that regard, are highly influenced or biased by economic private economic agendas and colonial policies that try to sustain the existing power and ideological structures. A way in which these policies the stakeholder mentions are manifested is regarding assessment practices, especially high-stakes tests such as international certification tests, as the analysis of the following extract can shed light on.

**Extract 25:** Int #2: TD/Ter-PB-RIS

“The certifications have a lot to do with adjusting to a series of international protocols and well, inviting those international institutions to review, but well, that invitation represents very high costs. And the public university offering this to a massive number of students does not have those resources, that is, there must be a great political will to do it because certification is expensive and, in any field, an international certification is much more expensive.”

This final extract illustrates the influence and control that transnational institutions, typically based in Europe and the US, may have on English proficiency certification. It suggests that international certifications are the main indicators of quality and excellence in language education, which has resulted in Colombian teachers and institutions spending substantial amounts of money on obtaining these certifications. This situation seems to evidence a power dynamic that exists between local educational institutions and foreign transnational testing centers and institutions, which maintain a virtual monopoly on the certification of people’s English proficiency. Besides, the socially-constructed need for international certification sustains the notion that only US and European institutions have the academic authority to assess language proficiency. This collective view may hinder efforts to develop and implement locally tailored proficiency certifications that are more responsive to the unique needs and challenges faced by Colombian students and teachers, such as the affordability factor.

The phenomenon of "businessification," as described by Le Gal (2019), refers to the global trend of education becoming a for-profit activity adopted by corporations. In the Colombian context, this trend is evident in the proliferation of private institutions and ad hoc agents that determine policies, administer tests, and profit from selling various educational products (Gómez, 2017; Usma, 2015). This market-driven approach to education raises concerns about the prioritization of profit over the best interests of students, teachers and society.

Entrusting overseas companies with national tests, as noted by Le Gal (2019), exemplifies the importation of educational technology and testing tools from transnational institutions. This reliance on foreign models may lead to the adoption of standardized, outcome-oriented approaches that do not fully account for the unique needs and challenges of the local context. Moreover, the dominance of transnational institutions in determining language proficiency further reinforces their control over educational policies and practices, while undermining the authority of local educators and institutions.

As Valencia (2005) suggests, it is crucial to critically analyze the discourse on investment in English and bilingual education to determine who truly benefits from the promotion of "bilingualism." This critical approach should aim to unpack the power dynamics and economic interests at play, as well as assess the potential consequences of market-driven policies on educational quality and equity.

### ***CEFR Power and Influence on Language Policies and Stakeholder's Discourses***

Another relevant aspect to analyze regarding transnational forces is the powerful influence that the Common European Framework of Reference (CEFR) has exercised on language education and assessment practices in Colombia. The CEFR has become a dominant force shaping language learning and assessment discourses and practices, which has led to the shadowing of the need for local adaptations and context-relevant approaches. Ironically, however, data (e.g. Int #1: PM-PB-DC, Int #9: PM/Adm-PB-QUI, Int #5: Sch/Adm-PB-ANT) seems to suggest that, even though the CEFR is frequently present in discourses, when it comes to the practices, it is often a superficial part of the CEFR that gets to be implemented. The proficiency labels of the framework seem to be the main focus of implementations so as to show compliance with the policy rather than a thorough designing process or courses and assessment tools. To illustrate the powerful and influential presence that the CEFR seems to hold in official discourses and language policies, the following extracts are useful.

**Extract 26:** *Int #1: PM-PB-DC*

*“Well, basically the Common European Framework is the basis of the basic standards of English Competence and the basic standards are the basis of the Suggested English Curriculum and therefore it is the origin of many things.”*

**Extract 27:** Int #9: PM/Adm-PB-QUI

*“And the amount of time that I need to take them from A1 to A2 or A2 to B1 or vice versa or going forward or backward.”*

**Extract 28:** Int #5: Sch/Adm-PB-ANT

*“the problem is wanting to base everything on the whole [Common European] framework, that is, everything has to be the Framework that is, you can no longer think of a course without thinking about the framework. You can no longer think of an exam without thinking about the framework, or a textbook without thinking about the frame.”*

The three extracts denote the considerable impact and authority of the CEFR in language policy and stakeholder’s perspectives and practices in Colombia. As stated by the policymaker from Programa Nacional de Bilingüismo in extract twenty-six, the CEFR is the foundation for many of the main policy documents, guidelines and curriculums. It has permeated the system in such a way that stakeholders typically refer to learner’s English or language development processes in terms of the CEFR labels (A1, A2, B1, B2, C1...), as shown in extract twenty-seven. Moreover, the third extract (28) illustrates the extent to which the CEFR has become default or de-facto policy when it comes to language education and assessment in the country, in such a way that attempting to propose alternatives or even develop context-relevant approaches outside that framework has become very unlikely.

As mentioned in the theoretical framework, Berni and Salvati (2012) have raised concerns about the political uses of the CEFR, which have been used, for example, to justify high-stakes standardized testing or pervasive policies. The authors explain that the

framework can give more power to tests (and policies) since they offer a socially-trusted narrative of what it means to know a language, to such a point that, as the data seems to indicate, people refer to the CEFR labels regularly. Byram and Parmenter (2012) also acknowledge the CEFR as a written policy bearing values and intentions. Nonetheless, authors cautioned that the intentions of the authors might not always be fully adopted by its users, leading to partial (and instrumental) implementation for specific purposes, which in a sense, is what has happened in the country, where the framework has become an omnipresent force shaping policy and practices and transnational institutions have taken advantage to monopolize the market by claiming their alignment to that framework.

In light of the three subfindings earlier discussed, it becomes apparent that second language education in Colombia has been significantly shaped by international standards, frameworks, and private, often transnational, economic interests. As Usma (2017) highlights, language teaching and learning in the country have become a matter of meeting a standard and advertising oneself as a product in the "free" market. Historically, the subordinated position of Colombia to more powerful countries has facilitated the proliferation of these trends, with education increasingly becoming a lucrative business rather than a space for fostering meaningful learning experiences. However, it is crucial to advocate for a more sovereign approach to second language education and bilingualism, in which local research and developments by qualified professionals are valued and favored.

Granados-Beltran (2016) advocates for designing situated pedagogical and assessment practices that consider local historical, social, political, and educational conditions. By emphasizing the value of local knowledge and production, the education community can be empowered towards a more diverse and contextually relevant understanding of bilingualism. Similarly, Clavijo (2009) agrees on the importance of becoming more intellectually independent as a nation, distancing itself from imposed dominant ideologies and supporting local knowledge and production, which, among other things, require to be more critical about the policies we borrow, favor students and society over private interests and devise our own frameworks or at least carefully adapt them.

## **Emerging Practices on Language Assessment and Language Assessment Policies**

This finding highlights an emerging shift in the language assessment policies and practices in the country. Despite the predominance of traditional-monolingual practices heavily driven and influenced by high-stakes standardized tests, some emerging trends in language assessment policies and practices that offer alternative approaches to language assessment were identified: the promotion of formative and alternative assessment practices by language policies, emergence of heteroglossic bilingual assessment practices, and the increasing integration of technology in assessment.

### ***Relevance of Formative and Alternative Assessment on Language Policies and Stakeholders' Discourses and Practices.***

The data suggests that recent policies and policymakers have begun to acknowledge formative and alternative assessment practices, recognizing the need for diversified assessment procedures, aligning with views of some stakeholders who make a call for multiplicity practices in language assessment and report doing so in their assessment practices. This new approach starts to move policies and practices away from the reliance on high-stakes testing, focusing, instead, on broader and diversified assessment procedures that can better capture a comprehensive picture of students' actual language learning and development process. The following extracts serve as illustration:

**Extract 29:** INT #1: PM-PB-DC

“One of the main aspects that we have tried to do from the Ministry is, one, to demystify the concept of evaluation that currently exists in the educational system. Assessment is not just an exam. And so, we have been generating some tools and strategies to be able to make the assessment more of a process, and to make it part of the student 's learning process.”



**Extract 30:** Plan Sectorial Mineducación 2018-2022

“In the same way, tools are put into consideration so that teachers have instruments that strengthen their diagnosis and monitoring strategies, and so that students are active participants in these processes.”<sup>5</sup>

The data might evidence that policymakers and written policies are shifting their approach and are starting to promote other tools than exams for diagnosis and monitoring which strengthen and democratize their assessment practices, where students have more active roles in their learning process. The first extract (29) suggests a desire to *demystify* the concept of assessment as solely meaning tests and a commitment to develop tools and strategies that emphasize assessment as an ongoing process. By engaging in tools and strategies developments that promote alternative assessment, they seem to aim at integrating assessment more closely with the student’s learning process. Extract thirty, by the written policy, seems to emphasize the importance of providing students with instruments that strengthen their diagnosis and monitoring strategies and involving students in the active participation of such processes. The data suggests an effort to relieve pressures on high-stakes test preparation in classroom time by offering parallel test-taking training through the technological tools, so that classroom time and public resources can be invested in more sustainable learning projects. They also acknowledge a bias in the educational system towards test-based assessment practices and claim to be tackling that aspect by offering alternative tools and training sessions on language assessment literacy.

The following extract suggests that stakeholders may be starting to diversify their assessment practices.

**Extract 31:** INT#6: SCH/ADM-PB-CAL

“So, what the tests can do, at any level, massive or in class, is limited in terms of the data they collect. So, let's say that, at a more educational level, more than in the

---

<sup>5</sup> Translated from Spanish. See original version in Appendix 4.

classroom, teachers are called to collect data on the communicative competence of our students from different points of view. ... So I can do a formal test where I collect a bit of information, I use a portfolio where I collect a bit of information and I do a self-assessment where my students give me a bit of information. That is what is called the perspective of multiplism.”

**Extract 32:** Int #9: PM/Adm-PB-QUI

“I will say that most of the time formative [assessment] is the deal. Summative is really important but formative assessment should be at all times.”

These two extracts may reveal an increasing awareness among stakeholders of the limitations of tests, both standardized and classroom-based, in capturing the full extent of student’s language development; therefore, they recognize the need for diversified assessment practices that draw on multiple perspectives and strategies to provide a more holistic understanding of students’ communicative competence. The first extract (31), in particular, highlights the importance of multiplism in (bilingual) language assessment, which can represent a fairer and well-rounded way to collect data on a student’s language development and make more informed instructional decisions. The second stakeholder (Extract 32) emphasizes on the criticality of formative assessment in education to support students’ learning process rather than focusing on mere achievement.

These data seem to indicate that besides language policy and stakeholders’ discourses adopting later views on bilingualism and bilingual education (see finding 1), they have also aligned to scholar’s discourses on assessment practices such as that of prioritizing alternative assessment (e.g., Giraldo, 2018; Muñoz et al., 2012). Nonetheless, there might be a discrepancy between language policy’s discourses for alternative assessment and practices in which such efforts are directed towards high-stakes test preparation. Additionally, these later discourses for alternative assessment seem to be shadowed by traditional policies or high-stakes tests as de-facto (Menken, 2018) policies such as the *Saber Pro* which drive classroom practices away from formative assessment practices, and towards teaching to the test practices (Koretz, 2017).

***Emerging Heteroglossic Bilingual Assessment Practices***

Another emerging shift that was identified, despite the great presence of monoglossic practices still, is the exploration of heteroglossic bilingual assessment practices that give more importance to students' first language (L1). This shift acknowledges the crucial role that students' complete language repertoire has on their language and content learning process and the assessment possibilities that it represents, such as getting a better appraisal of student's actual content knowledge or boosting students' performance in the target language. The following extracts illustrate this shift in stakeholder's view.

**Extract 33:** INT #3: SCH/TD-PV-DC

“As I told you, we give a lot of importance to the foreign language when we talk about bilingualism and we cannot forget about the mother tongue, especially assessing it. From this process of language use, many of these skills or competencies can be transferred to the second language.”

**Extract 34:** INT: #6 SCH/ADM-PB-CAL

“But let's say that the big trend, which will continue to exist, is the use of task-based assessment to activate students' language competence. In other words, it is not evaluating bilingualism, but what can be done with bilingualism ... If the focus is not philosophy, but English, as a foreign language, how can I create, think about pedagogical strategies where I can evaluate in English as a foreign language activating the bilingualism of my students? That is, if I want to evaluate an aspect of my students' listening comprehension in English, how can I design an evaluation instrument that takes, say, the Spanish of my students so that I can look at and evaluate their listening comprehension in English? So let 's say bilingualism as a resource in the area of learning English as a foreign language.”

These extracts seem to show a growing awareness among stakeholders of the importance of heteroglossic bilingual assessment practices in Colombia (Arias, 2016), which acknowledges the value of both the student's first language and target language during assessment processes. In the first extract (33), the stakeholder identifies that there is a tendency to prioritize the target language in assessment, but he seems to reveal awareness of the importance of considering a student's mother tongue in such processes since, for example, some skills in the first language can be transferred to the second one. In a similar line, the second stakeholder (extract 34) elaborates on this idea by discussing the potential of task-based assessments to assess bilingualism or to assess a target language bilingually, that is, strategically using student's L1. He insists on the relevance of incorporating students' L1 to support their understanding and improve their performance in the target language. Therefore, stakeholders view bilingualism as a valuable resource in language learning and assessment processes.

The identified emerging shift towards heteroglossic bilingual assessment in the country is consistent with the multilingual turn in applied linguistics that has been coined by Conteh and Meier (2014). This wider shift in language education ideologies in academia may be influencing the perspectives and approaches of stakeholders in Colombia as well. This multilingualism calls for a "more holistic view of language acquisition, multilingualism and multilingual development" (Block, 2003; Lafford, 2007). Even though in academia this shift may have started several years before, when it comes to assessment, bilingualism is a rather recent construct, as was noted by one of the interviewed stakeholders: the big discussions about bilingualism as a construct in foreign language assessment started about 2017 or 2019, so it is relatively new"<sup>6</sup> Int #6: Sch/Adm-PB-CAL. This bilingual turn recognizes the value of incorporating multiple languages in the learning and assessment process, and stakeholder's growing awareness of the importance of bilingualism as a resource in assessment aligns with such a view.

---

<sup>6</sup> **Quote in original language:** "*digamos que las grandes discusiones en cuanto al constructo bilingüismo en evaluación de lenguas extranjeras vienen más o menos desde el 2017, 2019 y entonces esto es relativamente reciente*"

The heteroglossic approach to assessment, as suggested by López et al. (2017) allows for the presence of multilingual practices, including language choice, translanguaging, code-switching, and code-mixing. It contrasts with mainstream assessment practices in which students are expected to perform monolingually in tests even though they might give better responses if it was possible for them to answer partly in their stronger language (Shohamy, 2011; Shohamy et al, 2015). So, by adopting a more heteroglossic bilingual assessment approach, Colombian stakeholders may be creating more inclusive and effective assessment practices that take into account students' diverse and rich existing linguistic resources in learning and assessment processes.

### ***Increasing Tech Integration to Language Assessment Practices***

An increasing integration of technology, and identified forthcoming integrations, in language assessment to tackle current problems and barriers was identified. This indicates opportunities for better assessment of linguistic skills and assessment practices, adaptation to learners' needs and goals, and possibilities to provide more holistic appraisal of student's language development process. It also could allow for reduction of the stress that may be associated with traditional assessment procedures. Firstly, the following extracts identify language assessment needs in regards.

**Extract 35:** INT#2: TD-TER-PB-RIS

“I would think that if the national government really wanted to have [...] a true political will, they would have a test that combined the four skills and that, perhaps using technology. [...] I know it may sound crazy, but I don't think it will be soon, [to use] Artificial Intelligence to evaluate, at least within a range, the speaking competence and the writing competence and the listening one”

**Extract 36:** INT #3: SCH/TD-PV-DC

“You can use more simulations, for example, in which you are interacting with an avatar. ... So in some way it allows you to evaluate in a more appropriate way all those

linguistic skills than before with pencil and paper. ... Today the reality is what they call multimodality in English, which means that you not only read different written media on the Internet, but also videos, erratic table views, and so on.”

The previous extracts could evidence that stakeholders have identified language assessment challenges that could be addressed through some technological integrations, which they identify as a not-so-far possibility. For example, the first stakeholder (extract 35) suggests the use of technologies such as artificial intelligence which could play a role in the assessment of productive skills like speaking and writing, as well as listening competence, on a large scale. The second stakeholder (extract 36) expands on this idea as he proposes the use of simulations and interaction with avatars to create a more immersive and dynamic assessment environment. Both stakeholders identify a potential of such integration in the improvement of assessment practices; however, they also acknowledge the challenges and resource demands associated with implementing these technology integrations, as well as the need for actual political will to make such advancements a reality. Despite these difficulties, stakeholders emphasize the importance of moving in the direction of technology-enhanced language assessment to improve the overall effectiveness and relevance of assessment practices in Colombia and respond to current challenges.

Even though some of these identified necessities are not yet identified in the data, the following extracts may evidence how some integrations are already happening, especially in regard to gamification, and analytics that may represent some benefits both for students and other stakeholders' actions.

**Extract 37:** Plan sectorial Mineducación 2018-2022

“Be The One Challenge: digital application that seeks to strengthen learning in English, close the gaps and generate a culture of continuous improvement in learning English.”

**Extract 38:** INT #1: PM-PB-DC

“There is still a very long way to go. I think that with this little game (be the one challenge), as many teachers call it, we have been somewhat demystifying the way assessment is perceived; it does not necessarily have to be a stressful process where the student has to sit in front of an exam, involving the whole issue of pressure and socio-emotional factors to demonstrate the results. Assessment can also be done through this type of game-based tool which provides some data so that the teacher is constantly monitoring what is happening with student learning.”

**Extract 39:** Plan Sectorial Mineducación 2018-2022

“In the same way, tools are put into consideration so that teachers have instruments that strengthen their diagnosis and monitoring strategies, and so that students are active participants in these processes. To such ends, the strategy “*Supérate con el Saber*” (Succeed with the *Saber* Test) plays an important role. Students from official and unofficial educational establishments enrolled in grades two to eleven participate in this strategy, through an online or offline application. ... *Supérate con el Saber* allows certified territorial entities to monitor the results and progress of their institutions and redirect the resources that are currently invested in training for the State Tests, particularly *Saber 11*, towards actions that favor sustainable learning rather than orienting efforts towards test-taking training.”

These extracts from written policies and a policymaker highlight the ongoing efforts to integrate technology in language development and assessment practices in Colombia through initiatives such as *Be the One Challenge* and *Supérate con el Saber*. Extract thirty-seven introduces *Be the One Challenge* as a digital application aimed at strengthening English language learning, closing gaps in the education system and fostering a culture of continuous improvement. As it was earlier mentioned by the stakeholder, it resembles the English *Saber 11* test and prepares students to take it. Similarly, Extract 39 discusses further the role of *Supérate con el Saber*, an online or offline application that involves students from

various grade levels and allows certified entities to monitor the progress of their institution to strategically focalize improvement efforts.

Extract thirty-eight emphasizes on the potential benefits of using game-based tools like be the One Challenge in assessment since they can help to demystify the assessment process and reduce the stress and negative emotions that students may associate with traditional exams. Additionally, these tools may provide data to stakeholders on student's language development without the need of high-pressure high-stakes measures.

These data are relevant since authors and researchers have also identified the benefits and the potential of integrating technology in assessment practices. For instance, (Llosa, 2018) explains how technology can be used to design content and language integrated tasks, a form of bilingual assessment, allowing for the dynamic integration and separation of both constructs. Similarly, (Fox, 2018) illustrates how technology advances have allowed for the increase of alternative assessment practices that allow for the delivery of achievement, competence and ability evidence. Furthermore, he foresees that "technological advancement will continue to extend alternative assessment approaches." (p. 144), something that Kuhail et al. (2022) corroborated, for example, in the field of natural language processing powered by artificial intelligence.

Similarly, (Chappelle et al., 2018) identified the growing use of artificial intelligence powered language processing tools to allow for the assessment of productive skills. This growing field in language assessment, the author claims, may offer an alternative to the over-reliance of *selected-response items* and offer more linguistically rich tasks to test-takers. Finally, technology represents a great opportunity for incurring and exploring heteroglossic assessment designs. For instance, Lopez et al., (2014) proposed a technology-enhanced platform that permits bilingual speakers to make use of multiple assessment features such as listen to, write, or record answers in multiple languages to make room for the strategic use of whichever linguistic resources students may have at their disposal.

Even though these integration efforts may be somewhat misspent since they are heavily focused on high-stakes tests training, as was presented in the second finding, they do



suggest a possibility for local stakeholders to collect data on students language development in an alternative way and focus resources and efforts on language learning instead of high-stakes test preparation. It should be noted, however, that the integration of technologies such as analytics may represent new challenges to assessment practices when misused since they could be a possibility to strengthen accountability measures with an ubiquitous assessment that always watches over local stakeholder's practices. For instance, Fox (2018) acknowledges the danger of technology-mediated alternative assessment practices such as digital portfolios to increase surveillance and control since they sometimes represent a vast repository of written compositions that may be misused by those holding such big data.

The data earlier presented suggest emerging language assessment policies and practices such as the embracement of alternative assessment practices, contemplation and exploration heteroglossic assessment practices, and integration of emerging technologies in assessment practices. This represents a potential shift in language assessment and language policy in the country, which may challenge the traditional emphasis on high-stakes tests and international standards and monolingual practices.

### **Stakeholders' Active Engagement in Policy Enactment**

The interaction with the different stakeholders seems to reveal that, despite the constraints imposed by certain traditional, powerful and de facto policies on stakeholders' agency, it has been observed that they have adopted various strategies to participate in policy discussions, exercise autonomy in their practices and actively engage in the policy enactment process. By questioning, debating and resisting some policies, stakeholders strive for more contextually-relevant policies that cater specific needs of their contexts. Additionally, recent policies were found to be more participatory, placing greater emphasis on teacher's voice and stakeholder articulation, yet, similar to later official views on bilingualism, they have been shadowed by traditional more powerful policies. The data highlights the importance of stakeholder involvement in shaping and enacting language policies in local contexts.

**Extract 40:** *Int #5: Sch/Adm-PB-ANT*

*Issues such as evaluation, which is already very normalized, that is, there are rules for everything. Which tests are accepted? Which ones are not? What are the required levels? Why those tests? Where do these tests have to be presented? What are the scores, the relationship of the scores to the [Common European] Framework? all that is already normalized. ... In fact we receive a lot of criticism because sometimes they say: "but if you are the University of Antioquia, you have University autonomy. Why do you require, for example, a teacher to have a C1 [certificate] to be a teacher trainer here in the bachelor 's degree?" Man, because that is what the government demands.*

**Extract 41:** *Int #3: Sch/TD-PV-DC*

“So, it was a way of forcing the schools to react, like, "here's a test and a language policy, you'll see what they do.”

These extracts appear to highlight the ways in which policies can hinder stakeholder's agency, compelling them to align with practices they do not identify with. The first extract (40) could reveal the highly normalized and regulated nature of assessment and evaluation in the country, such as policies regarding the valid international certifications for both teachers and students in order to access job and education opportunities, it suggests that assessment and evaluation decisions and practices beyond the classroom are highly constrained and little room for autonomous practices is available. The second extract (41), similarly, illustrates how high-stakes tests, as de facto policies, can further constrain stakeholders' agency since stakeholders must adapt to these tests or face potential consequences, limiting their possibilities to make choices based on their own view of fair and appropriate practices. Despite all this, as the following extracts may illustrate, stakeholders keep seeking to actively engage and participate in policy discussions and policy enactment opting for more contextually-relevant language policies.

**Extract 42:** *Int #2: TD/Ter-PB-RIS*

“Both [directors] have been very committed to the Institute being an active player in regional bilingualism strategies. So, the Institute does participate in the regional bilingualism roundtables and also participates ... in the discussions that take place at the University, in the municipality and has always tried to be immersed and aligned with those policies. It has participated in several bilingualism projects, including those of Colombia Bilingüe. For example, I have been part of several of those projects, both in the formulation and in the execution.”

**Extract 43:** INT #8: ADM-PV-VAC

“We created an institutional policy a policy that helps us see where we are going in terms of foreign language.”

**Extract 44:** Int #5: Sch/Adm-PB-ANT

“If there is something that characterizes a space like this (the university), it is diversity; some professors do one thing and others do another, that is, there is no uniformity here, but that is simply part of the richness of the institutions. As dictatorial as a principal may be, it is very difficult for him to absolutely control the work of teachers. Yes, or at least in a country like Colombia it does not happen, maybe it's possible in other countries, yes, but in a country like Colombia it's very difficult.”

The previous extracts illustrate the different ways in which stakeholders have managed to actively participate in language policy planning or specially enactment and exercise their agency in so-doing, despite the constraints imposed by certain (de-facto) policies. By engaging in discussions, joining bilingualism roundtables, or contributing to regional language policy strategies, stakeholders have been able to influence and shape, to some extent, the policies that attempt to manage their work. Extract forty-two, for instance, highlights the commitment of directors to ensure their institution is an active agent in

regional bilingualism policy making and enactment, participating in various committees and projects related to bilingualism and foreign language development.

Extract forty-three, similarly, may evidence that stakeholders have also reflected upon their immediate context and attempted to develop institutional language policies that target the language needs of their populations and surroundings. This has allowed them to tailor wider policies and direct them towards their specific ends. The final Extract (44) acknowledges the limitations of restrictive policy practices since it can be difficult to completely control stakeholder's actions, particularly within diverse educational spaces such as universities. The stakeholder values the richness of this diversity as positively contributing to democracy and society building. This highlights the active and important role stakeholders can have even under restrictive policy practices, differentiating the written or official policies from the ones that are brought to practice by stakeholders.

**Extract 45:** Int #5: Sch/Adm-PB-ANT

*“Well, in general terms, we as a university have distanced ourselves from those official discourses around bilingualism. We do not share that view. And usually, we have had a critical position, and in our actions, we have marked a difference with those official discourses. We do not believe in them, but that doesn't mean ... that everything the government does is bad, no.”*

This previous extract denotes a critical position towards the government's approach to bilingualism “official discourses” and agency manifestation in the institution's practices which are guided not only by powerful policies but also by their critical view of bilingualism and official policies. The stakeholder (an institution) acknowledges that they do not share the same view as the government, implying that they have analyzed and evaluated the government's policies and actions related to bilingualism and taken a stance about it. By distancing themselves from official discourses around bilingualism and marking a

“difference” from official discourses in their practices, the institution showcases their agency exercising and active decision-making in shaping their own practices and perspectives. It suggests that they are not blindly following the government's approach but rather actively analyzing and determining the positive and negative aspects of language policies and government actions related to bilingualism. This further denotes their criticality, since, even with the ideological differences, they are able to discern the good and bad official policy practices of the government as well.

This finding is consistent with the theory and research on language policies as sociocultural phenomena, where, regardless of the policy approach, stakeholders end up involved and actively appropriating, interpreting, resisting and/or resignifying the policies to their contexts and within their possibilities. As Peláez and Usma (2017) found, the role of the different education stakeholders, especially teachers, is essential in the *enactment* of language policies. The data, similar to that of Usma's (2015), and García and Velásco's (2012) may evidence the highly complex nature of policy making and the frequent disconnection of some national mandates or practices with local realities and needs, the reason why local stakeholder's participation gains special need.

The extracts that were analyzed seem to show that stakeholders in the western-center Colombian education system are aware of the restrictions imposed by given (de-facto) language policies, yet they seem to have also exercised their agency by questioning, debating and resisting those policies, corroborating that policy making and enacting is a contested, dynamic, and unpredictable process (Honing, 2006; Ozga, 2000), in which stakeholders are rarely passive actors but rather critical ones in the manifestation and resignification of policies (Ricento & Hornberger, 1996). The data suggests that most stakeholders have appropriated to some extent language policies and reappropriated them to meet their local needs and interests, highlighting the role they have in shaping official policies and discourses on language education and language assessment.

### Conclusions

This study explored and described the interplay of language policies with assessment and evaluation practices in West-center Colombia from the view and experiences of a variety of stakeholders in the language education system and information in selected policy documents. This study evidences the complexities of the relationship of language policies and language assessment/evaluation especially in light of the analysis of language policy manifestations from sociocultural (e.g., Levinson & Sutton, 2001) and critical (Toffelson, 2006) lenses that view policies as conflicting terrains in which stakeholders, in this case mainly high- and medium-level stakeholders, play a relevant role in the final policy enactment (Usma, 2015). Yet, significant insights were obtained to answer the established questions and the research objectives and draw some conclusions.

Firstly, in regards to the question about the influence of language policies in assessment and evaluation practices, it can be concluded that, even though recent language policy discourses have been shifting to new, more alternative and heteroglossic paradigms, traditional powerful language policies and official practices such as high-stakes standardized tests, which have become de-facto policies (Menken, 2017) due to its accountability uses, are still exercising great pressures and influence on institutions, specially schools. These pressures have driven schools to highly focus on getting the best performance possible in national tests (tests offering a limited profile of students English competence of bilingualism profile), limiting, thus, language development and assessment possibilities, narrowing curricula by focusing on teaching to the test and other washback conditions like replicating national tests at the school and classroom level, and maintaining monoglossic practices towards the English language (Guerrero, 2008).

Not only are public and private schools highly influenced by national tests as de facto policies, but other instances were shown to be influenced as well. On the one hand universities that view national tests as an opportunity to portray higher status in their institutional competition or facing high-quality certification processes also respond to some

extent to these policy pressures. On the other hand, even local and national stakeholders were found to be influenced in their rush to improve test results and showing management results. For instance, the fact that an app development effort was designed around the *Saber 11* Test format sheds light on the extent to which the test has become a de-facto policy, guiding not only language teaching and learning practices, but even shaping to some extent the policy strategies towards language development in the country.

Additionally, other factors were found to be influencing both policy planning and assessment views and practices, namely transnational trends and market interests. From the data analysis, it can be concluded that transnational trends such as (foreign) standard and outcome-oriented approaches (Molstad et al., 2020) teaching and assessment have highly influenced language policies and, in turn, stakeholders' views and especially practices of language assessment, since they have to respond to such policies. These trends can guide assessment practices towards a standardized, de-contextualized approach rather than context-sensitive approach that is responsive to student's needs and learning objects, which may vary across the country.

A very illustrative aspect of this trend is the high influence that the CEFR has played in the Colombian language education narratives and practices. It has become the base for language policy planning and most formal assessment practices, and the extent to which discourses on language development were influenced by aspects such as the CEFR labels were highly noticeable. However, as the data also shows, many times only some specific aspects of the framework are taken into account, that being the labels, so as to show that programs and practices comply with government requirements. Even though the framework itself offers many positive aspects, what becomes problematic is the adoption rather than adapting practices (Cadavid et al., 2004), the way it is homogeneously fostered across the country and the marketable and accountability process that have surrounded the incorporation of the foreign framework (Int #5: Sch/Adm-PB-ANT).

A final factor that influenced assessment practices are the economic forces of private, typically, transnational policies, which have acquired significant authority and swayed over policy planning and implementation, resulting in policies and practices that prioritize certain language types, teaching methods, and assessment decisions, such as favoring foreign language proficiency tests, entrusting assessment tool developments to transnational companies rather than local experts and driving language teachers and other language users to pay expensive proficiency tests to access job and education opportunities. These approaches are frequently driven by financial interests rather than authentic efforts to promote language growth and cultivate inclusive, democratic teaching, and assessment practices in local settings, and have led to the belief that outer developments, such as proficiency tests or material designs, are better than the ones that could be developed by local talent. Therefore, this influence of language policies and related factors on language assessment practices highlights the need to critically examine the influence of language policies on stakeholders' practices and to develop more inclusive and contextually relevant approaches to national and local (language assessment) practices in Colombian bilingual education.

Despite the complex powerful forces that influence stakeholder's language assessment and evaluation practices, from the different cases, it can be concluded that many stakeholders have also played a very active role in policy discussions, enactment and even influencing language policies themselves. In doing so, emergent practices were identified, such as alternative assessment practices, increasing inclusion of technologies to language assessment and an emerging heteroglossic view of language assessment. The fact that later policy efforts, even though they are not very influential yet, are starting to focus on more formative views of language assessment, evidences an influence of stakeholders such as teachers and scholars who have urged for this alternative approach to language assessment.

By embracing formative and alternative assessment practices in language policy, incurring in and exploring heteroglossic bilingual assessment, and integrating emerging



technologies, bilingual education in Colombia may develop a more inclusive, contextually-relevant, and effective approach to language education and language assessment, making thus, fair, democratic, reliable and valid language assessment and evaluation practices an essential part of a bilingualism development in the country. It is worth noting that this increasing advance-guard tech integration that has started already and stakeholders foresee in the near future has both 1) the potential to enhance alternative assessment practices (Fox, 2018) and heteroglossic approaches to bilingual education (Lopez, et al. (2014), and 2) the potential to increase accountability measures if used from traditional top-down practices of language policies and language assessment.

Regarding the language ideologies of both selected stakeholders and language policies, a variety of views on bilingualism and bilingual education were found. For example, diverging views relating to several dimensions of language such as the linguistic, educative and social aspects were noticed, some getting closer to monoglossic views and some others to heteroglossic ones. Nonetheless, from the data it can be concluded that there is an emerging shift in the country in the views of bilingualism and bilingual education, portraying the positive influence that scholars on the field of bilingualism have played. This exploration of policies' and stakeholder's ideologies was relevant since "Tests [and other assessment procedures] are deeply embedded in culture and ideology" (Brown & Abeywickrama, 2018) and understanding language and assessment ideologies helps to the overall understanding of assessment practices.

Additionally, it is worth noting that there are significant disparities between ideologies reflected in recent written policies, the effect of official (or de-facto) language policy implementations, and the perspective stakeholders have towards government's approach to language policies. Even though the written policy documents such as the *plan sectorial 2018-2022* (MEN, 2021), portray a "functional plurilingual" view of languages, official practices suggest hierarchies among languages, favoring foreign powerful ones, languages are addressed apart from each other (the National Bilingualism Program focuses

mainly in English or other foreign languages, cultural branches of the government focus on Indigenous and Sign languages, and Spanish is covered in the general education system). Additionally, language development and assessment strategies such as Be the One challenge focus on the type of English and skills fostered by the *Pruebas Saber 11*. Furthermore, the way lower-level and mid-level stakeholders perceive the policies is significantly apart from the written views or that of the national policymaker, with the former ones perceiving official approaches to bilingualism as market-led, monoglossic, and de-contextualized for some contexts of the country's realities. This shows how traditional powerful policies and assessment and evaluation practices such as the national test exert a great influence on the collective view of what language policies in the country are, shadowing later strategies.

Overall, this project has contributed to the understanding of the interplay between language policies and stakeholder's assessment practices in the west-center side of Colombia, from the insights of stakeholders from a diversity of regions and positions in the language education system. It has also contributed to the understanding of the complex policy dynamics of the country under the understanding of stakeholders' potential to actively participate and influence policy enactment, and in this case, even policy planning. Finally, it has shed light on the language ideologies behind policies and stakeholders, and the ideological shifts that are taking place in west-center Colombian language education. All this calls for a number of recommendations and suggestions which will be posed in the following sections.

### **Implications**

In light of the findings, a number of implications can be put forward. Firstly, even though language ideology diversity is natural and even expected in a plural and diverse country, it might be the case that the lack of a consensual approach to bilingualism and bilingual education by stakeholders and language policies can end up in inconsistent language policy efforts, lack of stakeholder articulation towards common ends, and power imbalances among stakeholders. This, and the gap between medium- and low- stakeholders, and policymakers and official policies, calls for the need of a more active conversation among stakeholders, from policymakers to teachers (and even students), to both engage in conversations on what bilingual education policies should look like and aim for in the country, and the different regions of the country, now, in light of current theoretical and practical paradigms as well as local stakeholder's experiences and needs. In this regard, policy makers should, for example, consider the development of technologies that generate the conditions for a more democratic participation of stakeholders. These technologies could allow students to be heard and participant in the writing and enactment of the policies.

Some insights that emerge about the nature of these policies to come, is that they should be highly participatory and sensitive to local contexts, understanding that language needs vary across the country and so should language policies and their goals, and local stakeholders can give relevant insights on such needs and necessary goals. Similarly, policy makers should be more congruent on what they define in the written documents and what they actually end up promoting when they are officially implemented, and the consequences of traditional policies and official assessment practices such as national tests should be reconsidered so that they do not shadow newer efforts.

Additionally, it is important to acknowledge and question the influence and power of certain transnational policies and economic interests that have also directed to some extent national practices and even views of language, teaching and assessment. It is important to critically assess the impact of these transnational institutions and policies and ensure that

local policies prioritize language learner's actual learning needs rather than transnational interests or solely profit-driven goals. In this sense, it is also important to reconsider the place of foreign language assessment tools or assessment designs which have become an economic barrier and an excessive public budget expense. Were it concluded that assessment practices such as proficiency certifications were necessary, articulated language assessment development efforts could be promoted in the country taking advantage of the now highly competent and experienced local academic human power.

In a similar line, it is important to take a closer look at these emerging assessment practices: foster the diversification of teaching and assessment procedures, and more deeply explore, through research, how heteroglossic practices are emerging in the country. Also, as stakeholders noticed, heteroglossic bilingual assessment is a relatively new field, especially in the country, so there is a need to keep studying and developing assessment processes and tools under this new approach in the country. In this direction, emerging technologies may come handy in coping with the adaptability and multimodal nature that heteroglossic bilingual assessment requires.

Regarding school level implications, the findings call for the broadening of language learning curricula to go beyond high-stakes tests, and the limited skills they assess, to a more meaningful and holistic approaches to language learning. Similarly, the findings also suggest a context-sensitive and democratic approach to curriculum development (including assessment) understanding the possibilities and language needs of the population to impact since it may diverge to some extent from national homogeneous goals. Finally, findings suggest the relevance of incurring in theoretically- and contextually-informed heteroglossic teaching and assessment practices in which languages are brought together and used to nurture one another.

### **Limitations**

Even though this research process was rigorous and transparent, trying to follow the best researcher practices to offer reliable findings and insights, a number of limitations should be acknowledged. Firstly, the nature of the study, that is, a multi-case qualitative study, offers an in-depth insight from a few cases but broader or narrower approaches could be useful to complement the findings. Similarly, there were restricted data sources, semi-structured interviews and document analysis, which can offer numerous insights on the phenomenon. However, they do not draw the whole picture since relevant data can be collected from other sources as well. Besides, the data may be subject to biases since it relies on stakeholders' views and experiences as an understanding of the system and the policies in practice.

For the purposes, scope and possibilities of this research, a limited number of policies from an 8 year-span were selected, which do not portray the whole picture of language policies. Therefore, there may be other factors influencing language policies and bilingualism in West-center Colombia that were not considered or explored in this study. Finally, as it was presented in the positionality statement (See method section), the researcher himself holds biases towards language, education and assessment that may influence data interpretation, and such views have added a criticality view to some of the findings.

### Recommendations and Future Directions

A number of recommendations and suggested lines of research can be set forth. Firstly, from the data collection process it was noticed that there is not a recent bilingualism policy in place, but rather policy documents such as the *plan sectorial 2018-2022*, which sheds lights on the governments approach and views on language development to be implemented from the National Bilingualism program. When the researcher kindly requested the MEN to share the current policy, they shared PowerPoint presentations reporting on the *plan sectorial* and the National Bilingualism Program. Therefore, there is clearly a need to develop a national bilingualism policy that emerges from an ample conversation among stakeholders and a high sensitivity to what the language realities and needs are in the different localities of the country.

In that regard, the national policy maker suggested to develop in the short term an intersectoral “national bilingualism policy (*política intersectorial de bilingüismo*)” (Int #1: PM-PB-DC). That is, a great policy planning effort that goes beyond the ministry of education to other ministries such as industry and commerce, tourism, technology, companies, the national education service (*SENA*) and universities. Other stakeholders that could be added to this policy efforts are teachers and administrators (of the different types of language education, such as public and private schools or language centers), and scholars. It is worth noting that such effort should end up in a flexible yet long term policy effort that is not dependent on elected governments that should implement short-term policy strategies<sup>7</sup> but rather a state policy or program with a long-term vision of language development needs and goals.

This project has also evidenced the need for more collaboration and articulation among universities public schools supported by the government. Provided the conditions, universities have the potential to come up with assessment and evaluation instruments that

---

<sup>7</sup> “There would have to be an institution belonging to the state rather than the government. ... If there was an independent organism for bilingualism, probably more long-term policies could be put forward.” (INT #2: TD/Ter-PB-RIS)

both inform governments on language developing processes and enhance language learning from heteroglossic perspectives, empowering, thus, national stakeholders in general instead of undermining local abilities and misdirecting resources towards transnational institutions.

In regards to research in the field of language assessment, it is recommended to keep exploring and spreading heteroglossic assessment practices from a multiplicity perspective, with a close look at the particularities of the contexts and the learning objectives/needs. In this regard, there is a great research gap to be explored, due to the recentness of this approach to language assessment, especially in the country. Besides, articulated efforts could be posed to set forth locally-developed standardized assessment tools for such instances as they are necessary, so that local knowledge is more valued and economic or cultural barriers to accessing job or education opportunities are diminished.

Overall, it is crucial to continue exploring in research the complex relationships between language policies and assessment practices (and language education in general). For example, by extending the focus to other regions within the country, a more comprehensive understanding of relationship between policies and practices and how they shape the education (and overall assessment) system. Additionally, adopting a broader approach to this matter by including a wider range of stakeholders possibly through quantitative research methods, can add valuable insights into the factors that come into play in this assessment-policy relationship. Conversely, digging deeper into single cases or smaller contexts can offer more detailed understanding of specific contexts and illustrate the unique realities and needs of stakeholders in different settings. Employing both a wider and a narrower approach to this phenomenon can significantly contribute to the body of knowledge that informs practitioners, researchers and policymakers. Finally, this research should also take more into account teachers' assessment practices in different contexts.

### References

- Alonso Cifuentes, J., Díaz Mejía, D., & Estrada Nates, D. (2018). Achievements of a bilingual policy: The Colombian journey. In B. S. Chumbow (Ed.), *Multilingualism and Bilingualism* (pp. 57-74). IntechOpen. <https://doi.org/10.5772/intechopen.68609>
- Anderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14, 115-129.
- Arias, C. E. (2016). *Translingüismo y aprendizaje integrado de lengua y contenido como modelo de educación bilingüe dinámica*. [Tesis de doctorado]. Universidad Tecnológica de Pereira.
- Ayala Zárate, J., & Álvarez V., J. A. (2005). A perspective of the implications of the Common European Framework implementation in the Colombian Socio-cultural context. *Colombian Applied Linguistics Journal*, 7, 7-26. <https://doi.org/10.14483/22487085.162>
- Bachman, L.F. (1990). *Fundamental Considerations in Language Testing*. Oxford University Press.
- Bailey, K., (2009) Issues in Language Teacher Evaluation. . In Long, M., & Doughty, C. (Eds.) *The Handbook of Language Teaching*. Blackwell Publishing Ltd. ISBN: 978-1-405-15489-5
- Baker, C. (2001). *Foundations of Bilingual Education and Bilingualism*. Multilingual Matters LTD. 3rd ed.
- Ball, S. J. (2003). The teacher's soul and the terrors of performativity. *Journal of Education Policy*, 18(2), 215-228.
- Ball, S. J. (2007). Introduction: The seething and swirling of education policy. In M. Sutton (Ed.), *Education, Globalisation and New Times* (pp. ix-xvi). Routledge.
- Barni, M., & Salvati, L. (2017). The Common European Framework of Reference (CEFR). In E. Shohamy. (Ed.), *Language Testing and Assessment*. pp. 417-426. Springer
- Barrance, R., & Elwood, J. (2018). National assessment policy reform 14–16 and its consequences for young people: student views and experiences of GCSE reform in



- Northern Ireland and Wales. *Assessment in Education: Principles, Policy & Practice*, 25(3), 252-271. <https://doi.org/10.1080/0969594X.2017.1410465>
- Bialystock, E. (2001). *Bilingualism in development: Language, literacy and cognition*. Cambridge: Cambridge University Press.
- Bidwell, C., & Quiroz, P. (1991). Organizational control in the high school workplace: A theoretical argument. *Journal of Research on Adolescence*, 1, 211-229.
- Birdsong, D. (2015). Dominance in bilingualism: Foundations of measurement, with insights from the study of handedness. In C. Silva-Corvalán & J. Treffers-Daller (Eds.), *Language Dominance in Bilinguals: Issues of Measurement and Operationalization* (pp. 85-105). Cambridge University Press.
- Block, D. (2003). *The social turn in second language acquisition*. Washington, DC: Georgetown University Press.
- Bloomfield, L. (1935). *Language*. Allen & Unwin.
- Bonilla Carvajal, C. A., & Tejada-Sanchez, I. (2016). Unanswered questions in Colombia's foreign language education policy. *Profile. Issues in Teachers' Professional Development*, 18(1), 185-201. <https://doi.org/10.15446/profile.v18n1.51996>
- Bowen, G.A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, 9(2), 27-40. 10.3316/QRJ0902027
- Bray, M., & Thomas, R. (1995). Level of comparison in educational studies: Different insights from different literatures and the values of multilevel analysis. *Harvard Educational Review*, 65(3), 472-490.
- Broadfoot, P. (1987). *Introducing Profiling: A practical guide*. Macmillan Education.
- Brown, H. D., & Abeywickrama, P. (2018). *Language Assessment: Principles and Classroom Practices* (3rd ed.). Pearson Education.
- Brown, J. D. (2005). *Testing in Language Programs: A Comprehensive Guide to English Language Assessment*. Prentice Hall Regents.
- Byram, M., & Parmenter, L. (Eds.). (2012). *The Common European Framework of Reference. The globalization of language education policy*. Multilingual Matters.

- Cadavid Múnera, I. C., McNulty, M., & Quinchía Ortiz, D. I. (2004). Elementary English language instruction: Colombian teachers' classroom practices. *Profile*, 5, 37-55.  
<https://repositorio.unal.edu.co/handle/unal/25291>
- Cazden, C & Snow, C. (Eds.). (1990). English Plus: issues in bilingual education. *The annals of the American Academy of Political and Social Science*. V-508
- Cenoz, J. & Gorter, D. (2012). Language policy in Education. B. Spolsky (Ed.), *The Cambridge Handbook of Language Policy*. (pp. 301-319). Cambridge University Press.
- Chapelle, C.A., Kremmel, B., & Bridley G. (2020). Assessment. In N. Schmitt (Ed.), *An Introduction to Applied Linguistics*. Hodder Education. 247-267
- Chapelle, C.A., & Plakans, L. (2012). Assessment and Testing: Overview. In C. A. Chapelle (Ed.), *The Encyclopedia of Applied Linguistics*. Wiley-Blackwell.  
<https://doi.org/10.1002/9781405198431.wbeal0603>
- Choi, I. C. (2008). The impact of EFL testing on EFL education in Korea. *Language Testing*, 25, 39-62.
- Colombia, Ministerio de Educación Nacional (MEN). (2005). Programa Nacional de Bilingüismo. Colombia 2004- 2019. Retrieved from [https://www.mineduacion.gov.co/1621/articles-132560\\_recurso\\_pdf\\_programa\\_nacional\\_bilinguismo.pdf](https://www.mineduacion.gov.co/1621/articles-132560_recurso_pdf_programa_nacional_bilinguismo.pdf)
- Colombia, Ministerio de Educación (MEN). (2006) Guía No. 22: Guía Estándares Básicos de Competencias en Lenguas Extranjeras: Inglés. Retrived from <https://www.mineduacion.gov.co/portal/men/Publicaciones/Guias/115174:Guia-No-22-Estandares-Basicos-de-Competencias-en-Lenguas-Extranjeras-Ingles>
- Colombia, Ministerio de Educación Nacional (MEN). (2014). Programa nacional de inglés 2015-2025. Retrieved from <https://www.ucn.edu.co/institucion/sala-prensa/noticias/Documents/2014/Documento%20PROGRAMA%20NACIONAL%20DE%20INGL%C3%89S%202015-2025.pdf>

- Colombia, Ministerio de Educación (MEN). (2016). Currículo Sugerido de inglés para sexto a undécimo y Derechos básicos de Aprendizaje. Retrieved from <http://aprende.colombiaaprende.edu.co/es/colombiabilingue/88005>
- Colombia, Ministerio de Educación (MEN). (2021) Plan sectorial 2018-2022: Pacto por la equidad, pacto por la Educación. Retrived from [https://www.mineducacion.gov.co/1759/articles-349495\\_recurso\\_140.pdf](https://www.mineducacion.gov.co/1759/articles-349495_recurso_140.pdf)
- Common European framework of reference for languages: learning, teaching, assessment.* (2002). Cambridge University Press.
- Conteh, J. & Meier, G. (2014). *The multilingual turn in language education: Opportunities and challenges.* Multilingual Matters.
- Copp, D. (2016). Teaching to the test: a mixed methods study of instructional change from large-scale testing in Canadian schools. *Assessment in Education: Principles, Policy & Practice*, 25(5), 468-487. 10.1080/0969594X.2016.1244042
- Correa, D., & Usma Wilches, J. (2013). From a bureaucratic to a critical-sociocultural model of policymaking in Colombia. *HOW Journal*, 20(1), 226-242. <https://www.howjournalcolombia.org/index.php/how/article/view/3>
- Cummins, J. (1981). Four misconceptions about language proficiency in Bilingual Education. *NABE: The Journal of the National Association for Bilingual Education*, 5, 31-45.
- Cummins, J. (1984) *Bilingualism and special education: Issues in assessment and pedagogy.* Multilingual Matters.
- De Houwer, A., & Ortega, L. (eds.). (2019). *The Cambridge handbook of bilingualism.* Cambridge University Press. 10.1017/9781316831922
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies.* OUP Oxford.
- Döpke, S. (1992). *One parent, one language: an interactional approach.* J. Benjamins Publishing Company.
- Education, D. o. B. a. S. S. a., Committee on Foreign Language Assessment for the U.S. Foreign Service Institute, & National Academies of Sciences, Engineering, and

- Medicine. (2020). *A Principled approach to language assessment: Considerations for the U.S. Foreign Service Institute*. National Academies Press.  
<https://doi.org/10.17226/25748>
- Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Cambridge University Press.
- Escobar, C., & DillarD-Paltrineri, E. (2015). Professors' and students' conflicting beliefs about translanguaging in the EFL classroom: Dismantling the monolingual bias. *Revista de Lenguas Modernas*, 23, 301-328.  
<https://doi.org/10.15517/rlm.voi23.22355>
- Extra, G., Spotti, M., & van Avermaet, P. (Eds.). (2009). *Language testing, migration and citizenship: Cross-national perspectives*. Continuum.
- Fairclough, N., & Wodak, R. (1997). Critical discourse analysis. In T. A. Van Dijk (Ed.), *Discourse as Social Interaction*. SAGE Publications.
- Fettes, M. (2003). "The geostrategies of interlingualism." In J. Maurais, and M. Morris (Eds.), *Languages in a globalising world*. Cambridge University Press, Cambridge, pp. 37-46
- Ferguson, C. A., Houghton, C. & Wells, M.H. (1977) Bilingual education: An international perspective. In B. SPOLSKY & R. COOPER (Eds.), *Frontiers of bilingual education*. (pp. 159-174). Newbury House.
- Fishman, J. A. (1977). *Bilingual education: Current perspectives*. Arlington Center.
- Foucault, M. (1979). *Discipline and punish: The birth of the prison* (trans: from the French by Alan Sheridan). Vintage Books.
- Fox, J. (2018). Using portfolios for assessment/alternative assessment. In E. Shohamy, L. G. Or, & S. May. (Eds.), *Language testing and assessment. (3rd ed., pp. 135-148)*. Springer.
- Fraenkel, J., Wallen, N., & Hyun, H. (2011). *How to design and evaluate research in education*. McGraw-Hill Education.
- García, O. (2009). *Bilingual Education in the 21st Century: A Global Perspective*. Wiley.

- García, O. (2014). Multilingualism and language education. In C. Leung & B. V. Street (Eds.), *The Routledge Companion to English Studies*. Routledge.
- García, O. (2019). Doing and Undoing Bilingualism in Education. In De Houwer, A. & Ortega, L. (Eds.), *The Cambridge Handbook of Bilingualism*. Cambridge University Press.
- García, O., Kleifgen, J. A., & Falchi, L. (2008). From English language learners to emergent bilinguals. *Equity Matters: Research Review No. 1 New York: A Research Initiative of the Campaign for Educational Equity*.
- García, O., & Velasco, P. (2012). Insufficient language education policy: Intercultural bilingual education in Chiapas, Diaspora. *Studies of Migration, Integration, Equity, and Cultural Survival*, 6(1), 1-18. <http://dx.doi.org/10.1080/15595692.2011.633129>
- Giraldo, F. (2018). Language assessment literacy: Implications for language teachers. *Profile: Issues in Teachers' Professional Development*, 20(1), 179-195. <https://doi.org/10.15446/profile.v20n1.62089>.
- Gómez Sará, M. M. (2017). Review and analysis of the Colombian foreign language bilingualism policies and plans. *HOW*, 24(1), 139-156. <http://dx.doi.org/10.19183/how.24.1.343>.
- Graddol, D. 2006. *English next: Why global English may mean the end of English as a foreign language*. British Council.
- Gronlund, N. E. (1998). *Assessment of student achievement* (6th ed.). Allyn & Bacon.
- Grosjean, F. (1985, a). The bilingual as a competent but specific speaker-hearer. *The Journal of Multilingual and Multicultural Development*, 6(4), 67-77.
- Grosjean, F. (1989). Neurolinguists beware! The bilingual is not two monolinguals in one person. *Brain and Language*, 36(1), pp.3-15. [https://doi.org/10.1016/0093-934X\(89\)90048-5](https://doi.org/10.1016/0093-934X(89)90048-5)
- Grosjean, F. (1996). Living with two languages and two cultures. In I. Parasnis (Ed.), *Cultural and language diversity and the deaf experience*. Cambridge University Press.

- Guerrero, C. H. (2008). Bilingual Colombia: What does it mean to be bilingual within the Framework of the National Plan of Bilingualism. *Profile: Issues in Teacher's Professional Development*, 10, 27-45.
- ICFES (2021). Informe nacional de resultados del exam Saber 11 2020. Vol 1.
- Hamers, J. F., & Blanc, M. (2004). *Bilinguality and Bilingualism* (2nd ed.). Cambridge University Press.
- Hanson, A. (1993). *Testing, Testing*. University of California Press. ISBN:0-520-08060-2.
- Heck, R. (2004). *Studying educational and social policy: Theoretical concepts and research methods*. Lawrence Erlbaum Associates Publishers.
- Herazo, J., Jerez, S., & Lorduy, D. (2012). Opportunity and incentive for becoming bilingual in Colombia: Implications for Programa Nacional de Bilingüismo. *Íkala, revista de lenguaje y cultura*, 17(2), 199–213
- Herdina, P. and Jessner, U. (2002) *A dynamic model of multilingualism*. Multilingual Matters
- Honig, M.I. (Ed.). (2006). *New directions in education policy implementation: Confronting complexity*. State University of New York Press.
- Hopfenbeck, T. N., Flórez Petour, M. T., & Tolo, A. (2015). Balancing tensions in educational policy reforms: Large-scale implementation of assessment for learning in Norway. *Assessment in Education: Principles, Policy & Practice*, 22(1), 44-60.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge University Press.
- Hurie, A. H. (2018). ¿Inglés para la paz? Colonialidad, ideología neoliberal y expansión discursiva en Colombia *Bilingüe. Íkala, Revista de Lenguaje y Cultura*, 23(2), 333-354. DOI: 10.17533/udea.ikala.v23n02a09
- Janssens, G., & Meier, V. (2013). Establishing placement test fit and performance: Serving local needs. *Colombian Applied Linguistics Journal*, 15(1), 100-113.
- Jernudd, B., & Nekvapil, J. (2012). History of the field: a sketch. In B. Spolsky (Ed.). *The Cambridge handbook of language policy*. Cambridge University Press. (pp. 16-36)

- Jiménez, M., Rodríguez, C., & Rey Paba, L. (2017). Standardized test results: An opportunity for English program improvement. *HOW Journal*, 24(2), 121-140.  
<https://doi.org/10.19183/how.24.2.335>
- Jones, G., Jones, B., Chapman, L., Yarbrough, T., & Davis, M. (1998). The impacts of high-stakes testing on teachers and students in North Carolina. *Phi Delta Kappan*, 81(3), 199-203.
- Koretz, D. (2017). *The testing charade: Pretending to make schools better*. University of Chicago Press.
- Kuhail, M.A., Alturki, N., Alramlawi, S. et al. (2023) Interacting with educational chatbots: A systematic review. *Educ Inf Technol* 28, 973–1018 <https://doi.org/10.1007/s10639-022-11177-3>
- Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. Sage.
- Lafford, B. (2007). Second language acquisition reconceptualized? The impact of Firth & Wagner. *The Modern Language Journal*, 91, 735–756.
- Lambert, W. E. (1955) Measurement of the linguistic dominance in bilinguals. *Journal of Abnormal and Social Psychology*, 50, pp. 197–200.
- Lasagabaster, D., & García, O. (2014). Translanguaging: towards a dynamic model of bilingualism at school/Translanguaging: hacia un modelo dinámico de bilingüismo en la escuela. *Culture and Education*, 26(3), 557-572.  
10.1080/11356405.2014.973671
- Levinson, B. A. U., & Sutton, M., (2001). Introduction: Policy as/in practice- a sociocultural approach to the study of educational policy. In M. Sutton (Ed.), *Policy as practice: Towards a comparative sociocultural analysis of educational policy* (pp. 1-22). Ablex.
- Lopez, A. (2009). Impact of a policy-driven test effort on foreign language programs in Colombian universities. *Paper presented at Pensando Educación., Universidad de los Andes, Bogotá, Colombia.*

- Lopez, A., & Arendina, R. (2009). Language testing in Colombia: a call for more teacher education and teacher training in language assessment. *Profile*, 11(2), 55-70.
- Lopez, A. A., Guzman-Orth, D., & Turkan, S. (2014, March). *A study on the use of translanguaging to assess the content knowledge of emergent bilingual students*. Paper presented at the Invited Colloquium, Negotiating the Complexities of Multilingual Assessment, at the annual American Association for Applied Linguistics (AAAL) Conference.
- Lopez, A., & Janssen, G. (2010). Validation study of Colombia's ECAES English exam. *Lenguaje*, 38, 423-448.
- López, A. A., Pacheco, J. R., & Peralta, J. C. (2011). Estudio de validez del examen de estado Saber 11 de inglés. *Folios*, 34(2), 77-91.
- Lopez, A., Turkan, S., & Guzman-orth, D. (2017) Assessing multilingual competence. In E. Shohamy, L. G. Or. & S. May (Eds.), *Language testing and assessment*. (3rd. Ed.), Springer.
- Lovón-Cueva, M. A., & Quispe-Lacma, A. P. (2020). ¿Quién tiene derecho a opinar sobre política lingüística en Perú? Un análisis crítico del discurso. *Íkala, Revista de Lenguaje y Cultura*, 25(3), 733-751. <https://doi.org/10.17533/udea.ikala.v25n03a12>
- Mackenzie, L. (2020). A critical evaluation of the social justice implications of the Colombian government's English-Spanish bilingualism policies. *Current Issues in Language Planning*. 21(2), 117-134. DOI: 10.1080/14664208.2019.1615743
- McKay, S., & Brown, J. D. (2016). *Teaching and assessing EIL in local contexts around the world*. Routledge.
- McNamara, T. (1998). Policy and social considerations in language assessment. *Annual Review of Applied Linguistics*, 18, 304–319.
- McNamara, T. (2000). *Language Testing*. Oxford University Press.
- McNamara, T. (2009). Validity in language testing: The challenge of Sam Messik's legacy. *Language Assessment Quarterly*, 3(1), 31-35



- Menken, K. (2017) High-Stakes tests as de facto Language education policies. In E. Shohamy, L. Or, & S. May (Eds.), *Language testing and assessment*. (3rd ed., pp. 185-196). Springer.
- Menken, K and Shohamy, E. (2015). Invited colloquium on negotiating the complexities of multilingual assessment, AAAL Conference 2014. *Language Teaching*, 48, 421-425  
doi:10.1017/S0261444815000166
- Messick, S. (1980). Test validity and the ethics of assessment. *American Psychologist*, 35, 1012-1027.
- Messick, S. (1981). Evidence and ethics in the evaluation of tests. *Educational Research*, 10, 9-20.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (pp. 13-103). American Council on Education & Macmillan.
- Messick, S. (1996). Meaning and values in test validation: The science and ethics of assessment. *Educational Research*, 18, 5-11.
- Messick, S. (1996). Validity and washback in language Testing. *Language Testing*, 13(3), 241-256.
- Metz, M. H. (2003). *Different by design: The context and character of three magnet schools*. Teachers college press.
- Meyer, J. (2008). Building education for a world society. In M. Pereyra (Ed.), *Changing knowledge and education: Communities, mobilities, and new policies in global societies* (pp. 31-49). Peter Lang.
- Molstad, C., Proitz, T., & Dieude, A. (2020). When assessment defines the content: Understanding goals in between teachers and policies. *The Curriculum Journal*, 32(2), 290-314.
- Mora, R. A., Chiquito, T., & Zapata, J. D., (2019). Bilingual Education Policies in Colombia: Seeking Relevant and Sustainable Frameworks for Meaningful Minority Inclusion. In: Johannessen, B. (ed). *Bilingualism and Bilingual Education: Politics, Policies and Practices in a Globalized Society* (pp. 55–77). Springer.

- Morrow, K. (Ed.). (2004). *Insights from the Common European Framework*. Cambridge University Press.
- Mousavi, S. A. (2009). *An Encyclopedic Dictionary of Language Testing* (4th ed.). Rahnama Publications.
- Muñoz, A., Palacio, M., & Escobar, L. (2012) Language Assessment Literacy: Implications for Language Teachers. *PROFILE Issues of Teacher's Educational Development* 14(1), pp.143-158. ISSN 1657-0790.
- Norris, J. (2009) Task-Based Teaching and Testing. In Long, M., & Doughty, C. (Eds.) *The Handbook of Language Teaching*. Blackwell Publishing Ltd. ISBN: 978-1-405-15489-5
- Pan, Y.-C. (2009). Evaluating the appropriateness and consequences of test use. *Colombian Applied Linguistics Journal*, 11, 93-105.
- Patton, M. Q. (2002). *Qualitative research and evaluation methods*. SAGE Publications.
- Peláez, O., & Usma, J. (2017). The crucial role of educational stakeholders in the appropriation of foreign language education policies: A case study. *PROFILE Issues of Teacher's Educational Development*, 19(2), 121-134.  
<http://dx.doi.org/10.15446/profile.v19n2.57215>.
- Peñafort, C. (2002). Experiencias en la construcción de propuestas curriculares bilingües para colegios monolingües de la ciudad de Cali. In A.-M. De Mejía & R. Nieve (Eds.), *Nuevos caminos en educación bilingüe en Colombia*. Universidad del Valle, Escuela de Ciencias del Lenguaje.
- Perrow, C. (1986). *Complex organizations: A critical essay* (3rd ed.). Random House.
- Quintero Polo, A. H. (2009). ¿Es usted bilingüe? Concepciones y alternativas para la educación en idiomas en el contexto colombiano [Are you bilingual? Conceptions and alternatives for language teaching in the Colombian context]. *El Educador*, 7, 4-10.  
Retrieved from <http://es.calameo.com/read/000085494e33c599e8268>.

- Reyes, Y. (2015, February 2). Colombia, very well! El Tiempo. Retrieved from <http://www.eltiempo.com/opinion/columnistas/colombia-very-well-yolanda-reyes-columnista-el-tiempo/15181161>
- Romaine, S. (1989). *Bilingualism*. B. Blackwell.
- Ricento, T. K., & Hornberger, N. H. (1996). Unpeeling the onion: Language planning and policy and the ELT professional. *Tesol Quarterly*, 30(3), pp. 401-427. <http://doi.org/10.2307/3587691>.
- Salma, N. F. (n.d.). *Exploring Van Dijk: Critical Discourse Analysis's Aims*. FKIP Universitas Lancang Kuning.
- Schüppert, A., & Gooskens, C. (2012). The role of extralinguistic factors for receptive bilingualism: Evidence from Danish and Swedish preschoolers. *International Journal of Bilingualism*, 16(3), 332-347.
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Longman.
- Shohamy, E. (2006). *Language policy: hidden agendas and new approaches*. Routledge.
- Shohamy, E. (2011). Assessing multilingual competencies: Adopting construct valid assessment policies. *The Modern Language Journal*. 95. pp. 418-429. <https://doi.org/10.1111/j.1540-4781.2011.01210.x>
- Shohamy, E. (Ed.). (2017) *Language testing and assessment*. (3rd ed.). Springer.
- Silva-Corvalán, C., & Treffers-Daller, J. (2015). Digging into dominance: A closer look at language dominance in bilinguals. In C. Silva-Corvalán & J. Treffers-Daller (Eds.), *Language Dominance in Bilinguals: Issues of Measurement and Operationalization* (pp. 1-14). Cambridge University Press. doi: 10.1017/CBO9781107375345.001
- Smith, M. L. (1991). Put to the test: The effects of external testing on teachers. *Educational Researcher*, 20(5), 8-11. doi:10.3102/0013189X020005008
- Spolsky, B. (2004). *Language Policy*. Cambridge University Press.
- Stavans, A. & Hoffmann, C. (2015). *Multilingualism*. Cambridge University Press.

- Steiner-Khamsi, G. (Ed.). (2004). *The global politics of educational borrowing and lending*. Teachers College Press
- Steiner-Khamsi, G. (Ed.). (2012). *Policy borrowing and lending in education*. Routledge.
- Steiner-Khamsi, G. (2000). Transferring education, displacing reforms. In J. Schriewer (Ed.). *Discourse formation in comparative education* (pp. 155-187). Lang publishers.
- Stake, R. E. (1995). *The art of case study research*. SAGE Publications.
- Stake, R. E. (2013). *Multiple case study analysis*. Guilford Publications.
- Stobart, G. (2003). The impact of assessment: Intended and unintended consequences. *Assessment in Education*, 16, 139-140.
- Toffelson, J. W. (2006). Critical theory in language policy. In T. Ricento (Ed.), *An introduction to language policy: Theory and method*. Blackwell, Malden.
- Treffers-Daller, J. (2015). Language dominance: The construct, its measurement and operationalization. In C. Silva-Corvalán & J. Treffers-Daller (Eds.), *Language dominance in bilinguals: Issues of measurement and operationalization*. Cambridge University Press.
- Treffers-Daller, J. (2019). The measurement of bilingualism. In A. De Houwer, & L. Ortega (Eds.), *The Cambridge handbook of bilingualism* (pp. 289-306). Cambridge University Press.
- Tsai, Y., & Tsou, C.-H. (2009). A standardised English Language Proficiency test as the graduation benchmark: student perspectives on its application in higher education. *Assessment in Education: Principles, Policy & Practice*, 16(3), 319-330.  
<http://doi.org/10.1080/09695940903319711>
- Valencia, S. (2005, October). Bilingualism and English language teaching in Colombia: A critical outlook. Paper presented at the Conference on English Language Teaching in Colombia, Universidad del Quindío.
- Usma, J. (2015). *From transnational language policy to local appropriation: The case of the National Bilingual Program in Medellín, Colombia*. Deep University Press.

Wall, D. (1997). Impact and washback in language testing. In C. Clapham & D. Corson (Eds.), *Encyclopedia of language and education* (pp. 291–302). Dordrecht: Kluwer Academic Publications

## Appendices

### Appendix 1: Informed Consent

#### *Formato de Consentimiento Por Participantes del Estudio.*

#### **Título del proyecto: INFLUENCIA DE LAS POLÍTICAS EDUCATIVAS BILINGÜES EN LAS PRACTICAS EVALUATIVAS DE LA EDUCACIÓN EL BILINGÜE EN COLOMBIA**

Yo, Luis F. Jaramillo, de la Universidad Tecnológica de Pereira (Teléfono: #####) estoy realizando un proyecto de investigación el programa de Maestría en Educación Bilingüe de la Universidad Tecnológica de Pereira sobre el tema arriba escrito.

El propósito general de esta investigación es explorar las relaciones que existen entre las políticas lingüísticas y las prácticas evaluativas en la educación bilingüe en Colombia. Con esta entrevista se pretende consultar a los actores interesados o relacionados con las políticas lingüísticas en torno al bilingüismo en el país y la manera en que ejercitan su agencia en su formulación, su aplicación y las prácticas evaluativas.

Si acepta participar de este estudio, se le solicitará participar de dos entrevistas semiestructuradas. Los datos recolectados pueden ser utilizados como información para el proyecto de tesis de Maestría. Su participación en este estudio es voluntaria y se puede retirar a cualquier momento. No tiene que dar ninguna justificación para retirarse del estudio. Antes de que el reporte final del proyecto sea entregado, se le enviará un resumen de lo que haya escrito con base a los datos recolectados de la entrevista les solicitará hacer los comentarios que crean necesarios sobre las descripciones e interpretaciones que ustedes crean que no sean exactas o acertadas. Los corregiré si es necesario. Cuando hagamos reportes de la investigación, nos aseguraremos de que usted no sea identificado como individuo; no se utilizarán referencias a nombres personales. Yo y el asesor de tesis seremos las únicas personas que tendrán acceso a los datos recolectados para el proyecto, a menos que se deban presentar en público, en cuyo caso se aplicará la confidencialidad y el anonimato. Cualquier dato que se utilice en los

reportes y publicaciones sólo serán con propósitos ilustrativos. Si desea tener una copia de la versión final me encargaré de hacerla llegar.

## **Appendix 2: Interview Questions Examples**

### **Stakeholder Interview: Test-designer**

1. ¿Podría contarme un poco sobre su perfil profesional y su historia en este campo?
2. ¿En qué procesos evaluativos en Colombia ha participado y cuál ha sido su función?
3. ¿Cuál es su marco referencial para diseñar evaluaciones? ¿Cambia en algo si la evaluación es para el contexto colombiano?

Si ¿De qué manera cree que las políticas del país, los objetivos de aprendizaje, las guías y mallas curriculares sugeridas influyen su proceso de diseño de evaluaciones?

Según respuesta a pregunta anterior:

¿Cuál es la importancia del MCER en los procesos de diseño de instrumentos de evaluación?

4. ¿Cómo cree que las políticas lingüísticas en materia de bilingüismo y lenguas extranjeras influyen la manera en que se enseña y evalúa en el país?
5. ¿Cree que es posible o apropiado evaluar con un único instrumento las competencias lingüísticas de los estudiantes a lo largo del país?

Si sí,

6. ¿Cuál es su opinión respecto a las pruebas nacionales como las pruebas Saber 11 o las pruebas Saber Pro en materia de inglés?
7. Desde su perspectiva, ¿Qué implicaciones tiene para la evaluación en lenguas el hallarse en contextos de educación bilingüe o que el sistema educativo tenga como objetivo el bilingüismo?
8. ¿Cuál es el papel del español en los procesos de valoración bilingüe en Colombia?
9. ¿De qué manera usted identifica la relación entre la lengua materna y la lengua objeto en los procesos de evaluación bilingüe en Colombia?
10. ¿Qué retos proyecta en la construcción de evaluación del bilingüismo en Colombia?

**Interview Policy Maker:**

1. ¿Podría contarme un poco sobre su perfil profesional y su historia en este campo?
2. ¿Cómo concibe el bilingüismo? ¿Qué rol cree que tiene en nuestra sociedad?
3. ¿Cuál cree que es la importancia para un país como Colombia el aprendizaje de una lengua extranjera como el inglés?
4. ¿Cuál es su opinión referente a la implementación de las políticas bilingües en Colombia?
5. ¿Cuál es la relevancia de crear e implementar políticas encaminadas al aprendizaje y uso de una lengua extranjera como el inglés en el país?
6. ¿Cómo es su participación en la creación de políticas lingüísticas en torno al bilingüismo y el aprendizaje de lenguas extranjeras?
7. ¿(de acuerdo con las políticas actuales) Cómo cree que se debería evaluar el bilingüismo?



8. Desde su perspectiva, ¿cómo se configura la propuesta evaluativa de las políticas lingüísticas actuales del país?
9. ¿Cómo cree que se pueden conciliar las políticas lingüísticas nacionales con las necesidades particulares de las diferentes regiones y localidades del país?
10. ¿Cuál es la importancia del MCER en el proyecto de bilingüismo y su propuesta evaluativa?
11. (preguntar si no se toca el tema) ¿Cómo coexisten las pruebas estandarizadas e internacionales con las políticas actuales en materia de bilingüismo?
12. ¿Cómo cree que varían las necesidades educativas y lingüísticas de los estudiantes en los diferentes lugares del país?
13. ¿Cuál es el papel del español en los procesos de valoración bilingüe en Colombia?  
¿han participado en procesos de diseño de evaluación que contemplen la lengua materna de los estudiantes en el país?

**Appendix 3: Codebook**

Access to Google folder where codebook is stored:

[https://drive.google.com/drive/folders/1TwOOK7Io3TNewcGlpKpLJw9v2bUNP7Q5?usp=share\\_link](https://drive.google.com/drive/folders/1TwOOK7Io3TNewcGlpKpLJw9v2bUNP7Q5?usp=share_link)

#### **Appendix 4: Extracts in Original Language**

##### **Finding one**

*"desde la [visión] clásica lo que se espera que la persona pueda ser Igualmente competente en ambas lenguas, eso es la interpretación clásica y tradicional de bilingüismo, que pues personalmente no encuentro disonante, pues me parece que es muy claro, si alguien es bilingüe es porque tiene la competencia para desenvolverse ... para mí una persona bilingüe es una persona que es capaz de asimilar y reaccionar e interactuar con el mundo en más de una lengua, no necesariamente inglés de manera equivalente"* Int #2: TD/Ter-PB-RIS

*"También estoy familiarizado y creo en el tema del bilingüismo como un fenómeno social de idiomas o named languages en contacto que llevan a que haya procesos comunicativos entre personas de diferentes países que va obviamente atado a temas de inmigración ... Entonces los idiomas son entes vivos que hacen parte de una sociedad en donde interactuamos los humanos entonces aprender un idioma es también aprender esa parte cultural"* Int #6: Sch/Adm-PB-CAL

*"También somos hijos de ese paradigma donde pensábamos que en nuestro contexto el español de los estudiantes era un impedimento para aprender la lengua inglesa. Digamos que eso ya está cambiando un poquito y estamos pensando en cómo idear, cómo implementar estrategias pedagógicas en las cuales los estudiantes utilicen su español para aprender inglés de manera efectiva ... Cómo puedo hacer uso yo como profesor de esos recursos lingüísticos genuinos, legítimos y reales que tienen los estudiantes para evaluar, la lengua inglesa o aprenderla"* Int#6: Sch/Adm-PB-CAL

*"pero para poder lograr ese uso de la lengua Extranjera, debemos entender que el inglés no solamente es una competencia que se trabaja en un área que se llama inglés, sino que debe ser algo más transversal en donde se debe hacer un trabajo más institucional por lo general"* Int #1: PM-PB-DC

*"Tú necesitas espacios donde necesites utilizar el idioma para interactuar para aprender de algo, para cumplir una función, para cumplir con una tarea y pues es*

*lo que hemos estado tratando de hacer con los procesos de formación, de capacitación, con los materiales que hemos producido es que el inglés, más que ser objeto de estudio, se convierta en un instrumento para aprender otras cosas.” Int #1: PM-PB-DC*

*Written Policy vs Enacted Policy*

*“El Ministerio avanza en el establecimiento de un enfoque de plurilingüismo funcional que reconozca y promueva la riqueza cultural, étnica y lingüística del país con sus más 65 lenguas nativas, 2 lenguas criollas, lengua rom y lengua de señas colombiana; y que abra las puertas al desarrollo de la ciudadanía global y el intercambio cultural que aporta el aprendizaje de lenguas extranjeras, como el inglés.” Plan Sectorial 2018-2022 Min Educación*

*“hay que hacer más precisión sobre el tema de Educación bilingüe dado que educación bilingüe, pues es una amplia gama y yo solamente manejo una parte de esa amplia gama que es inglés como lengua extranjera en instituciones educativas oficiales, el resto lo que se entiende como por educación bilingüe en términos etnográficos y en términos de segundas lenguas, de lenguas criollas, de lengua de señas, eso no está a mi alcance, Pues hay unos equipos y unas entidades encargadas específicamente del tema de lenguas nativas, de lengua de señas y un área también aquí en el Ministerio que se encarga de Educación bilingüe colegios privados, es diferente ... solamente nosotros nos encargamos del Programa Nacional de Bilingüismo que yo la llamaría plurilingüismo pero nosotros nos encargamos, específicamente del tema de lenguas extranjeras en el sistema educativo oficial.” Int #1: PM-PB-DC*

*“...Es que una cosa es el bilingüismo y otra cosa es el discurso o el concepto o la visión de bilingüismo que ha venido proponiendo el gobierno nacional En compañía del Consejo británico y otras instituciones que han venido proponiendo en los*

*últimos años cierto, eso sí es otra cosa, pero eso ya es el discurso que ellos han venido promoviendo. Pero eso no significa que el bilingüismo como tal sea algo malo no, es algo deseable y algo que debemos promover”*

*“Ahora bien esta política de Colombia no explora el tema del bilingüismo como un recurso personal e incluso social ... entonces allí no se explora de cómo digamos que el enfoque de la política lingüística en Colombia no es bilingüismo, sino el aprendizaje de inglés como lengua extranjera. Porque si fuera un enfoque de bilingüismo, entonces en ese enfoque de bilingüismo dentro de la misma política se explicarían estándares de aprendizaje de los idiomas indígenas en Colombia de lengua de señas colombiana y de español” Sch/Adm-PB-CAL*

**Finding 2:**

“se pone la prueba para que refuerce la política y mucha gente lo que hace es modificar lo que está haciendo en la práctica del aula de clase para que estén alineadas con la prueba. ... Entonces, infortunadamente como te dije, se utiliza la evaluación como una herramienta para obligar a todo el mundo a reaccionar a esas políticas en vez de trabajar con los usuarios, los profesores, estudiantes, en cómo crear esas políticas y cuáles políticas serían ideales, entonces infortunadamente se recurrió a esa práctica que ha sido criticada por muchos años.” Int #3: Sch/TD-PV-DC

*“Entonces digamos que en los Colegios eso [la influencia de las políticas] es muy fuerte, se ve tan fuerte que por ejemplo usted le dice a muchos maestros les dice, venga, vamos a enseñar de otra manera. Vamos a integrar las cuatro habilidades. Vamos a poner a los muchachos y las muchachas a que hablen más, hagamos esto un poco más comunicativo, no enseñamos tanta gramática vocabulario comprensión de lectura y le dicen a uno pero es que si eso es lo que van a evaluar, yo, para qué les voy a enseñar lo otro y si Yo no Les enseñó eso y les va mal por*

*ponerme a hacer con ellos conversaciones y sale y les va mal y entonces a la persona que le van a echar la culpa es a mí acá en el colegio.” Int #5: Sch/Adm-PB-ANT*

*“El colegio, como te conté, quedó en número uno en la ciudad de Dosquebradas, y ese es uno de los objetivos que siempre tienen los administradores y el rector. Entonces nosotros, al menos desde que estoy yo en la institución, se ha intentado que se empiece un proceso al menos desde temprana edad con los niños de que reconozcan cómo sería el formato de ese tipo de evaluaciones. La influencia es total, o sea, todo lo que te mencionaba ahorita sobre las pruebas bimestrales, aunque se hacen individuales, nos basamos en el formato del ICFES, sí. Nos esforzamos para que los niños desde muy pequeños tengan ese conocimiento [de toma de exámenes] y puedan presentarlas y sepan qué hacer en el momento de que tengan que presentar [el examen].” Int #4: Ter-PV-RIS*

*“En el colegio damos cuatro horas de inglés y como es un colegio enfocado a las pruebas hay que ver una hora de Pruebas Saber 11 de inglés, cierto? ... En cuando a la evaluación obviamente todos es muy procedimental, sin embargo, obligatoriamente y por la modalidad del colegio, al final de cada periodo se tiene que dar un tipo de pruebas Saber, o sea, una prueba con opción múltiple cierto, todos los niños desde sexto hasta 11 tienen ese tipo de prueba de evaluación” Int #7: Ter-PB-RIS*

*“Bueno uno es el efecto que tienen las pruebas en el currículo de las instituciones públicas en el país es que hay instituciones que encarrilan o enfocan su currículo a que los estudiantes se preparen para presentar una prueba y subir indicadores en la institución parte de esas consecuencias es que haya más prestigio para las instituciones lo cual digamos es un poco irónico porque sube el prestigio tomando como base los resultados de una prueba que es limitada porque es la naturaleza de las pruebas.” Int #6: Sch/Adm-PB-CAL*

*“Be the One Challenge, que este sí me gustaría mostrartelo, para nosotros es muy importante porque en este momento ya más de 400.000 estudiantes la han*

*descargado y la han utilizado. Estamos en la versión 3.0, es una versión de una aplicación que es gamificada, y si tú ves la app y la abres, es básicamente la Prueba Saber de Inglés. Hay cuatro niveles, el estudiante escoge en qué nivel están, los cuatro niveles que son pre A1, A1, A2, B1 y cuando los niños escogen cada una de esas misiones, dentro de cada misión tiene siete sub-misiones corresponden a la prueba saber en inglés.” Int #1: PM-PB-DC*

*“[Las pruebas nacionales] se utilizan digamos que esto es grave, en serio esto es grave, que se utilizan los resultados de las pruebas para evaluar, la calidad de los docentes, las pruebas están diseñadas ... para un propósito específico las pruebas estandarizadas en Colombia son para medir una parte de los conocimientos de los estudiantes en el país, no para medir competencias docentes esas no es el propósito. Ya que las personas y las instituciones o las personas en las instituciones lo utilicen para este propósito digamos que no hay conocimiento suficiente para entender, que eso es un grave error.” INT #6: Sch/Adm-PB-CAL*

### **Finding 3**

*“El año pasado quedamos de primeros en dos quebradas en las pruebas saber 11 y a nivel de Risaralda quedamos terceros. Entonces tiene muy buen nivel académico.”*

*Int #4: Ter-PV-RIS*

*“Lo que procuramos hacer es que los currículos estén actualizados y alineados a diferentes estándares internacionales.”*

*“Finalmente esos datos [resultados Saber Pro] se vuelven importantes cuando la lógica es hablar de resultados de aprendizaje. ¿Entonces qué va a pasar? que si a uno le va mal la próxima vez no le dan la acreditación de alta calidad a la universidad. ¿Qué va a pasar? Que la administración universitaria exija a todo el mundo que le presten atención a esas pruebas, que se preparen los pelados para esas pruebas, que hay que mejorar en esas pruebas. Sí me hago entender? Entonces eso finalmente sí tiene un efecto.”*

*"Algunas instituciones educativas por un contexto particular lo están apuntando a un B2, pero son instituciones que son a veces más técnicas, en temas relacionados con turismo o tienen de pronto un técnico de servicio al cliente y entonces tienen una alianza interinstitucional en un call center y ese call center los apoya, financia, brinda becas, bueno, oportunidades para los estudiantes de educación media, principalmente, entonces hay unos colegios a nivel nacional que lo están apuntando al nivel B2.*

*"Nosotros quisimos generar otro Banco Nacional de evaluaciones, que lo puedes encontrar en eco web donde se evalúan todas las competencias, todas las habilidades comunicativas de los estudiantes y hay pruebas de entrada y de salida, están para los niveles A1 A2 y B1, y los maestros lo pueden descargar, y de hecho, pues es uno de los recursos que más se descarga a nivel nacional, porque quisimos tener también pruebas estandarizadas, estas las desarrollamos con una aplicación a través del British Council, y pues son unas pruebas que están calibradas para poder comprobar el nivel de los estudiantes, según las cuatro habilidades básicas de la comunicación" Int #1: PM-PB-DC*

*"mira que lo primero que hizo el ministerio de educación nacional fue pagarle bastante plata al British Council para que evaluara a los profesores, cuando todo el mundo sin siquiera hacer el examen se supo la respuesta sabía la respuesta: los profesores no tienen el nivel. Y se hizo una inversión en comprar y aplicar los exámenes, se hizo y "uh, sorpresa, todos tienen nivel bajo" en vez de invertir en otras cosas." Int #3: Sch/TD-PV-DC*

*"entendemos que son miradas muy parcializadas, muy marcadas por unas agendas económicas políticas coloniales que no nosotros no compartimos." Int #5:*

*Sch/Adm-PB-ANT*

*"las certificaciones tienen mucho que ver con ajustarse a una serie de protocolos Internacionales y pues invitar a esas instituciones internacionales a que revisen*



*pero pues esa invitación representa costos y costos altísimos y la universidad pública al estarle ofreciendo esto a una cantidad de estudiantes masiva no tiene esos recursos, o sea, tiene que haber una voluntad política grande para hacerlo porque pues certificarse es costoso y en cualquier campo, una certificación internacional muchísimo más costosa.” Int #2: TD/Ter-PB-RIS*

*“Pues básicamente el Marco Común Europeo es la base de los estándares básicos de competencia y los estándares básicos de competencia son la base del currículo sugerido de inglés y por ende es el origen de muchas cosas.”*

*“el problema es querer volver todo el marco, o sea, todo tiene que ser El Marco. Es decir, usted ya no puede pensar un curso sin pensar en el marco, usted ya no puede pensar un examen sin pensar en el marco, un texto guía sin pensar en el marco.” Int #5: Sch/Adm-PB-ANT*

#### **Finding 4**

*“Uno de los principales aspectos que hemos tratado de hacer desde el Ministerio es, uno, desmitificar el concepto de evaluación que existe actualmente en el sistema educativo. Evaluar, no solamente es un examen. Y entonces hemos estado generando unas herramientas y unas estrategias para poder hacer que la evaluación sea más un proceso, y que haga parte del proceso de formación del estudiante.” INT #1: PM-PB-DC*

*“De igual forma, se ponen a consideración herramientas para que los docentes cuenten con instrumentos que fortalezcan sus estrategias de diagnóstico y seguimiento, y para que los estudiantes sean partícipes activos de estos procesos.”*  
Plan Sectorial Mineducación 2018-2022

*“Entonces lo que hacen las pruebas, en cualquier nivel, masiva o en clase, son limitadas en cuanto a los datos que ellas recogen. Entonces digamos, ya a nivel más educativo, más del aula de clases, los docentes somos llamados a recoger datos sobre la competencia comunicativa de nuestros estudiantes desde diferentes puntos*

*de vista. ... Entonces yo puedo hacer una prueba formal donde recojo un poco de información, utilizo un portafolio donde recojo otro poco de información y hago una autoevaluación donde mis estudiantes me dan otro poco de información. Eso es lo que se llama la perspectiva del multiplismo.” INT#6: SCH/ADM-PB-CAL*

*“Como te digo, le damos mucha importancia cuando hablamos de bilingüismo a la lengua extranjera y no nos podemos olvidar de la lengua materna, sobre todo evaluarla. De ese proceso del uso del lenguaje muchos de esas habilidades o competencias se pueden transferir a la segunda lengua” INT #3: SCH/TD-PV-DC*

*“Pero digamos como la tendencia grande, que seguirá habiendo es el uso de la evaluación por tareas para activar la competencia de lengua de los estudiantes. O sea que no es evaluar el bilingüismo, sino lo que se puede hacer con el bilingüismo ... Si el enfoque no es filosofía, sino inglés, como lengua extranjera ¿Cómo puedo yo crear, pensar estrategias pedagógicas donde yo pueda evaluar en inglés como lengua extranjera activando el bilingüismo de mis estudiantes? Es decir, si yo quiero evaluar un aspecto la comprensión de escucha de mis estudiantes en inglés, ¿Cómo puedo diseñar un instrumento de evaluación que tome digamos el español de mis estudiantes para poder yo mirar y evaluar su comprensión de escucha en inglés? Entonces digamos bilingüismo como un recurso en el área de aprendizaje de inglés como lengua extranjera.” INT: #6 SCH/ADM-PB-CAL*

*“Creería yo que el gobierno nacional sí de verdad quisiera tener (De nuevo falta la voluntad política) una verdadera voluntad política tendría que ser una prueba que juntara las cuatro habilidades y que tal vez utilizando la tecnología, como se dice esto, la tecnología, sé que suena loco, pero no creo que lo sea de aquí a poco, Inteligencia Artificial para evaluar, al menos dentro de un rango la competencia de habla y la competencia de escritura y la de listening.” INT#2: TD-TER-PB-RIS*

*“se puede utilizar más simulaciones, por ejemplo que estás interactuando con un avatar. ... Entonces de alguna manera permite evaluar de una manera más apropiada todas esas habilidades lingüísticas que antes con lápiz y papel ... hoy en*

*día la realidad es que lo llaman multimodality en inglés que es que tú lees no solamente diferentes medios escritos en internet, sino que también es videos, miras tablas erráticas, etcétera.” INT #3: SCH/TD-PV-DC*

*“Be The 1 Challenge: aplicación digital que busca fortalecer los aprendizajes en inglés, cerrar las brechas y generar una cultura de mejoramiento continuo en el aprendizaje del inglés.” Plan sectorial Mineducación 2018-2022*

*“Todavía hay un camino larguísimo yo creo que con este jueguito (be the one challenge), como muchos profesores lo llaman, hemos estado desmitificando un poco lo que se entiende por evaluación, que no necesariamente tiene que ser un proceso donde hay estrés y el estudiante se tiene que sentar frente a un examen y entra a jugar como todo el tema de la presión y el tema socioemocional para que demuestre los resultados, sino que también se puede dar a través de este tipo de herramientas que aportan algunos datos para que el maestro [y las secretarías] esté constantemente monitoreando qué está pasando con el aprendizaje de los estudiantes.” INT #1: PM-PB-DC*

“La estrategia “Supérate con el Saber” cumple un papel importante. En esta estrategia participan, mediante aplicación en línea o por fuera de línea, estudiantes de establecimientos educativos oficiales y no oficiales matriculados en los grados 2° a 11°. ... Supérate con el Saber les permite a las entidades territoriales certificadas monitorear los resultados y avances de sus instituciones y redireccionar los recursos que hoy invierten en entrenamiento para las Pruebas de Estado, en particular de Saber 11°, hacia acciones que favorezcan aprendizajes sostenibles, menos orientadas al entrenamiento para responder a ellas.” Plan Sectorial Mineducación 2018-2022

### **Finding 5**

“Asuntos como la evaluación, eso está muy normalizado ya, o sea, hay normas para todo. ¿Cuáles son las pruebas que se aceptan? ¿Cuáles no? ¿Cuáles son los niveles exigidos? ¿Por qué esas pruebas? ¿En dónde se tienen que presentar esas pruebas?

¿Cuáles son los puntajes, la relación de los puntajes con el Marco [Común Europeo]? todo eso ya está normalizado. ... De hecho recibimos muchas críticas porque a veces dicen pero si usted es la universidad de Antioquia, usted tiene autonomía Universitaria. ¿Usted por qué le exige, por ejemplo a un docente que tiene que tener un [certificado] C1 para ser formador de maestros acá en la licenciatura? Hombre, porque eso es lo que exige el gobierno.”

*Int #5: Sch/Adm-PB-ANT*

“Entonces era una forma como de obligar a los colegios a reaccionar, como bueno, “aquí hay una prueba y una política lingüística, ustedes verán lo que hacen.” *Int #3: Sch/TD-PV-DC*

“Ambas [directoras] han estado muy comprometidas con que el Instituto sea un jugador activo de las estrategias de bilingüismo regionales. Entonces el Instituto sí participa de las mesas del bilingüismo regionales y también participa ... en las discusiones que se llevan a cabo en la Universidad, en el municipio y pues siempre ha tratado de estar inmerso y alineado a esas políticas. Ha participado en varios de sus proyectos de bilingüismo, en las de Colombia Bilingüe Por ejemplo, yo hice parte de varios de esos proyectos tanto en la formulación como la ejecución.” *Int #2:*

*TD/Ter-PB-RIS*

“creamos una política institucional, una política que nos ayude a ver para dónde vamos en términos de lengua extranjera.” *INT #8: ADM-PV-VAC*

“Si hay algo que caracteriza un espacio como este (la universidad) es la diversidad y unos profesores hacen una cosa y otros hacen otra, o sea, aquí no hay uniformidad, pero esa es parte de la riqueza de las instituciones, simplemente. Por dictatorial que pueda ser un rector es muy difícil que controle absolutamente el quehacer de los docentes. Sí, o al menos en un país como Colombia no lo hace, quizá eso es posible

*que en otros países, sí, pero en un país como Colombia es muy difícil.” Int #5:*

Sch/Adm-PB-ANT

*“Pues en términos generales nosotros como universidad nos hemos desmarcado de esos discursos oficiales en torno al bilingüismo. No compartimos esa mirada. Y usualmente hemos tenido una posición crítica, y en nuestras acciones hemos marcado una diferencia con esos discursos oficiales, no creemos en ellos. Pero eso no significa ... que todo lo que el gobierno hace es malo, no.”*