

An empirical validation study of Spolsky's language policy model

Victoria Van Oss, Piet Van Avermaet, Esli Struys & Wendelien Vantieghem

To cite this article: Victoria Van Oss, Piet Van Avermaet, Esli Struys & Wendelien Vantieghem (2021): An empirical validation study of Spolsky's language policy model, Current Issues in Language Planning, DOI: [10.1080/14664208.2021.1907060](https://doi.org/10.1080/14664208.2021.1907060)

To link to this article: <https://doi.org/10.1080/14664208.2021.1907060>



Published online: 07 Apr 2021.



Submit your article to this journal [↗](#)




View related articles [↗](#)



View Crossmark data [↗](#)



An empirical validation study of Spolsky's language policy model

Victoria Van Oss ^a, Piet Van Avermaet ^a, Esli Struys ^b and Wendelien Vantieghem ^a

^aCentre for Diversity and Learning, Ghent University, Ghent, Belgium; ^bCentre for Linguistics, Vrije Universiteit Brussel, Brussels, Belgium

ABSTRACT

This study investigates Spolsky's conception of language policy as the amalgam of language beliefs, language practices and language management. To the best of our knowledge, no quantitative evidence has ever been produced for his theoretical model. The current paper addressed this void, albeit specifically in the domain of Early Childhood Care and Education – scarce in terms of language policy research, but vital for children's (multilingual) language development. An Exploratory Factor Analysis yielded three valid and reliable scales that capture language policy from the perspective of early childhood professionals ($N=305$) associated with infant welfare clinics in the Flemish Community of Belgium. Furthermore, a Confirmatory Factor Analysis indicated that the three components are 'independent' as well as 'interconnected,' thereby quantitatively corroborating Spolsky's theory. We conclude that language policy is no unitary concept; therefore, if one intends to study language policy holistically, the three indispensable components are to be identified. As a corollary, one simply cannot investigate management without considering beliefs and practices, and vice versa.

ARTICLE HISTORY

Received 30 October 2020
Accepted 19 March 2021

KEYWORDS

Language policy;
multilingualism; early
childhood; health care;
linguistic diversity

Introduction

Today's multilingual reality, engendered by the ever-growing mobility across the globe, is reflected in all spheres of modern society. Against this backdrop, research into language policy (LP) has been steadily gaining momentum. According to Spolsky (2004), LP can be defined as the amalgam of 'three interrelated but independently describable components: practices, beliefs, and management' (Spolsky, 2007, pp. 3–4). Language practices are the language behaviour of individuals, or, 'what people actually do.' Language beliefs relate to what people think is 'appropriate or desirable' language use and language management refers to people's attempts to modify the practices or beliefs of members of a speech community (Spolsky, 2017). A manifestation of language management is *advice giving* – an outside advocate's effort to (un)intentionally change the beliefs and practices of an individual or family (Spolsky, 2007). A pertinent example is the New Zealand early childhood

CONTACT Victoria Van Oss  victoria.vanoss@ugent.be  Centre for Diversity and Learning, Ghent University, Ghent, Belgium

workforce, whose advice to families contributed to the shift from Māori to English in the 1960s – which, twenty years later, was to be corrected by the immersion preschool programmes (Spolsky, personal communication, December 22, 2019).

While LP has been studied in multifarious domains, one domain that received surprisingly little attention is Early Childhood Care and Education (ECCE) (Caporal-Ebersold, 2018, p. 57). The broad domain of ECCE comprises childcare, social welfare, child protection and health care for (expectant) parents, neonates and young children (Marope & Kaga, 2015). Nevertheless, the linguistic diversity inherent in early childhood settings poses challenges for its workforce (Chan, 2011). One instance of these challenges is that despite the overwhelming evidence for the importance of home language development for multilingual children (Cummins, 1991; Hélot, 2010; Verdon et al., 2014) and the far-reaching repercussions of home language neglect (Couëtoux-Jungman et al., 2010; De Houwer, 2017), ECCE professionals feel uncertain about how to promote home languages (Kirsch & Aleksic, 2018, p. 148). Be that as it may, ECCE professionals could be instrumental in inducing children's home language maintenance during those critical years of language acquisition (Rodríguez, 2015; Verdon et al., 2014). After all, ECCE professionals' language management – in the form of advice to parents – can be a determining factor in the development of children's language skills – or lack thereof – as the advice may affect multilingual families' language beliefs, practices and decisions apropos the language(s) in which they want to raise their children. Hence, professionals' Early Childhood Language Policy (henceforth ECLP) vis-à-vis the multilingual family – warrants thorough investigation.

To the best of our knowledge, however, no quantitative evidence has ever been provided for Spolsky's LP model – not in ECCE, nor in any other domain for that matter. The primary purpose of this research is therefore to explore whether language policy can be decomposed into three interrelated but independently describable components – as advanced by Spolsky – albeit applied to an ECCE setting, and viewed from a grassroots perspective. More concretely, we aimed to verify whether advice giving as a language managerial act is related to ECCE professionals' beliefs and practices with regard to children's home languages. To this end, the first aim of this study was to develop three valid scales measuring professionals' ECLP. The second aim was to test Spolsky's theory and establish whether the constructs are indeed 'independent' as well as 'interconnected' (2019, p. 323) by means of an evaluation of construct validity. The creation of these measures may be useful for future research to investigate mechanisms among a larger population.

Previous research on Spolsky's model

LP is a well-documented (Johnson, 2013; Shohamy, 2006) and flourishing field of study. Amidst the manifold approaches, we rely on two influential yet compatible frameworks. To begin with, we recognize that language policy can occur on different layers (i.e. micro, meso, macro) that permeate and interact with one another and which coalesce into an 'onion' – an apt analogy drawn by Ricento and Hornberger (1996). From this *multi-layered* perspective, the current paper regards LP as a *multidimensional* construct by investigating the dynamics in Spolsky's (2004) triangular model. As such, the inter-relationship between language beliefs, practices and management has been the focal point of a number of studies in disparate spheres.

In the national domain, Kristinsson (2012) reviewed one hundred years (from 1860 to 1960) of Iceland's implicit 'academy-free' language policy. He found that, until the establishment of a central language council in the 1960s – with the aim of consolidating the nation's language management – language beliefs, practices and management were interconnected. Björkman (2014) applied Spolsky's trilogy to the school domain. His analysis of language policy documents from Swedish universities showed interdependence between language beliefs (i.e. the worry that Swedish is losing ground due to the widespread use of English), language management (i.e. to protect and promote Swedish) and descriptions of language practices. Nevertheless, while the LP constituents have been identified in the scholarly realms in a variety of ways, the interrelationships among the components manifest quite differently depending on the research context. Angouri (2013) and Kingsley (2009), for instance, examined LP in the workplace; the former in multinational companies in Europe and the latter in international banks in Luxembourg – two loci with English as the formal working language. Their mixed methods research revealed some discrepancy between language management and employees' beliefs and practices, as staff adopted a multilingual 'flexible' approach when communicating at work, despite official recommendations stipulating the inverse. Consonance between the three pillars in the family domain does not seem evident either. Chatzidaki and Maligkoudi (2013) interviewed Albanian immigrant parents in Greece about their family languages policies (FLPs) in relation to ethnic language maintenance. They found that despite parents' positive beliefs about Albanian, the majority did not engage in language management efforts to transmit Albanian to their children. Mirvahedi and Jafari (2021) reported similar results; their interviews with bilingual parents in Iran revealed that despite parents' ideological attachment to Azerbaijani (i.e. their ethnic language), Farsi (i.e. the official language) was omnipresent at home and management efforts to alter this were in short supply. When we turn our gaze to the paucity of studies investigating LP in ECCE; it is noticeable that these are invariably conducted in childcare facilities, while other sectors (e.g. health, child protection, social welfare) seem absent from the literature. Bergroth and Palviainen (2017), for example, investigated child agency in the interplay between Finland's official monolingual language policy and its education policy (i.e. acknowledgement of the bilingual child's right to make use of both languages). They monitored bilingual children at Swedish-medium preschools in Finland and found that all children played an active role in the construction of both policies. They argue that the two policies (i.e. management) and perceptions of these policies (i.e. beliefs) were intertwined with language practices. Another study, by Caporal-Ebersold (2018), explored LP in a bilingual crèche in France, which operates a One Professional-One Language (OPOL) language policy to support children's bilingual development. Her ethnographic case study showed interrelatedness among the LP components as the crèche's official LP (i.e. management), and parents and professionals' language practices, were formed by their language beliefs.

On the whole, we found that whereas some – but definitely not all – studies have demonstrated the 'interrelatedness' between the LP components, it appears that the second part of Spolsky's definition has been somewhat overlooked, for no studies have concomitantly substantiated that the components are also 'independent.' While these – preponderantly qualitative – studies yielded enlightening insights into LP, it remains unclear – due to the lack of quantitative instruments – how *closely* the components

are connected and how *widespread* certain language beliefs, practices and management acts are; a void this paper aims to fill, in the hitherto neglected domain of ECCE.

Method

Research context

The locus of this project is the Flemish Community in Belgium – an entity of the federal state comprising all inhabitants from the officially monolingual Flemish Region and the Dutch-speaking inhabitants of the officially bilingual (Dutch-French) Brussels-Capital Region – which has legal responsibility for matters including welfare and education. The Flemish Community's language policy – marked by the preeminence of Dutch – can be illuminated by several factors; one of which is the fight for the recognition of Dutch (against French) as an official language in Flanders and Brussels in the nineteenth and twentieth century which triggered the prevailing notion that Dutch is in jeopardy and that, accordingly, it has to be preserved as the majority language in the Flemish Community (Mettewie & Mensel, 2020). Despite this official strict monolingual policy with scant attention for languages other than Dutch, its linguistic landscape is nevertheless increasingly heterogeneous. In 30% of the children born in 2019 the language use between mother and child was not the majority language (i.e. Dutch). Other languages most widely spoken in home environments are French (5.9%), Turkish (4.3%) and Arabic (2.7%) (Kind en Gezin, 2019).

The ECCE context for this study is Flemish infant welfare clinics. The institutions in question offer medical, psychosocial and educational prevention to parents with newborns and children up to three years old. The *raison d'être* of the infant welfare clinic is to create the necessary conditions to maximize young children's development opportunities by offering family support, entirely free of charge. Their outreach is therefore remarkably broad; of all babies born in Flanders in 2018, no less than 90% paid a visit to the infant welfare clinic ('Child and Family', n.d.). The clinics provide health care (e.g. administering of vaccines, monitoring child development, etc.) as well as (multilingual) parenting advice. The fact that these professionals reach nearly all families in the Flemish Community entails that their potential impact on multilingual parents' FLP is by no means marginal – and therefore highly interesting for our research. As for the clinics' LP; we should first mention that these facilities are under the auspices of 'Child & Family,' an agency of the Flemish government. However, in contrast to other domains of Flemish language policy, professionals in infant welfare clinics are allowed to use other languages than Dutch in their communication with non-Dutch-speaking parents, which is justified by the specific need to provide health care to vulnerable families (Van Gorp & Moons, 2014). In addition to such leeway, the Child & Family agency promulgates a multilingual LP for early childhood facilities and emphasizes the importance of home language development for young children (2012). This openness to linguistic diversity is unparalleled in other domains of Flemish language policy, such as education or government communication targeted at adults.

Survey development

The first step towards the creation of ECLP measures was the development of a questionnaire. This process was carried out as part of the research project 'Pro-M,' which was

launched in 2018 with the objective of promoting early multilingualism in childhood and childcare. Item generation was the fruit of deductive and inductive strategies (Kapusinski & Masters, 2010). The former involved a literature review and examination of pre-existing scales, whereas the latter included in-depth discussions between the project members (i.e. five policy officers and nine academics). The initial item pools were to reflect the desired ECLP constructs; viz. professionals' *beliefs* about the value of home languages, their language *practices* at the clinic, and their language *management* – in casu operationalized as advice on multilingual parenting. Face validity was evaluated by target population judges to verify whether the instrument was appropriate for the intended purpose (Morgado et al., 2018). Three professionals – one volunteer, one nurse and one district team manager – pilot-tested the survey and their subjective judgement was taken into consideration (e.g. withdrawal of ambiguous and formal language). In the penultimate stage an editor was hired to perform item redaction and to fine-tune the questionnaire.

Respondents and data collection

Data gathering commenced with a random sample selection by virtue of a stratified two-stage sampling method. In the first stage, nineteen municipalities located in Flanders and Brussels were selected based on the high prevalence of mothers who do not speak the dominant language (i.e. Dutch) to their child. In the second stage, one to two institutions were selected per municipality (depending on the population size) resulting in a sample of 34 institutions. The latter's staff were then invited to participate in our survey. Five teams declined and 29 facilities, including 337 professionals, consented – thus representing 10% of the 302 Flemish infant welfare clinics in Belgium. Data collection ran from March until August 2019, during which professionals with a coordinating function (32 from the 337 respondents) completed a survey comprising only management related questions. These 32 questionnaires were omitted from the dataset. Our final sample ($N = 305$) thus consists of 74 nurses (24%), 73 physicians (24%), 21 family supporters (7%) and 137 volunteers (45%), of which the vast majority are female (94%). The gender-skewed sample distribution does not undermine generalization of our findings, for at the time of the data collection virtually all (91%) of the employees working for the Child and Family agency were women ('Child and Family', n.d.). The mean age of the respondents was relatively high ($M = 52.58$, $SD = 15.96$) and ranged from 23 up to 80 years old. This is due to the high average age of volunteers working in the clinics, who are primarily retired women. Ensuing administration and collection, the authors embarked on a quest to find three suitable scales by dint of rigorous psychometric analyses as well as theoretical rationales.

Analytical plan

In view of the first aim of this paper – the development of three scales – we adopted a stepwise approach consisting of five phases. Firstly, response distributions of the individual items were inspected. Self-reports are known to elicit response bias (Paulhus, 2017); therefore, all items exhibiting limited variability were identified and disposed of (Clark & Watson, 1995). With regard to the batteries of binary items assessing language beliefs and management, if more than 85% of respondents either endorsed or denied the statements

administered, a ceiling or floor effect was considered present (Bot et al., 2003). Conversely, the battery of potential 'language practices' comprised items that were presented in a five-point frequency Likert rating format; hence, if more than 85% of participants responded identically to the items (e.g. if the vast majority ticked 'sometimes'), they were excluded from the item pool. In the second step, an Exploratory Factor Analysis (EFA) was conducted (in SPSS Statistics, V26.0) to ensure unidimensionality of each scale. More specifically, a Principal Component Analysis (PCA) with orthogonal rotation (Varimax) was performed to evaluate each of the angles' internal structure. In the third step, a screening of the factor loadings of the items in the rotated solution was carried out. Indicators which loaded above .30 were retained, a minimum acceptable cutoff point that is recommended for a sample size of 300 respondents (Field, 2009). Furthermore, only factors with a minimum of three items with factor loadings above the cutoff point were retained in view of the subsequent factor analysis which necessitates at least three items per factor to ensure minimum coverage of the theoretical domain and to properly identify the construct (Hair et al., 2014, p. 608). Finally, all items in the retained factor solution which cross-loaded above .30 were eliminated as these items are affected by more than one factor and were thus deemed too intricate (Worthington & Whittaker, 2006). Next, internal consistency reliability was assessed by computing Cronbach's alphas and corrected item-total correlations (CITC). Whilst the acceptable level of reliability remains a moot point, an alpha value below .60 is considered questionable and above .75 demonstrating good reliability (Field, 2009). As regards CITC, which is sometimes considered a more straightforward measure of internal consistency (Clark & Watson, 1995), items with coefficients below .30 and above .80 were excluded as the former suggests a lack of homogeneity (Viswanathan, 2005) whereas the latter is indicative of multicollinearity.

To fulfil the second aim of this study – testing Spolsky's theory – we conducted a Confirmatory Factor analysis (CFA) (using SPSS Amos, V22.0). To assess how well a measurement model fits the empirical data, a plethora of fit indices is available. We report at least one incremental index and two absolute fit indices to provide sufficient evidence of model fit (Hair et al., 2014). Comparative Fit Index (CFI) values above .90 and Root Mean Square Errors of Approximation (RMSEA) and Standardized Root Mean Square Residual (SRMR) values smaller than .08 are commonly acceptable cutoff criteria and were thus applied (Hooper et al., 2008). Chi-Square values along with their *p* values and degrees of freedom are reported but disregarded, due to limitations of the Chi-square test (Hooper et al., 2008).

Following the assessment of the overall fit of our measurement model, the second objective of the CFA entailed an evaluation of construct validity. Therefore, three validity aspects were explored including convergent, divergent and nomological validity. Convergent validity (i.e. whether the items are related to the latent construct they are theoretically predicted to be) was examined by employing two methods. Firstly, factor loadings of the items were inspected, thereby eliminating standardized loading estimates that were not statistically significant. In addition, all factor loadings below the cutoff value of .50 were removed (Hair et al., 2014), provided that contraction did not have a detrimental effect on the construct in its entirety. Secondly, construct reliability (CR) was assessed, thereby accepting all values above .70. Subsequently, divergent validity (i.e. the extent to which the constructs are unequivocally different from one another) was examined

by comparing the original three-factor-model with a unidimensional model in which the same items load directly onto one latent construct. If the original measurement model yields a better fit, evidence for divergence is supported since this finding implies that the items represent different constructs (Hair et al., 2014, p. 619). Thirdly, we sought evidence for nomological validity by exploring the model's nomological network (Cronbach & Meehl, 1955). The latter was established if the constructs in the measurement model related to one another in a theoretically consistent manner and if this interrelatedness was in line with Spolsky's theory. Finally, we relied on the modification indices to ascertain whether the model could be further improved.

Results

Development and validation of three scales measuring Spolsky's triad

In this section, we first portray the methods that were employed to develop three scales measuring ECLP from the professional's vantage point. In the second part, we cast light on how we established construct validity of the scales that emerged in the EFA and we elucidate how the vertices of the language policy triangle are related.

Language management

As previously stated, the management component of the model was operationalized as 'language advice' of the ECCE professional to the multilingual family. The administered questionnaire included twelve statements formulated as recommendations to parents with regard to multilingual parenting. Professionals were asked what kind of advice they offer families by ticking all the boxes that apply to them (in an 'agree/disagree' format). A first inspection of the variability in the response scores led to the removal of one item which exhibited a floor effect, as 91.6% of respondents disagreed with the statement 'I advise multilingual parents to only speak Dutch to their child.' Next, PCA was performed on eleven items. The KMO value was .83, exceeding the recommended value of .50 (Field, 2009) and Bartlett's Test of Sphericity displayed statistical significance (Chi-square = 833.90, $df = 55$, $p < .001$). We then relied on the Kaiser criterion and inspection of the screeplot to identify the number of significant components (Yeomans & Golder, 1982). Two components with eigenvalues greater than one were generated. Factors I and II yielded eigenvalues of 3.50 and 2.35, respectively, together accounting for 53.19% of the variance. The first factor contained six items with factor loadings $>.30$ and the second factor five items. Table 1 portrays the rotated factor solution. For clarity of presentation, factor loadings that did not meet the standard cutoff point were excluded from this table. Subsequently, Cronbach's alphas were computed to examine whether the pools of items exhibited coherence. Two items were omitted from factor I to enhance internal reliability, thereby yielding an alpha coefficient of .86. The set of items of factor II obtained a Cronbach's alpha of .70. The corrected item-total correlations of the items in both factors exceeded the threshold of .30, suggesting good item consistency. Scrutiny of the psychometric properties of both potential scales suggested opting for factor I, yielding the highest alpha value (.86 $>$.70), and the juxtaposition of factor contents only confirmed this tendency. Some extent of subjectivity was required to label the factors based on our interpretation of the items (Morgado

Table 1. Loadings of items measuring management on the two rotated factors.

Indicators	F I	F II
I advise multilingual parents to read in their home language to their child	.86	
I advise multilingual parents to read in Dutch to their child		.66
I advise multilingual parents to speak Dutch and their home language interchangeably, but to consciously plan which language is used and when.		.74
I advise multilingual parents to let one parent speak one language to the child, and the other parent the other language	.68	
I advise multilingual parents to only speak their home language with their child	.53	
I advise multilingual parents to speak the language they know best with their child	.80	
I advise multilingual parents to speak Dutch and their home language interchangeably, but to consciously plan which language is spoken depending on the moment		.70
I advise multilingual parents to mix Dutch and their home language when speaking to the child, but to consciously plan which language is spoken depending on the location		.70
I advise multilingual parents to speak Dutch and the home language interchangeably		.56
I give multilingual parents tailored advice on language and multilingual parenting	.77	
I spontaneously give multilingual parents advice on language and multilingual parenting and indicate what is best	.80	
Eigenvalues	3.50	2.35
% variance explained	31.83	21.37

Note: Factor loadings from PCA. Factors with a minimum of three items with loadings $>.30$ are listed.

et al., 2018). It seemed suitable to label factor I as ‘management towards the home language’ and factor II as ‘management towards the majority language.’ The latter factor included items such as ‘I advise multilingual parents to read in Dutch to their child’ and ‘I advise multilingual parents to speak Dutch and the home language interchangeably’ and accentuated the use of the dominant language (i.e. Dutch). In stark contrast, the factor summarized as ‘management towards the home language’ most closely resembled the theoretical construct we envisaged. That is, the statements explicitly involved the home language environment, example items are ‘I advise multilingual parents to speak the language they know best with their child’ and ‘I give multilingual parents *tailored* advice on language and multilingual parenting.’ Validity is an estimation based on empirical grounds as well as theoretical rationales (Messick, 1995; Viswanathan, 2005) and a deviation from the proposed conceptual domain may result in ‘technical sophistication in isolation’ (Van Maanen et al., 2007, p. 1146). As the focus of this study is on the use of home languages, factor I labelled ‘management towards the home language,’ was found to best capture the intended construct’s content domain and was thus retained for further analysis.

Language practices

The administered questionnaire comprised twenty items measuring professionals’ language practices. First, two items exhibiting limited variability and three items which concerned only one profession in particular were removed; resulting in a battery of fifteen items on which PCA was performed. The KMO value was .79 and Bartlett’s Test of Sphericity displayed statistical significance (Chi-square = 918.24, $df = 105$, $p < .001$). Factor loadings were screened and all items that loaded below .30 were discarded. Five components with eigenvalues greater than one were generated, accounting for 61.56% of the variance. Eigenvalues were 3.78 for factor I (explaining 25.17% of the variance), 2.10 for factor II (explaining 13.97% of the variance), 1.26 for factor III (explaining 8.40% of the variance), 1.08 for factor IV (explaining 7.22% of the variance)

and 1.02 for factor V (explaining 6.80% of the variance). However, inspection of the screeplot suggested retaining only two to three factors. Moreover, factors IV and V displayed two items with factor loading above our cutoff value of $|\lambda| \geq .30$, thus failing to meet the established standard of a minimum of three items. The first three factors, however, each comprised more than three items with factor loadings $> .30$. Thus, a three-factor solution was retained (see Table 3), thereby first eliminating cross-loadings. Consequently, after removal of cross-loadings, only one item in factor III remained, and the factor was thus discarded. Cronbach's alphas were computed for the pools of items which loaded above $.30$ in factor I (number of items = 6, after removal of one item, $\alpha = .81$) and factor II (number of items = 3, $\alpha = .65$). However, factor II was somewhat of a curate's egg since its alpha value increased to $.70$ after removal of one item, but in that case only two items remained, thus failing to adhere to the three-item-rule. Nevertheless, the two remaining factors (I & II) were subjected to scrutiny. Factor I was labelled 'language use with the multilingual family,' and factor II was summarized as 'deployment of resources to facilitate the multilingual interaction.' Besides the fact that 'language use with the multilingual family' exhibited higher internal reliability than 'deployment of resources to facilitate the multilingual interaction,' content inspection of the items in factor II revealed that the 'resources' in factor II concerned multimedia resources as well as involvement of an interpreter during the service encounter. The scale was thus found to be less straightforward than factor I. Consequently, the latter factor labelled 'language use with the multilingual family' was retained for further analysis (Table 2).

Table 2. Loadings of items measuring practices on the three-factor solution.

Indicators	F I	F II	F III
I greet multilingual children in their home language	.75		
I only speak Dutch with the children	-.68		
In one-on-one situations with a multilingual child, I consciously use a few words from the child's home language	.77		
I disapprove when children speak a language other than Dutch			.66
If a child uses a word in their home language, I repeat that word in Dutch		-.36	.66
I greet multilingual parents in their home language	.78		
I speak Dutch with the multilingual parents	-.51		
I speak to multilingual parents in their home language	.68		
I make use of resources if I do not share a language with multilingual parents (e.g. image dictionary or Google Translate)		-.78	
I use pictograms or photographs in oral communication with multilingual parents		.84	
I use pictograms or photographs in written communication with multilingual parents		.73	
I ask multilingual parents to share a few words from the child's home language with me		.35	-.42
I ask for an interpreter to communicate with multilingual parents		.55	.33
Eigenvalues	3.76	2.10	1.26
% variance explained	25.17	13.97	8.40

Note: PCA factor loadings $> .30$ are listed.

Language beliefs

All items designed to assess beliefs about language use of and with the multilingual family were selected from the overall survey resulting in a battery of 34 items. Subsequently, all items with limited variability were omitted, curtailing the initial battery to 19 variables. PCA yielded a KMO value of $.72$, and Bartlett's Test of Sphericity showed statistical significance (Chi-square = 670.52, $df = 171$, $p < .001$). A screening of the items' factor loadings was carried out of which the retained factors after Varimax rotation are displayed in

Table 3. Seven components with eigenvalues greater than one accounted for 60.88% of the variance. Eigenvalues were 3.16 for factor I (explaining 16.65% of the variance), 2.40 for factor II (explaining 12.64% of the variance), 1.49 for factor III (explaining 7.86% of the variance), 1.22 for factor IV (explaining 6.42% of the variance), 1.16 for factor V (explaining 6.10% of the variance), 1.10 for factor VI (explaining 5.78% of the variance) and 1.03 for factor VII (explaining 5.43% of the variance). However, inspection of the screeplot suggested retaining only three to four components. Factor VII displayed only one item with factor loading above our cutoff value of $|\lambda| \geq .30$. The sixth factor contained three items which loaded above $.30$, although one item cross-loaded with an item from factor III, and was therefore omitted. This resulted in a factor with only two items with factor loadings $>.30$, thus failing to meet the established standard of a minimum of three items per factor. Factors V and IV each displayed two items with factor loadings $>.30$, hence they were discarded. Subsequently, reliability of the three remaining factors was examined. Cronbach's alphas were computed for the retained items in factor I (number of items = 4, $\alpha = .67$), factor II (number of items = 5, $\alpha = .64$) and factor III (number of items = 4, after removal of one item, $\alpha = .51$). Factor III showed poor reliability, so was discarded. Conversely, the alpha coefficients of the two remaining factors (I & II) suggested moderate reliability. Consequently, both factors were considered. The difference between their psychometric properties was negligible, which is why further inspection of response distributions and contents was required. The four retained items in factor I were summarized as 'benefits of multilingualism.' Most respondents, however, denied the administered statements. Only 19% of respondents endorsed item 'multilingual children have a stronger bond with their family than monolingual children' and 25% thought that 'multilingual children do better at school than monolingual children.' Interestingly, the response scores were incongruent with those of the pool with the corresponding polar opposite reverse worded items, for example 3% of the respondents agreed that monolingual children who only speak Dutch, have a stronger bond with their family than multilingual children. Accordingly, the vast majority of respondents denied the statements from both pools. These contradictory results made it difficult to draw conclusions regarding the interpretation of the

Table 3. Loadings of items measuring beliefs on three rotated factors.

Indicators	F I	F II	F III
A multilingual child is best addressed in their home language at the IWC		.72	
A multilingual child is best addressed in Dutch at the IWC		-.70	
IWCs should reflect the different realities of multilingual children		.38	
At the IWC, a multilingual child is best addressed in the language in which he or she feels safe		.63	
multilingual children do better at school than monolingual children	.62		
multilingual children make friends more easily than monolingual children	.65		
multilingual children have a stronger bond with their family than monolingual children	.73		
multilingual children have more self-confidence than monolingual children	.75		
My own language use at the IWC has a great influence on the language development of the children.			.59
I find it important to offer multilingual parents advice on multilingual parenting			-.45
It is best for multilingual parents to mix Dutch and their home language freely			.72
I find it important that multilingual parents can be assisted in their home language at the IWC.		.65	
At least one of the multilingual parents must speak Dutch to the children			.69
Eigenvalues	3.16	2.40	1.49
% variance explained	16.65	12.64	7.86

Note: PCA factors with a minimum of three items with loadings $>.30$ are listed. IWC = infant welfare clinic.

scales and might have led to an underestimation of professionals' multilingual beliefs. On the grounds of this, the 'benefits of multilingualism' factor was omitted and factor II was examined. The latter was summarized as 'beliefs about language use with the multilingual family during the service encounter' and exhibited sufficient variability. Hence, the item pool of factor II was retained for further analysis.

Corroboration of Spolsky's theory

As the proof of the pudding is in the eating – one can only produce empirical evidence for Spolsky's theory by testing it. Accordingly, in view of the second aim of this paper, a CFA was performed. We applied a maximum likelihood estimation method which required raw and complete data. We found 9.8% missing values for management, 7.5% for beliefs and 4.6% for practices, but Little's MCAR test (1988) established that the data were missing completely at random, $\chi^2 = 419.689$ (377), $p = .064$. The issue was addressed by employing a regression-based imputation technique, thereby generating a new dataset with imputed values, based on all available data for a given respondent, to replace isolated missing values (Kline, 2011). Next, management, beliefs and practices were integrated as constructs of a reflective measurement model including fourteen items. The chi-square statistic, with 74 degrees of freedom, was 171.31 ($p < .001$). The value for RMSEA was .066 and SRMR displayed a value of .057, both below the adopted standards of .08. The CFI estimate was .93, thereby surpassing the standard of .90. Ergo, the three fit indices provided support for overall acceptable model fit. The subsequent stage involved an assessment of the construct validity by investigating convergent, discriminant and nomological validity (Hair et al., 2014).

First, we verified whether the items tapping into their corresponding constructs shared a significant amount of variance, thus exhibiting convergent validity. All path estimates were statistically significant ($p < .001$), and a screening of the factor loadings revealed that the indicators showed appropriate factor loadings ($>.50$), with the only exception of one item ('infant welfare clinics should reflect the different realities of multilingual children'), with factor loading .44, which was retained in the model because the reliability of the beliefs factor decreased if the item was deleted. Thereafter, construct reliability was calculated manually for each latent construct under investigation. The CR values for management, beliefs and practices were .86, .65 and .82, respectively. Thus, all constructs exhibited acceptable to high reliability thereby suggesting that all of the individual items were measured consistently. Next, it was examined whether the three constructs positively differed from one another. To this end, the three-factor model was transformed into a unidimensional model in which all fourteen items tapped directly into a single latent construct. Comparison of the fit of the one-factor model to that of the original three-factor model demonstrated substantial discrepancy between the two model fits. In fact, the unidimensional model showed a very poor fit, χ^2 (77) = 618.796, $p < .001$; CFI = .62; RMSEA = .15; SRMR = .12, which indicates that the fourteen indicators represent three separate and unique constructs (i.e. management, practices and beliefs), thus providing evidence for divergent validity.

The correlations between the three constructs were all significant ($p < .001$), ranging from .37 between management and practices, .43 between management and beliefs, up



Figure 1. Correlations ($p < .001$) of the three sides of the ECLP triangle.

to .74 between beliefs and practices (Figure 1). As our measurement theory predicted the three measures to be positively associated with one another, a certain degree of nomological validity could be established. Ultimately, a minor and theoretically justifiable alteration was made to fine-tune the measurement model (Arnaú & Thompson, 2000). More specifically, we allowed one pair of measurement errors to covary, as suggested by the modification indices. It was indeed plausible that the error terms of ‘I greet multilingual parents in their home language’ and ‘I speak to multilingual parents in their home language’ were correlated as the formulation as well as the meaning of both items were not entirely dissimilar. Therefore, the permission of within-construct error variance posed no threat to the proposed theoretical structure. A visual diagram depicting the modified model is displayed in Figure 2. The rectification resulted in a slightly better fit with the data, $\chi^2(100) = 153.94$, $p < .001$; CFI = .94; RMSEA = .06; SRMR = .057. The latter model could have been further refined by adding more correlation error terms. Nevertheless, without solid theoretical justification, this may have led to acceptance of a new model while concealing a more meaningful theoretical structure, and is often advised against (Gerbing & Anderson, 1984; Kline, 2011). Therefore, no other (unnecessary) modifications were carried out (Table 4).

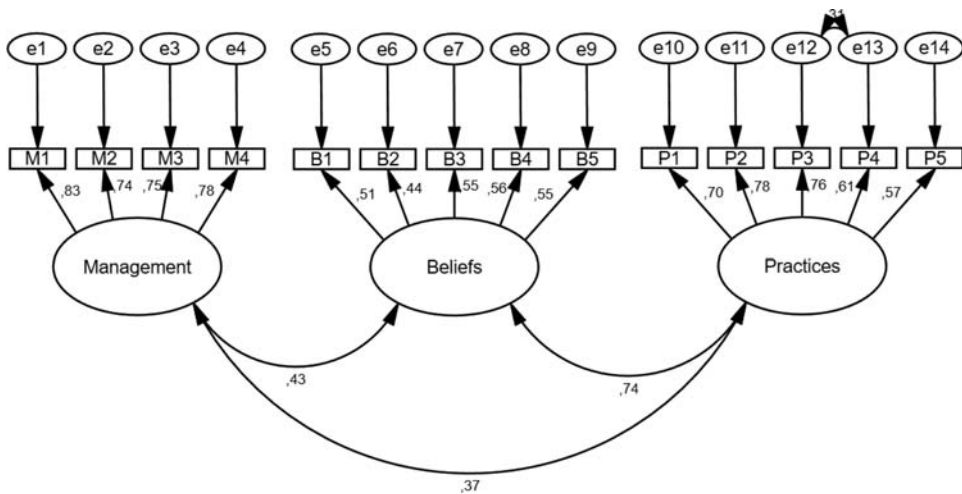


Figure 2. Final three-factor measurement model.

Table 4. Standardized factor loadings of the final measurement model.

Factor	Item contents	Factor loading
Management	M1 I advise multilingual parents to read in their home language to their child	.83
	M2 I give multilingual parents tailored advice on language and multilingual parenting	.74
	M3 I spontaneously give multilingual parents advice on language and multilingual parenting and indicate what is best	.75
	M4 I advise multilingual parents to speak the language they know best with their child	.78
Beliefs	B1 At the IWC, a multilingual child is best addressed in the language in which he or she feels safe	.51
	B2 IWCs should reflect the different realities of multilingual children	.44
	B3 I find it important that multilingual parents can be assisted in their home language at the IWC	.56
	B4 At the IWC, a multilingual child is best addressed in their home language	.56
	B5 At the IWC, a multilingual child is best addressed in Dutch *	.55
Practices	P1 I greet multilingual children in their home language	.70
	P2 In one-on-one situations with a multilingual child, I consciously use a few words from the child's home language	.78
	P3 I greet multilingual parents in their home language	.76
	P4 I speak to multilingual parents in their home language	.61
	P5 I only speak Dutch with the children *	.57

Note: $N = 305$. IWC = infant welfare clinic. * = reversed coding.

Discussion

Language policy (LP), according to Spolsky (2004), comprises three interconnected but independently describable components: language beliefs, language practices and language management. This paper challenges his claim by seeking empirical and quantitative evidence which – to date, at any rate – appears absent from the scholarly realms. We addressed this void, albeit specifically within the domain of Early Childhood Care and Education (ECCE) and viewed it from a grassroots perspective. To this end, we collected data from professionals ($N = 305$) employed at Flemish infant welfare clinics in Belgium. In this regard, the LP model was interpreted as follows: language management was operationalized as advice to parents on multilingual parenting. Language beliefs specifically concerned beliefs about children's multilingual development and language use with the family and language practices encompassed professionals' language use with multilingual families at the facility.

The purpose of this paper was twofold. First, to build three valid scales measuring professionals' Early Childhood Language Policy (ECLP). Second, to test Spolsky's theory and establish whether the constructs are indeed 'independent' as well as 'interconnected' (2019, p. 323) by means of an evaluation of construct validity. As regards the first objective, we successfully launched three unidimensional and reliable scales which can now be utilized to measure beliefs, practices and management.

With regard to the second objective, we made a theoretical contribution to the field by providing solid support for Spolsky's conceptual model. The results of the Confirmatory Factor Analysis (CFA) involved evidence for three subtypes of construct validity: convergent, divergent and nomological validity. We first demonstrated that all items converge adequately and consistently on their respective constructs. Next, by offering proof for divergent validity we established that there are three components (i.e. beliefs, practices and management) of an overarching construct (i.e. language policy) that are sufficiently distinct from one another to be considered as separate components. Language policy is no unitary concept; therefore, if one intends to study language policy in its entirety, all of

its indispensable ingredients are to be identified. As a corollary, one simply cannot investigate management without considering beliefs and practices, and vice versa.

In addition, evidence was obtained for nomological validity as our results revealed the expected positive correlations between the three constructs. This empirical finding corroborates that beliefs, management and practices are indeed ‘interrelated,’ in accordance with Spolsky’s theory. In other words, when applying Spolsky’s model to the multilingual reality of an ECCE context, this means that professionals who strongly believe in the value of home languages (i.e. beliefs) are more inclined to advise multilingual parents to speak the language they know best with their child (i.e. management) and tend to make use of home languages during the service encounter (i.e. practices). Conversely, professionals who are somewhat sceptical about the benefits of home languages (i.e. beliefs), are less likely to communicate with multilingual parents in any other language than the dominant language (i.e. practices). And, professionals who advise multilingual parents to only read in the dominant language to their child tend to show their disapproval when children speak their home language during the service encounter. While we admit that this article focuses exclusively on the dynamic interaction between the determinants that *form* language policy – we are, of course, aware that language policy does not exist in a vacuum. Subsequent quantitative studies could therefore probe into internal or external forces that *drive* language policy. Whilst the latter has already been done qualitatively, and at the macro level (Albury, 2016; Guerini, 2011), this has not yet been done quantitatively – and at the micro level.

By no means do we pretend to have generated turnkey measures that are universally applicable. Moreover, ‘validation is an ongoing process’ (Zumbo & Chan, 2014, p. 4), which means that while ‘authentically valid’ scales cannot be produced in a single study, they can be achieved through replication and refinement in other research environments (Venkatraman, 1989). Therefore, it would be invaluable to continue our tentative steps by testing, improving and further validating the three scales. Needless to say, the rich data of the present study suggest avenues for future research. After examining the Average Variance Extracted (AVE) of each factor we discovered that the beliefs’ average communality was very low, which is indicative of inordinate measurement error. However, as AVE estimates are often found to be subpar, especially in exploratory research (Hair et al., 2014), the matter was disregarded for the purpose of this study. Also, validity is a matter of degree, rather than a question of all or nothing (Viswanathan, 2005; Zumbo & Chan, 2014), so we did not want to throw the baby out with the bathwater. After all, the construct was found to be unidimensional and yielded adequate internal consistency reliability. The authors did not intend to turn a blind eye to the matter at issue, yet the reader should bear in mind that no measure is entirely free from measurement error (Schmidt & Hunter, 1996) and that it is difficult to pinpoint what lies at the root of this bias (e.g. halo effects, proximity errors, social desirability). The beliefs construct thus requires some adjustment and we urge researchers to circumvent measurement error at item level (e.g. evaluate content, purify and if necessary reword the items) and at measure level (e.g. examine sequencing, introduce filler items from other scales between items) (Viswanathan, 2005).

A noteworthy finding was that the Principal Component Analysis (PCA) of the ‘management’ scale revealed two entirely different components, rather than some kind of continuum. This may indicate that in the respondents’ minds there are actually two separate

management dimensions, that is, ‘management towards the minority language’ (factor I) and ‘management towards the majority language’ (factor II). Consequently, although our preference for factor I was warranted, for this study focuses on the role of home languages, the scale that emerged in the second factor may be useful in other research contexts where the focus is more on the dominant language or on language mixing. It would also be interesting to determine the grounds for this management bifurcation by investigating what elicited professionals to opt for one of the advice directions.

And two final remarks. First, we acknowledge that the dichotomy between the Flemish Community’s *macro* LP – with its strong focus on using only the majority language in the public domain – on the one hand and the facilities’ *meso* LP – with an explicit openness to linguistic diversity – on the other, may have permeated respondents’ *micro* level LP dynamics. That is, professionals may have responded to the questions in terms of what they think is expected of them, whether that may be from the Flemish authorities or from the Child and Family agency. Nonetheless, this multilayered permeation (Ricento & Hornberger, 1996) is not only challenging to grasp solely by means of a questionnaire, but is also beyond the scope of this paper. Secondly, we draw the attention to a conspicuous feature of the parent-professional relationship which cannot be passed over in silence: power disparity (Hall & Slembrouck, 2009; Maynard, 1991). The very nature of advice giving engenders an asymmetric power relationship between both actors, whether the advice was solicited or not (Schaerer et al., 2018). This does not necessarily mean that the asymmetry is a resultant of the relationship; rather, it is naturally inherent, without negating the possible reciprocity between the actors (Vandenbroeck, 2015, p. 111). It is potentially a two-way street; for parents (e.g. with high socioeconomic status) can not only reject the offered advice, but they can in turn (un)consciously influence professionals’ ECLP. As a matter of fact, it would be interesting to establish which of the LPs (macro or meso) most influences the LP dynamics within the parent-professional relationship. That being said, both limitations could be addressed via an ethnographic enquiry into early childhood professionals’ language beliefs and everyday language practices as well as interviews with multilingual parents – in order to paint a more holistic picture of the phenomenon under investigation. We believe that our study provides a solid foundation to embark on such an inquiry.

Conclusion

Linguistic plurality, engendered by the ever-growing mobility around the world, has become ubiquitous in every sphere of contemporary society. This affects the kind of language policies that are shaped and implemented in the domain of ECCE and poses challenges for the early childhood workforce. The empirical contribution of this study lies in introducing a measurement model – that satisfied rigorous validity criteria as well as theoretical rationales – to capture Early Childhood Language Policy (ECLP) from a grassroots perspective. For the first time, it is now possible to measure LP in ECCE using valid scales which can be employed to clarify the triangular interconnections, and which are conducive to obtaining information on a large number of individuals. The three measures constitute a rough cut that may contribute to more quantitative research on LP in different spaces and layers. As regards our theoretical contribution to the field; we dare say that our endeavour to empirically test Spolsky’s conceptual model

proved worthwhile. We produced quantitative evidence by demonstrating that the vertices of the language policy triangle are significantly and positively related to, but also truly distinct from, each other. From these findings, one may conclude that language policy functions as a *gestalt* – generated by the interaction between its indispensable parts and where the whole is more than the aggregation of its ingredients.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by Fonds Wetenschappelijk Onderzoek [grant number S006518N].

Notes on contributors

Victoria Van Oss is doing a PhD in Linguistics at Ghent University, Belgium. Her research interests include multilingualism, language policy and Early Childhood Care and Education (ECCE). She investigates health care professionals' language policy dynamics (i.e. beliefs, practices and management) in relation to multilingual families visiting Flemish (Belgium) infant welfare clinics.


Piet Van Avermaet is Professor in 'Language and Diversity' in the Linguistics Department of Ghent University, Belgium. He is also Head of the Research Centre for Diversity & Learning (CDL) at the same University. His expertise and research interests cover topics related to multilingualism in education, social inequality in education, language policy and practice in education, language policy and practice in contexts of (social) inclusion, language assessment, diversity and inclusion, integration and participation, discrimination in education.

Esli Struys is a professor of applied linguistics at the Vrije Universiteit Brussel. He teaches courses in the BA and MA programmes of (applied) linguistics on multilingualism (from cognitive and educational perspectives), psycho- and neurolinguistics, and interpreting studies. His research foci include the cognitive processes involved in second language acquisition, bilingualism, and interpreting; and multilingual education in family and school settings.

Wendelien Vantieghem is Professor in 'Language, diversity & learning' at the Faculty of Arts & Philosophy of Ghent University and vice-director of the Research Centre for Diversity & Learning (CDL). Wendelien's work researches diversity and inclusion from different angles, with a focus on the inequalities in the educational system in terms of gender, sexual orientation, multilingualism, SES, ethnicity and disability, and studies the attitudes and competences of teachers with regard to diversity.

ORCID

Victoria Van Oss  <http://orcid.org/0000-0001-9065-3539>

Piet Van Avermaet  <http://orcid.org/0000-0001-8970-2911>

Esli Struys  <http://orcid.org/0000-0001-8093-3422>

Wendelien Vantieghem  <http://orcid.org/0000-0001-5891-2131>

References

Albury, N. J. (2016). National language policy theory: Exploring Spolsky's model in the case of Iceland. *Language Policy*, 15(4), 355–372. <https://doi.org/10.1007/s10993-015-9357-z>

- Angouri, J. (2013). The multilingual reality of the multinational workplace: Language policy and language use. *Journal of Multilingual and Multicultural Development*, 34(6), 564–581. <https://doi.org/10.1080/01434632.2013.807273>
- Arnau, R. C., & Thompson, B. (2000). Second order confirmatory factor analysis of the WAIS-III. *Assessment*, 7(3), 237–246. <https://doi.org/10.1177/107319110000700304>
- Bergroth, M., & Palviainen, Å. (2017). Bilingual children as policy agents: Language policy and education policy in minority language medium Early Childhood Education and care. *Multilingua*, 36(4), 375–399. <https://doi.org/10.1515/multi-2016-0026>
- Björkman, B. (2014). Language ideology or language practice? An analysis of language policy documents at Swedish universities. *Multilingua*, 33(3-4), 335–363. <https://doi.org/10.1515/multi-2014-0016>
- Bot, S., Terwee, C., Van der Windt, D., Bouter, L., Dekker, J., & De Vet, H. (2003, June). Psychometric evaluation of self-report questionnaires: The development of a checklist. In *Proceedings of the second workshop on research methodology* (pp. 161–168). Vu University.
- Caporal-Ebersold, E. (2018). Language policy in a multilingual crèche in France: How is language policy linked to language acquisition beliefs?. In M. Siiner, F. M. Hult, & T. Kupisch (Eds.), *Language policy and language acquisition planning* (pp. 55–79). Springer.
- Chan, A. (2011). Critical multiculturalism: Supporting early childhood teachers to work with diverse immigrant families. *International Research in Early Childhood Education*, 2(1), 63–75.
- Chatzidaki, A., & Maligkoudi, C. (2013). Family language policies among Albanian immigrants in Greece. *International Journal of Bilingual Education and Bilingualism*, 16(6), 675–689. <https://doi.org/10.1080/13670050.2012.709817>
- Child and Family. (n.d.). <https://www.kindengezin.be/cijfers-en-rapporten/cijfers/over-kind-en-gezin/>
- Child&Family. (2012). *Vision statement. Language stimulation and multilingualism*. <https://www.kindengezin.be/img/Taalstimulering2015Engels.pdf>
- Clark, L. A., & Watson, D. (1995). Constructing validity: Basic issues in objective scale development. *Psychological Assessment*, 7(3), 309–319. <https://doi.org/10.1037/1040-3590.7.3.309>
- Couëtoux-Jungman, F., Wendland, J., Aidane, E., Rabain, D., Plaza, M., & Lecuyer, R. (2010). Bilinguisme, plurilinguisme et petite enfance. *Devenir*, 22(4), 293–307. <https://doi.org/10.3917/dev.104.0293>
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52(4), 281. <https://doi.org/10.1037/h0040957>
- Cummins, J. (1991). The development of bilingual proficiency from home to school: A longitudinal study of Portuguese-speaking children. *Journal of Education*, 173(2), 85–98. <https://doi.org/10.1177/002205749117300207>
- De Houwer, A. (2017). Minority language parenting in Europe and children’s well-being. In N. J. Cabrera & B. Leyendecker (Eds.), *Handbook on positive development of minority children and youth* (pp. 231–246). Springer International Publishing.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd edn). Sage publications.
- Gerbing, D. W., & Anderson, J. C. (1984). On the meaning of within-factor correlated measurement errors. *Journal of Consumer Research*, 11(1), 572–580. <https://doi.org/10.1086/208993>
- Guerini, F. (2011). Language policy and ideology in Italy. *International Journal of the Sociology of Language*, 2011(210), 109–126. <https://doi.org/10.1515/ijsl.2011.033>
- Hair, J. F. Jr., Black, W. C., Babin, B. J., & Anderson, R. E. (2014). *Multivariate data analysis* (7th ed.). Pearson Education Limited.
- Hall, C., & Slembrouck, S. (2009). Communication with parents in child welfare: Skills, language and interaction. *Child & Family Social Work*, 14(4), 461–470. <https://doi.org/10.1111/j.1365-2206.2009.00629.x>
- Hélot, C. (2010). “Tu sais bien parler maîtresse!”: Negotiating languages other than French in the primary classroom in France. In K. Menken & O. Garcia (Eds.), *Negotiating language policies in schools. Educators as policymakers* (pp. 52–71). Routledge.
- Hooper, D., Coughlan, J., & Mullen, M. R. (2008). Structural equation modelling: Guidelines for determining model fit. *Electronic Journal of Business Research Methods*, 6(1), 53–60.

- Johnson, D. C. (2013). *Language policy*. Palgrave Macmillan.
- Kapuscinski, A. N., & Masters, K. S. (2010). The current status of measures of spirituality: A critical review of scale development. *Psychology of Religion and Spirituality*, 2(4), 191–205. <https://doi.org/10.1037/a0020498> doi:10.1037/a0020498
- Kind en Gezin. (2019). *Kind in Vlaanderen 2019*. <https://www.kindengezin.be/img/KIV2019.pdf>
- Kingsley, L. (2009). Explicit and implicit dimensions of language policy in multilingual banks in Luxembourg: An analysis of top-down and bottom-up pressures on practices. *Language Problems and Language Planning*, 33(2), 153–173. <https://doi.org/10.1075/lplp.33.2.04kin>
- Kirsch, C., & Aleksic, G. (2018). The effect of professional development on multilingual Education in Early Childhood in Luxembourg. *Review of European Studies*, 10(4), 148–163. <https://doi.org/10.5539/res.v10n4p148>
- Kline, R. (2011). *Principles and practice of structural equation modeling* (3rd ed). Guilford Press.
- Kristinsson, A. P. (2012). Language management agencies counteracting perceived threats to tradition. *Language Policy*, 11(4), 343–356. <https://doi.org/10.1007/s10993-012-9253-8>
- Little, R. J. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, 83(404), 1198–1202. <https://doi.org/10.1080/01621459.1988.10478722>
- Marope, P. T. M., & Kaga, Y. (2015). *Investing against evidence: The global state of early childhood care and education*. UNESCO Publishing.
- Maynard, D. W. (1991). Interaction and asymmetry in clinical discourse. *American Journal of Sociology*, 97(2), 448–495. <https://doi.org/10.1086/229785>
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50(9), 741–749. <https://doi.org/10.1037/0003-066X.50.9.741>
- Mettewie, L., & Mensel, L. V. (2020). Understanding foreign language education and bilingual education in Belgium: A (surreal) piece of cake. *International Journal of Bilingual Education and Bilingualism*, 1–19. <https://doi.org/10.1080/13670050.2020.1768211>
- Mirvahedi, S. H., & Jafari, R. (2021). Family language policy in the city of Zanjan: A city for the forlorn Azerbaijani. *International Journal of Multilingualism*, 18(1), 1–23. <https://doi.org/10.1080/14790718.2018.1545019>
- Morgado, F. F., Meireles, J. F., Neves, C. M., Amaral, A. C., & Ferreira, M. E. (2018). Scale development: Ten main limitations and recommendations to improve future research practices. *Psicologia: Reflexão e Crítica*, 30(1), 3. <https://doi.org/10.1186/s41155-016-0057-1>
- Paulhus, D. L. (2017). Socially desirable responding on self-reports. In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of personality and individual differences* (pp. 1–5). Springer International Publishing.
- Ricento, T., & Hornberger, N. H. (1996). Unpeeling The Onion: Language planning and policy and the ELT professional. *Tesol Quarterly*, 30(3), 401–427. <https://doi.org/10.2307/3587691>
- Rodríguez, M. (2015). Families and educators supporting bilingualism in early childhood. *School Community Journal*, 25(2), 177–194.
- Schaerer, M., Tost, L. P., Huang, L., Gino, F., & Larrick, R. (2018). Advice giving: A subtle pathway to power. *Personality and Social Psychology Bulletin*, 44(5), 746–761. <https://doi.org/10.1177/0146167217746341>
- Schmidt, F. L., & Hunter, J. E. (1996). Measurement error in psychological research: Lessons from 26 research scenarios. *Psychological Methods*, 1(2), 199. <https://doi.org/10.1037/1082-989X.1.2.199>
- Shohamy, E. (2006). *Language policy: Hidden agendas and New approaches*. Routledge.
- Spolsky, B. (2004). *Language policy*. Cambridge University Press.
- Spolsky, B. (2007). Towards a theory of language policy. *Working Papers in Educational Linguistics (WPEL)*, 22(1), 1.
- Spolsky, B. (2017). Language policy in education: Practices, ideology, and management. In T. McCarty & S. May (Eds.), *Language policy and political issues in education* (pp. 1–14). Springer International Publishing.

- Spolsky, B. (2019). A modified and enriched theory of language policy (and management). *Language Policy*, 18(3), 323–338. <https://doi.org/10.1007/s10993-018-9489-z>
- Vandenbroeck, M. (2015). Ethnic diversity and social inclusion. In E. C. C. E. in Europe, I. P. T. M. Marope, & Y. Kaga (Eds.), *Investing against evidence. The global state of early childhood care and education* (pp. 102–114). UNESCO Publishing.
- Van Gorp, K., & Moons, C. (2014). Creating rich language environments for more than one language: A work in progress in Flemish childcare. *European Journal of Applied Linguistics*, 2(1), 53–78.
- Van Maanen, J., Sørensen, J. B., & Mitchell, T. R. (2007). The interplay between theory and method. *Academy of Management Review*, 32(4), 1145–1154. <https://doi.org/10.5465/amr.2007.26586080>
- Venkatraman, N. (1989). Strategic orientation of business enterprises: The construct, dimensionality, and measurement. *Management Science*, 35(8), 942–962. <https://doi.org/10.1287/mnsc.35.8.942>
- Verdon, S., McLeod, S., & Winsler, A. (2014). Language maintenance and loss in a population study of young Australian children. *Early Childhood Research Quarterly*, 29(2), 168–181. <https://doi.org/10.1016/j.ecresq.2013.12.003>
- Viswanathan, M. (2005). *Measurement error and research design*. SAGE Publications, Inc.
- Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist*, 34(6), 806–838. <https://doi.org/10.1177/0011000006288127>
- Yeomans, K., & Golder, P. (1982). The Guttman-Kaiser criterion as a predictor of the number of common factors. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 31(3), 221–229. <https://doi.org/10.2307/2987988>
- Zumbo, B. D., & Chan, E. K. H. (2014). *Validity and validation in social, behavioral, and health sciences*. Springer International Publishing.