JSSER
www.jsser.org

Journal of Social Studies Education Research

Sosyal Bilgiler Eğitimi Araştırmaları Dergisi

2024:15 (2), 273-304

Measuring a National Reading Program: Questionnaires Design, Validation and Pilot Testing

Jesús Honorato-Errázuriz[1] & Valentina Bastidas-Schade [2] & Maria-Soledad Ramírez-Montoya[3]*.

Abstract

During times of crisis, such as the COVID-19 pandemic, the evaluation of educational programs becomes crucial for making evidence-based decisions. This study aims to validate and pilot an assessment instrument tailored to evaluate an innovative national reading program in Chile, particularly during the critical phase of post-pandemic educational recovery. The focus of this study is to encompass three key areas: a. assessing reading proficiency in first-grade students, b. examining the utilization of technology and reading materials in early childhood education, and c. gauging the comprehension and utilization of program information by all stakeholders involved. The research design is grounded in a research and development (R&D) approach employing a mixed-method model. We engaged 11 experts in the field to assess the questionnaire's validity using the Delphi method, administered to a representative sample. Subsequently, through confirmatory factor analysis, we established the construct validity of the questionnaires, revealing goodness-of-fit indices such as comparative fit index (CFI) scores ranging from 0.90 to 0.92, a root means square error of approximation (RMSEA) of 0.08, and Tucker-Lewis's index (TLI) values between 0.88 and 0.90. Reliability analysis yielded robust outcomes, with Cronbach's alpha coefficients ranging between 0.89 and 0.93 for the questionnaire and McDonald's omega coefficients ranging between 0.92 and 0.93. Our study confirms the validity of the questionnaires, positioning them as dependable tools for evaluating the national reading literacy development program in Chile. These validated instruments hold the potential to facilitate further research endeavors and inform the formulation of effective public reading policies in Chile and beyond, across diverse educational landscapes.

Keywords: *Construct validation, educational innovation, educational measurement tools, reading competence*

Introduction

Reading plays an important role in the educational journey, profoundly shaping future academic and professional success and contributing significantly to a nation's progress. To effectively restore and elevate pre-pandemic levels of reading proficiency, governments must establish robust

[1] *Assoc. Prof., Faculty of Education, Universidad de las Américas; Chile, jhonorato@udla.cl*
[2] *Assoc Prof., Institute of Mathematics, Physics, and Statistics, Universidad de las Americas; Chile, vbastidas@udla.cl.*
[2]* *Research Prof. at EGADE Business School and Institute for the Future of Education, Tecnologico de Monterrey, Mexico solramirez@tec.mx*

national reading programs and conduct thorough evaluations to ensure timely and impactful outcomes (Orellana-García, 2018). Research highlights the importance of initiating early reading skill development, emphasizing the need to prioritize reading programs from the earliest stages of education for any nation seeking to recover lost learning (Scarborough, 2001). Furthermore, it is crucial to integrate innovative methodologies, such as leveraging technology (Mahardhani, 2023), to support comprehensive public policy initiatives encompassing information dissemination, training, and engagement, ensuring the assimilation of program knowledge by all stakeholders (Ministerio de Ciencias, Tecnología e Innovación [Minciencias], 2020). This approach strengthens the effectiveness of reading programs.

Governments bear the responsibility of assessing the effectiveness and societal implications of all public policies (European Commission, 2021; Howlett, 2019; Organization for Economic Co-operation and Development [OECD], 2023). It is essential to evaluate the impact of innovative early reading programs to potentially elevate these initiatives to the status of national public policies. Such policies can foster citizen engagement and aid in the restoration of reading skills, which are fundamental for the holistic development of nations.

Literature review

Early reading methods

Recent research has highlighted that advancements in early reading methodologies, especially those emphasizing crucial components like reading comprehension, hold the potential to substantially bolster learning recovery in early education. Therefore, a comprehensive assessment of reading programs should encompass pivotal elements such as phonemic awareness, phonological skills, fluency, vocabulary, comprehension, and metacognitive strategies, given their paramount importance to educational innovation (Ashby et al., 2023; McKenna & Stahl, 2020; Medina et al., 2014; National Reading Panel [NRP], 2000).

Within the innovation ecosystem, knowledge and technological outputs, alongside technology transfer (TT), play important roles (World Intellectual Property Organization [WIPO], 2023). They not only foster opportunities for participation, interest, and dissemination but also facilitate the social appropriation of scientific and technological knowledge (Daza-Caicedo et al., 2017; Ramírez-Montoya et al., 2023; Romero-Rodríguez et al., 2020), which can considerably affect

public policy (Cuéllar-Gálvez et al., 2018). Therefore, a reading program tailored for early childhood education should conduct a comprehensive assessment of its components and the involvement of key stakeholders in the learning process of young children.

Measuring the Implementation of Reading Programs Through Questionnaires

Ensuring measurement quality is paramount for the development of sustainable and impactful public education policies. Quality measurement encompasses all facets of a program, including its objectives and citizen engagement (Gertler et al., 2016). In experimental research, when conducting impact evaluations, validating, and piloting instruments is crucial. This process guarantees their validity and reliability, ensuring efficient utilization of collected data for analysis and interpretation (Middleton, 2022; Sireci & Benitez, 2023). Questionnaires serve as valuable tools for collecting data from a large population within a limited timeframe and at a lower cost (Pew Research Center, 2023). Therefore, ensuring the validity and consistency of questionnaire design is imperative to optimize the utilization of collected data and facilitate its analysis and interpretation.

The literature shows the extensive utilization of questionnaires in evaluating reading programs, serving as valuable tools for collecting diverse information about various aspects and stakeholders involved. As presented in Table 1, there are questionnaires tailored to collect data on aspects such as reading comprehension development (Bogaert et al., 2023) or literacy acquisition (Li et al., 2023), as well as perceptions of reading both at home and in educational settings (Fives, 2016). Moreover, questionnaires are instrumental in soliciting feedback from reciprocal reading programs (Cockerill et al., 2019) and assessing the fidelity of reading program implementation (Demers et al., 2020). At the national level in Portugal, questionnaires are employed to gauge these dimensions (Da Costa et al., 2015). These questionnaires are meticulously crafted based on benchmarks associated with the components of reading comprehension and its underlying processes.

Table 1

Questionnaires in Reading Programs - Wos and Scopus Articles

Program	Name of the instrument	To whom is this addressed	References
Wizards of Words	Questionnaire for Home Reading Attitude (ARH) and Classroom Reading Attitude (ARC)	Student	Measure Developed by Researchers (Fives, 2016)
Standard Project READS	Motivation Reading Questionnaire (MRQ)	Student	Guthrie et al. (2004)
Istation Reading Program	Teacher's Survey	Teachers	Researchers' measure (Luo et al., 2017)
Reciprocal Reading Program	Teacher and Assistant Survey	Teachers and Teacher Assistants	Researchers' measure (Cockerill et al., 2019)
Animal-Assisted Intervention	ERAS (Elementary Reading Attitude Survey)	Student	Version (McKenna & Kear., 1990)
Explicit Code-Based Reading Instruction	The Survey of Preparedness and Knowledge of Language Structure	Teachers	Instrument (Cohen et al., 2017) comprising items adapted from prior research (e.g., Georgiou et al., 2021; Goldfeld et al., 2022; Peytcheva, 2019; Piasta et al., 2009; Piasta et al., 2020; Porter et al.,2022; Puliatte & Ehri, 2018), in addition to newly developed items
Book-Reading Program	Questionnaires Assessing Structural Quality	Teachers	Developed by the Researchers (Pelatti & Piasta, 2017)
National Reading Plan in Portugal	School Survey	Teachers	The central offices of the Ministry of Education (Da Costa et al., 2015)
	Public opinion barometer	The General Public	The central offices of the Ministry of Education (Da Costa et al., 2015)

Assessing Innovation, Technology, and Social Appropriation in Reading Programs

Innovation and technology play pivotal roles in shaping modern public programs, policy design, and assessment frameworks. UNESCO (2023) has proposed a comprehensive set of indicators aimed at evaluating access to, utilization of, and proficiency in digital skills, as well as attitudes towards digital technology, to inform policy decisions. To assess the effectiveness of technology in fostering reading skills, surveys have been developed based on validated scales such as the ‘‘Technology Acquisition Scale’’ (Tuomi et al., 2023) and the ‘‘Teacher Metacognition and Technology Inventory’’ (Jiang et al., 2016; Ley et al., 2022). These instruments facilitate the collection of data regarding the utilization of technology in reading practices and the challenges encountered by educators in integrating technology into their instructional methods.

Social knowledge appropriation has emerged as a crucial dimension in evaluating program effectiveness, encompassing the perception and active involvement of various stakeholders,

including students, teachers, and parents. Specialized assessment tools, rooted in frameworks established by Daza-Caicedo et al. (2017), address a spectrum of domains, including technology-related knowledge. For instance, the ‘‘Social Appropriation Scale for Teachers in Higher Education’’ (Sahin et al., 2010) exemplifies this approach. This concept has been explored in diverse contexts such as entrepreneurship (Romero-Rodríguez et al., 2020) and technological convergence (Quevedo-Pinzón & Franco-Avellaneda, 2022), underscoring its escalating importance in science and technology fields over the past decade. However, existing instruments often fall short in comprehensively capturing data from all participants or providing an integrated assessment of a reading skills development program, encompassing reading skill enhancement, technology use, and social knowledge appropriation. To address this gap, we developed and validated questionnaires to evaluate the ‘‘Leo Primero' program.’’ These questionnaires were meticulously crafted and refined to systematically collect data on three critical variables: students' reading proficiency, technology utilization, and the extent of knowledge appropriation among stakeholders. Within this program, they served as primary tools for extensive data collection, ensuring broad participation among all stakeholders. These endeavors signify substantial advancements to this program compared to previous reading initiatives. The design and validation of these instruments align with robust public policy models, enriching the quality and comprehensiveness of data to inform government decision-making processes.

The scope of this research is clearly defined, aiming to confirm a hypothetical model and comprehend the causal relationships among measurable dimensions and individual characteristics. The primary objective is to develop, validate, and establish the reliability of questionnaires tailored to assess the implementation process of the national reading program. To accomplish this goal, we delineated three specific objectives: 1. Creating a preliminary version of the instrument, 2. Refining the instrument based on content validation, and 3. Proposing a definitive version of the instrument through construct validation and reliability analysis.

Method

Research Design

This study adopts a mixed-methods R&D approach proposed by (Gall et al., 1996). The research process entails 5 basic principles, with the fifth being optional: 1. Exploration to identify

weaknesses and benefits, and the creation of the prototype, 2. Expert judgment for the prototype, 3. Main field tests for a small sample, 4. Main field tests for an extended sample, and 5. Dissemination and implementation. Specifically, the study comprises three main phases: (I) Construction of a preliminary instrument version, (II) Expert field tests for content validation, and (III) Target audience field tests for construct validation and reliability analysis.

Participants

In this study, experts, teachers, and external professionals, participated across various stages of the research process. Their different perspectives and expertise contributed to the refinement and validation of the instrument evaluation.

In the initial stage of constructing the preliminary version of the instrument (stage I), the National Curriculum and Assessment Unit of the Ministry of Education initiated the process. This stage involved the collaboration of four experts from the unit, along with two experts from Queen's University Belfast. Additionally, at least four external professionals from an external consulting firm participated, contributing their expertise to the development of the instrument.

In the subsequent stage of field tests with experts for content validation (stage II), a total of 11 experts selected based on their extensive experience and expertise in the relevant field. Notably, 73% of these experts held doctorate degrees from prestigious Chilean and foreign universities, underscoring the high level of expertise and qualifications among the participants. (Refer to Appendix 3 for details on the experts.)

In the field tests conducted with the target audience for construct validation and reliability analysis (stage III), a total of 89 key informants, primarily teachers, participated in the study. Table 2 presents the demographic distribution of the participants, highlighting that 71professionals (approximately 79.78%) possessed extensive teaching experience spanning from 1 to 20 years. Moreover, 22 professionals (approximately 24.7%) either completed or were in the process of pursuing a postgraduate program, indicating a diverse range of educational backgrounds among the participants.

Table 2

Experience and Educational Level of Participants (Teachers)

	Educational level			
Experience (years)	Completed postgraduate	Incomplete postgraduate	Completed professional	Total
1-10	6	2	36	44
11-20	8	1	18	27
21-30	3	1	6	10
31-40	0	0	7	7
41-50	1	0	0	1
Total	18	4	67	89

Data Analysis

In the first stage, a preliminary version developed through literature review, aligned with objectives, and validated instruments.

In the second stage, Qualitative Delphi Prospective Method applied to validate instrument scales. Experts evaluate items for relevance, coherence, pertinence, clarity (Riaño & Palomino, 2015). Analysis includes statistical measures: mean, standard deviation, minimum, and maximum (Boubeta et al., 2008). Instrument adjusted based on results.

Following the third stage, the revised instrument version was utilized with 89 teachers across educational institutions. Descriptive analysis ensued, encompassing mean, range, standard deviation, standard error, skewness, and kurtosis, providing a comprehensive overview of the responses.

Construct validity was assessed through confirmatory factor analysis (CFA) utilizing the acquired data. The objective was to examine and validate whether the proposed theoretical model (Table 4) corresponds with the outcomes derived from its implementation. Six phases, as per Escobedo et al. (2016), were pursued: (1) specification, (2) identification, (3) parameter estimation, (4) fit evaluation, (5) model re-specification, and (6) interpretation of results.

Before proceeding with model identification and estimation, it's crucial to verify the suitability of the data for factor analysis. This will be accomplished by assessing the Kaiser-Meyer-Olkin (KMO) measure of sample adequacy. A value near 0 indicates suboptimal conditions for factor analysis, while values closer to 1 indicate suitability (Brand, 2021). Additionally, Bartlett's test of

sphericity will be employed to test the hypothesis that variables are uncorrelated in the population. Large test statistics reject the null hypothesis at a specific significance level, indicating non-zero item correlations.

In this study, despite the known theoretical constructs, a brief examination will be conducted to confirm the number of proposed factors using the scree plot (Cattell, 1966). This involves analyzing eigenvalues associated with the obtained data and their correlation matrix. Following model identification and parameter estimation, the model will be assessed using various goodness-of-fit measures. These statistical indicators measure the explanatory capacity of the model (Boubeta et al., 2008).

- Root mean square error of approximation (RMSEA) is a crucial indicator of model fit. Values below 0.05 are deemed excellent, while those surpassing 0.08 suggest poor fit (Lloret-Segura et al., 2014).
- Comparative fit index (CFI)) is another important measure of model fit. A CFI below 0.9 indicates poor, values between 0.9 and 0.95 are deemed good, while those surpassing 0.95 are considered very good (Brand, 2021).
- Tucker-Lewis's index (TLI) is also an important index for assessing model fit. Values close to 1.0 suggest a good fit (Matsumoto-Royo et al., 2021).

Following the provided values, model re-specification will be performed, followed by a subsequent evaluation utilizing the same goodness-of-fit indices.

Cronbach's alpha coefficient and McDonald's omega are computed to assess the internal consistency of the items within each scale. These coefficients enable evaluation of how the reliability of the test would change if a particular item was excluded (Frías-Navarro, 2022). For more comprehensive understanding of these coefficients, refer to McNeish (2018) and Vaske et al. (2017). Internal consistency will be analyzed for both the entire instrument and each of its dimensions using these coefficients.

As per convention, an alpha of 0.65 is typically deemed "adequate for a scale used in humanities research" (Vaske et al., 2017). Internal consistency values below this threshold suggest a low correlation between items. Item removal is warranted only when it significantly improves internal consistency or when a low value persists and removal yields an acceptable value (Zambrano et al., 2023). McDonald's Omega coefficient serves as an alternative to Cronbach's alpha, particularly

when the ordinal response scale is deemed more suitable. Notably, Cronbach's alpha assumes a continuous interval-scale response for variables (Frías-Navarro, 2022).

All statistical analyses were conducted using the base functions and the Lavaan package of the R software (Rosseel, 2012).

Findings

Stage 1: Preliminary Version of the Instrument.

The dimensions of the questionnaire in the preliminary version of the instrument are presented in Table 3.

Table 3

Questionnaire Dimensions

Dimensions	Name	Code	Items
1	Development of reading comprehension strategies in 1st grade texts	DSR	1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 1.10, 1.11, 1.12, 1.13, 1.14, 1.15
2	Impact of technological innovation on actors and learning	IMP	2.1, 2.2, 2.3, 2.4, 2.5, 2.6
3	Social appropriation of the program by the actors	APR	3.1, 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, 3.9, 3.10, 3.11, 3.12, 3.13, 3.14, 3.15, 3.16, 3.17, 3.18

Two Likert scales were defined for the instrument's questions, tailored to the nature of each item. One scale measures the level of agreement, while the other measures the frequency of activities. Comprehensive details of these scales are presented in Table 4.

Table 4

Likert Scales Used in the Instrument

Specification	Question	Scale	Categories
Level of agreement	How much do you agree with the following statements?	1 to 5	1: Strongly disagree 2: Disagree 3: Neither agree nor disagree 4: Agree 5: Strongly agree
Frequency	How often do you perform the following activities?	1 to 5	1: Never 2: Once every three months 3: 1 or 2 times a month 4: 1 or 2 times a week 5: More than three times a week

The relationship between each dimension and its items has been defined for each scale in Table 5.

Table 5

Dimensions, Scales, and Items of the Instrument (Initial Version)

Dimension	Name	Code	Number of questions	Section	Number of questions	Scale	Items
1	Development of reading comprehension in 1st grade text strategies	DSR	15	1. a	8	Level of agreement	1.7, 1.8, 1.9, 1.10, 1.11, 1.12, 1.14, 1.15
				1. b	7	Frequency	1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.13
2	Impact of technological innovation on actors and learning	IMP	6	2. a	2	Level of agreement	2.5, 2.6
				2. b	4	Frequency	2.1, 2.2, 2.3, 2.4
3	Social appropriation of the program by the actors	APR	18	3. a	10	Level of agreement	3.3, 3.4, 3.5, 3.9, 3.10, 3.13, 3.14, 3.15, 3.16, 3.18
				3. b	8	Frequency	3.1, 3.2, 3.6, 3.7, 3.8, 3.11, 3.12, 3.17

Stage 2: Field tests with experts for content validation

During this stage, experts assessed the proposed items for each dimension of the instrument utilizing a Likert scale. The details of each scale are presented in Table 5.

Table 6

Evaluation Scale, by Evaluation Criterion

REL	COH	PER	CLA
1: Not relevant at all	1: Not coherent at all	1: Not pertinent at all	1: Not clear at all
2: Slightly relevant	2: Slightly coherent	2: Slightly pertinent	2: Slightly clear
3: Sufficiently relevant	3: Sufficiently coherent	3: Sufficiently pertinent	3: Sufficiently clear
4: Relevant	4: Coherent	4: Pertinent	4: Clear
5: Highly relevant	5: Highly coherent	5: Highly pertinent	5: Very clear

Note. REL: Relevance - Measures the importance of the information obtained from the item for the research objective. COH: Coherence - Assesses how closely related the item is to the dimension or variable being measured. PER: Pertinence - Evaluates the relevance of the item to the actor being surveyed. CLA: Clarity - Determines whether the items are written in language understandable to the target audience.

The evaluation of the items according to each evaluation criterion is presented in Table 7.

Table 7

Statistics by Evaluation Criterion for the Instrument

	Relevance	Coherence	Pertinent	Clarity
Mean	4,43	4,39	4,45	4,47
Standard deviation	0,38	0,37	0,34	0,43
Minimum	3.00	3.00	3.17	3.33
Maximum	5.00	4.83	5.00	5.00

Stage 3: Field tests *were conducted* with *the* target audience for construct validation and reliability analysis.

The application of the instrument to the 89 participants (teachers) yielded descriptive statistics, as presented in Table 8.

Table 8

Descriptive Statistics of the Instrument Items

Item	n	Mean	Std. Dev.	Min.	Max.	Range	Kurtosis	Symmetry	Std. error
1.1	89	4.56	0.74	1	5	4	9.62	-2.64	0.0784
1.2	89	4.62	0.61	1	5	4	11.13	-2.50	0.0647
1.3	89	4.56	0.62	1	5	4	9.59	-2.20	0.0657
1.4	89	4.62	0.63	1	5	4	10.03	-2.46	0.0668
1.5	89	4.24	0.75	1	5	4	2.12	-1.04	0.0795
1.6	89	4.21	0.80	1	5	4	1.58	-1.04	0.0848
1.7	89	4.10	0.91	1	5	4	2.18	-1.38	0.0965
1.8	89	4.09	0.95	1	5	4	1.53	-1.28	0.1007
1.9	89	3.71	0.89	1	5	4	0.66	-0.82	0.0943
1.10	89	3.46	0.99	1	5	4	-0.04	-0.66	0.1049
1.11	89	4.10	0.78	1	5	4	1.49	-0.87	0.0827
1.12	89	3.43	0.99	1	5	4	-0.33	-0.36	0.1049
1.13	89	4.39	0.90	1	5	4	3.26	-1.77	0.0954
1.14	89	3.61	1.02	1	5	4	-0.18	-0.57	0.1081
1.15	89	4.12	0.85	1	5	4	1.17	-1.00	0.0901
2.1	89	3.55	1.21	1	5	4	-0.74	-0.44	0.1283

2.2	89	3.90	1.11	1	5	4	-0.35	-0.69	0.1177
2.3	89	3.65	1.22	1	5	4	-0.64	-0.56	0.1293
2.4	89	3.12	1.26	1	5	4	-1.01	-0.20	0.1336
2.5	89	4.28	0.84	2	5	3	0.75	-1.13	0.0890
2.6	89	3.48	1.13	1	5	4	-0.60	-0.40	0.1198
3.1	89	4.17	1.09	1	5	4	0.79	-1.27	0.1155
3.2	89	4.22	1.06	1	5	4	1.93	-1.57	0.1124
3.3	89	3.99	1.07	1	5	4	1.29	-1.29	0.1134
3.4	89	3.87	1.10	1	5	4	0.55	-1.05	0.1166
3.5	89	3.84	0.99	1	5	4	-0.08	-0.60	0.1049
3.6	89	2.27	1.25	1	5	4	-0.51	0.70	0.1325
3.7	89	1.67	1.17	1	5	4	0.77	1.45	0.1240
3.8	89	2.65	1.28	1	5	4	-0.92	0.34	0.1357
3.9	89	3.73	0.90	1	5	4	0.29	-0.47	0.0954
3.10	89	3.97	1.03	1	5	4	0.00	-0.80	0.1092
3.11	89	2.16	1.28	1	5	4	-0.41	0.84	0.1357
3.13	89	4.43	0.78	1	5	4	5.99	-2.02	0.0827
3.15	89	4.27	0.93	1	5	4	1.93	-1.40	0.0986
3.16	89	3.60	1.09	1	5	4	-0.43	-0.50	0.1155
3.17	89	3.72	1.14	1	5	4	-0.45	-0.63	0.1208
3.18	89	4.12	0.99	1	5	4	1.12	-1.23	0.1049

Some items of interest are presented in Figures 1, 2 and 3.

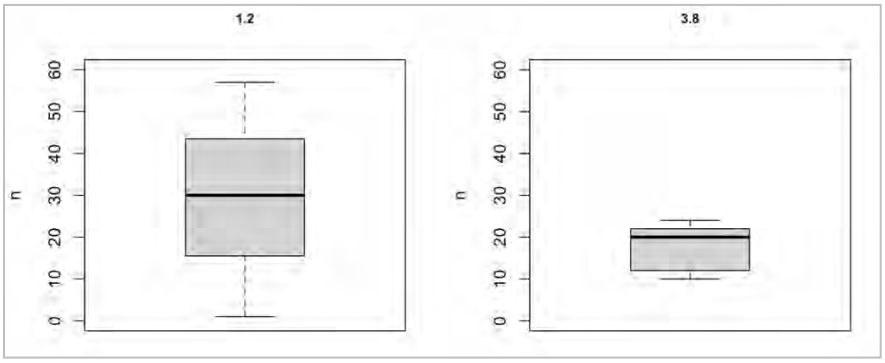

Figure 1 *Distribution of Items 1.2 and 3.8*

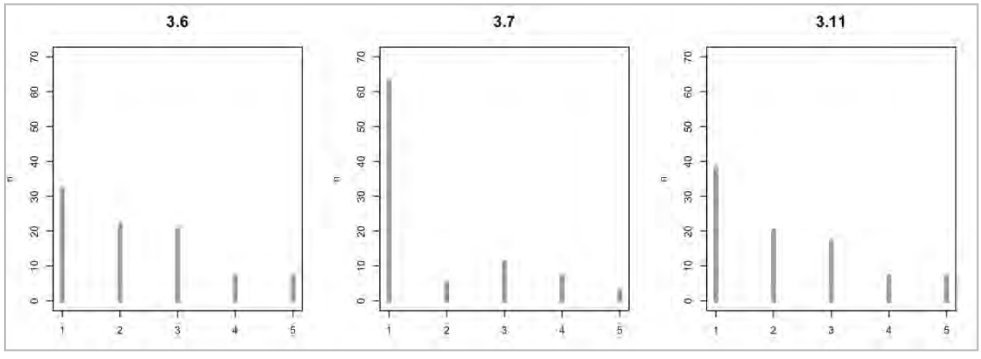

Figure 2 *Distribution of Items 3.6, 3.7 and 3.11*

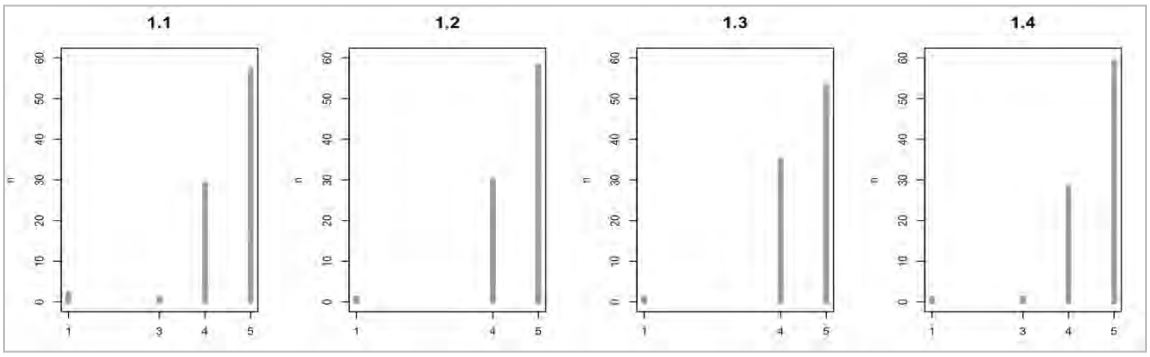

Figure 3 *Distribution of Items 1.1, 1.2, 1.3 and 1.4*

The number of dimensions of the questionnaire, was corroborated using the scree plot illustrated in Figure 4.

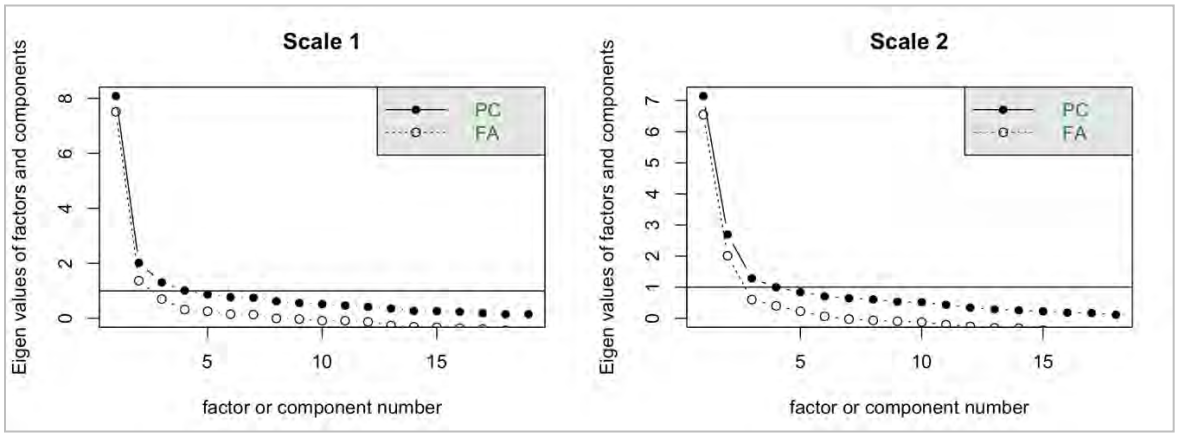

Figure 4 *Scree-plot of Both Scales*

Eight alternative models for Scale 1 and seven for Scale 2 were examined**.** The results of these fit measures are detailed in Tables 9 and 10.

Table 9

Evaluation of the Proposed Models for Scale 1

Goodness of fit measures of models for scale 1					
	CFI	CHISQ	RMSEA	SRMR	TLI
Model 1.1	0,79	328,04	0,12	0,09	0,76
Model 1.2	0,81	279,36	0,11	0,09	0,78
Model 1.3	0,82	270,7	0,11	0,09	0,79
Model 1.4	0,84	226,01	0,10	0,09	0,81
Model 1.5	0,84	202,75	0,11	0,09	0,81

Model 1.6	0,90	183,11	0,08	0,07	0,88
Model 1.7	0,88	200,36	0,09	0,09	0,86
Model 1.8	0,91	218,99	0,08	0,08	0,89

Table 10

Evaluation of the Proposed Models for Scale 2

Goodness of fit measures of models for scale 2					
	CFI	CHISQ	RMSEA	SRMR	TLI
Model 2.1	0,80	299,87	0,12	0,11	0,77
Model 2.2	0,82	250,92	0,11	0,11	0,79
Model 2.3	0,83	247,99	0,11	0,11	0,80
Model 2.4	0,86	221,24	0,10	0,10	0,83
Model 2.5	0,87	211,25	0,10	0,10	0,85
Model 2.6	0,87	211,25	0,10	0,10	0,85
Model 2.7	0,92	154,90	0,08	0,09	0,90

Two models have been identified as optimal (1.6 and 2.7), based on their respective fit indices. The structure of the final instrument is outlined in Table 11.

Table 11

Dimensions, Scales, and Items of the Instrument, According to Selected Models

Dimension	Name	Code	Number of questions	Section	Several questions.	Scale	Items
1	Development of reading comprehension in 1st grade text strategies	DSR	10	1. a	3	Level of agreement	1.8, 1.9, 1.12
				1. b	7	Frequency	1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.13
2	Impact of technological innovation on actors and learning	IMP	8	2. a	4	Level of agreement	2.5, 2.6, 1.14, 1.15
				2. b	4	Frequency	2.1, 2.2, 2.3, 3.17
3	Social appropriation of the program by the actors	APR	15	3. a	10	Level of agreement	3.3, 3.4, 3.5, 3.9, 3.10, 3.13, 3.15, 3.16, 3.18, 1.10
				3. b	5	Frequency	3.2, 3.6, 3.7, 3.8, 2.4

The factor loadings and standard errors for the selected model for Scale 1 and Scale 2 are presented in Tables 12 and 13, respectively.

Table 12

Factor Loadings and Standard Errors of the Model for Scale 1

	DSR		IMP		APR	
Item	Factor loadings	Standard error	Factor loadings	Standard error	Factor loadings	Standard error
1.8	0.701	0.000				
1.9	0.733	0.196				
1.12	0.815	0.234				
2.5			0.543	0.000		
2.6			0.507	0.394		
1.14			0.715	0.398		
1.15			0.697	0.223		
3.3					0.607	0.000
3.4					0.622	0.241
3.5					0.849	0.215
3.9					0.825	0.182
3.10					0.752	0.199
3.13					0.608	0.180
3.15					0.680	0.149
3.16					0.647	0.187
3.18					0.430	0.289
1.10					0.650	0.186

Table 13

Factor Loadings and Standard Errors of the Model for Scale 2.

	DSR		IMP		APR	
Item	Factor loadings	Standard error	Factor loadings	Standard error	Factor loadings	Standard error
1.1	0.664	0.000				
1.2	0.778	0.141				
1.3	0.871	0.174				
1.4	0.766	0.170				
1.5	0.674	0.177				
1.6	0.594	0.192				
1.13	0.671	0.251				
2.1			0.851	0.000		
2.2			0.621	0.131		
2.3			0.863	0.077		
3.17			0.620	0.121		
3.2					0.608	0.000
3.6					0.697	0.357
3.7					0.405	0.295
3.8					0.665	0.347
2.4					0.852	0.366

Both Cronbach's Alpha and McDonald's Omega coefficients are detailed in Table 14 and 15

Table 14

Cronbach's Alpha for the Complete Instrument

	Alfa coefficient	Omega coefficient	The average correlation between items
Scale 1	0.913	0.934	0.386
Scale 2	0.890	0.928	0.358

Table 15

Cronbach's Alpha Per Dimension for Each Scale

	Dimension	Alfa coefficient	Omega coefficient	The average correlation between items
Scale 1	DSR	0.793	0.797	0.562
	IMP	0.694	0.739	0.376
	APR	0.890	0.922	0.454
Scale 2	DSR	0.879	0.928	0.532
	IMP	0.826	0.854	0.539
	APR	0.782	0.877	0.412

Discussion

In the initial stage, three dimensions or constructs were proposed (Table 3), which were informed by two scales (Table 4), guiding the development of 37 questionnaire items (Table 5).

In the second stage, experts assessed the proposed items for each dimension of the instrument utilizing a Likert scale (Table 6). Following the experts' recommendations, items 3.12 and 3.14 were removed. Regarding the remaining items, overall, there is a good evaluation across each of the evaluation criteria, as presented in Table 7. Given that none of the items received a rating lower than 3.0, adjustments were made only to their wording and the presentation order within each dimension.

In the third stage, the findings (Table 8) reveal that item 3.8 exhibits the highest variability among responses, with a standard deviation of 1.28 and a maximum standard error of 0.1357 for the mean distribution. Conversely, item 1.2 exhibits the lowest variability, with a standard deviation of 0.61 and a minimum standard error of 0.0647 for the mean distribution. The distribution of these items is illustrated in Figure 1.

Most items demonstrate negative skewness coefficients, indicating leftward biases or tails skewed towards the lower end of the distribution. This trend is anticipated, as respondents tend to rate items positively, resulting in responses clustering towards the higher ranges of the scale. Notably, items 3.6, 3.7, and 3.11 exhibit pronounced right skewness, suggesting a positive bias or a right tail. These findings carry adverse implications for the instrument's utility, as they signify a concentration of responses towards lower ratings. Further details are delineated in Figure 2.

On the contrary, items 1.1, 1.2, 1.3, and 1.4 demonstrate kurtosis coefficients surpassing 9, indicating leptokurtic behavior characterized by a high concentration of responses around their means. This strong consensus among respondents proves beneficial for the instrument's objectives, reflecting agreement at higher scale values, as illustrated in Figure 3.

To assess the data's suitability for CFA, the Kaiser-Meyer-Olkin (KMO) measure and Bartlett's test were utilized (National Institute of Standards and Technology [NIST], 2024). For Scale 1, the KMO coefficient is 0.86, with Bartlett's test yielding a p-value of 0, confirming significant correlations between variables. Similarly, for Scale 2, the KMO is 0.84, and Bartlett's test produces a p-value of 0, also confirming significant correlations. These findings reveal that the data are well-suited for factor analysis.

Furthermore, the number of dimensions was corroborated using the scree plot illustrated in Figure 4, despite the known theoretical construct underpinning the instrument.

Graphical analysis suggests that the optimal number of constructs or factors falls between 2 and 4, consistent with the three theoretically proposed dimensions.

Consequently, factor analysis was executed for both scales, starting from theoretical base models, which exhibited inadequate fit (RMSEA 0.12 for both scales). To improve this, eight alternative models for Scale 1 and seven for Scale 2 were examined, involving modifications to item correlations or the instrument's structure. The results of these fit measures are detailed in Tables 9 and 10. Model 1.6 has been identified as optimal for Scale 1, while Model 2.7 is selected for Scale 2, based on their respective fit indices. Both models exhibit RMSEA values within the acceptable range of 0.08, while the CFI index indicates a good fit, ranging between 0.9 and 0.95. Additionally, the TLI approaches acceptable levels (Escobedo et al., 2016). For Scale 1, items 1.7 and 1.11 were removed due to their high correlation with other questionnaire items, indicating redundant variable measurement. This mirrors the earlier removal of item 3.14 during content validation. However, items 3.13, 3.15, and 3.18, despite displaying high correlations, were retained in the model with

acknowledged correlations among them, as they measure distinct variables. Furthermore, items 1.14 and 1.15 were relocated from the DSR dimension to the IMP dimension, and item 1.10 was moved from the DSR to the APR dimension.

In the revised model for Scale 2, items 3.1 and 3.11 were excluded, employing the same rationale as in Scale 1, akin to the prior removal of item 3.12. Despite their high correlation, items 1.1, 1.2, 1.13, 1.5, and 1.6 were retained in the model with acknowledged correlations. Moreover, item 3.17 was transferred from the APR dimension to the IMP dimension, while item 2.4 shifted from the IMP to the APR dimension. The updated instrument structure is outlined in Table 11.

The factor loadings predominantly range from moderate to large, indicating that the latent variables moderately explain the variance in the indicators (see Tables 12 and 13). The lowest factor loadings are observed for items 3.7 (0.405) and 3.18 (0.430).

Both Cronbach's Alpha and McDonald's Omega coefficients indicate high internal consistency, signifying a strong correlation among the items within the scales. The results of internal consistency for each scale are detailed in Table 14.

The analysis of Cronbach's alpha coefficient showed that removing individual items had minimal impact on the results. Internal consistency for each dimension of the instrument's scales was assessed, and Table 15 demonstrated good consistency across all dimensions. Therefore, no items were removed from the scales according to the selected models.

Conclusion and Implications

Learning to read stands as a crucial developmental milestone for children. Empirical evidence underscores that this learning process is not always seamless, emphasizing the importance of prioritizing reading programs in the early stages of education. These imperative holds particularly true for nations aiming to effectively recover lost learning (Orellana García, 2018). Furthermore, integrating innovative approaches, such as leveraging technology, becomes essential. In line with this, the Chilean national reading program, "Leo Primero," has been meticulously developed based on an extensive literature review on reading acquisition. In response to the challenges posed by the pandemic, the program has undergone updates aimed at enhancing distance learning and fostering the social appropriation of knowledge.

The instrument discussed in this study aims to measure three important latent variables pertinent to participating teachers, crucial for facilitating the program's implementation (Tables 2 and 4): the cultivation of comprehensive reading strategies in 1st-grade texts (Scarborough, 2001), the impact of technological innovation on stakeholders and learning outcomes (Tuomi et al., 2023), and the social appropriation of the program by its participants (Daza-Caicedo et al., 2017). These variables can be independently analyzed, enabling a comprehensive examination of the results.

During its inception, the instrument underwent a content validation process aimed at enhancing the relevance, coherence, pertinence, and clarity of its items. Subsequently, a confirmatory factor analysis was executed to evaluate the fit of the hypothesized models with the data. While initial fit indices were suboptimal, revisions to the models and adjustments to item variances notably enhanced these indices. Additionally, an analysis of internal consistency was conducted utilizing Cronbach's alpha and McDonald's omega coefficients, yielding robust outcomes for the questionnaire and its dimensions.

These adjustments have solidified a strong relationship between most items and their respective dimensions, affirming the construction validity of the questionnaires. This robust instrument now enables the continuous evaluation of the program, collecting vital data on the implementation process and teachers' perceptions, thus pinpointing strengths and weaknesses in program execution. Consequently, it facilitates informed decision-making aimed at achieving desired student learning outcomes.

The successful validation of these instruments carries substantial national and international implications. They have the potential to catalyze further research and offer crucial insights for the development of effective public policies in reading education. Their utility transcends the Chilean context, offering benefits to diverse educational settings worldwide. These tools prove particularly invaluable in comprehending behavioral responses during crises, furnishing lessons for future contingencies. In times of crisis, robust evaluation tools become indispensable for informed decision-making and the ongoing enhancement of educational practices.

The main limitation of this study arises from the challenges in sample selection imposed by the pandemic, which impeded effective communication with schools and their teaching staff. Future implementations should strive to engage larger sample sizes to enhance the reliability of the findings. Further studies could examine the construct validity of other related instruments

integrated into the public policy program, thereby providing a more comprehensive perspective on the program's overall implementation process.

Acknowledgment

This work is developed in the Doctorate program “Training in the Knowledge Society”, and GRIAL Group of the University of Salamanca. Also, the authors acknowledge the support of Tecnologico de Monterrey through the "Challenge-Based Research Funding Program 2023". Project ID #IJXT070-23EG99001, and, Writing Lab, Institute for the Future of Education.

References

Ashby, J., McBride, M., Naftel, S., O’Brien, E., Paulson, L. H., Kilpatrick, D. A., & Moats, L. C. (2023). *Teaching phoneme awareness in 2023: A guide for educators.* https://drive.google.com/drive/folders/1tPwupou7EqctDO50XtfvMGxfJAx7xPk9?usp=share_link

Bogaert, R., Merchie, E., Aesaert, K., & van Keer, H. (2023). The development of the reading comprehension - Progress monitoring (RC-PM) tool for late elementary students. *Frontiers in Education, 8*. https://doi.org/10.3389/feduc.2023.1066837

Boubeta, A. R., & Mallou, J. V. (2008). Estadística práctica para la investigación en ciencias de la salud. Netbiblo.

Brand, Y. (2021). *Modelos de ecuaciones estructurales: Conceptos y aplicaciones.* [Tesis de maestría, Universidad Nacional de Colombia]. https://repositorio.unal.edu.co/bitstream/handle/unal/80064/1059705148.2021.pdf?sequence=3

Cattell, R. B. (1966). The screen test for several factors. *Multivariate Behavioral Research, 1(2),* 245–276. https://doi.org/10.1207/s15327906mbr0102_10

Cockerill, M., Thurston, A., & Taylor, A. (2019). Protocol: An efficacy randomized controlled trial of reciprocal reading in high schools. *International Journal of Educational Research, pp. 97,* 99–106. https://doi.org/10.1016/j.ijer.2019.06.013

Cuéllar-Gálvez, D., Aranda-Camacho, Y., & Mosquera-Vásquez, T. (2018). A model to promote sustainable social change based on the scaling up a high-impact technical innovation. *Sustainability (Basel, Switzerland), 10(12)*. https://doi:10.3390/su10124532

Da Costa, A. F., Pegado, E., Ávila, P., & Coelho, A. R. (2015). Evaluating the Portuguese national reading plan: Teachers' perceptions on the impact in schools. *Educational Research for Policy and Practice, 14(2),* 119–138. https://doi.org/10.1007/s10671-014-9171-y

Daza-Caicedo, S., Maldonado, O., Arboleda-Castrillón, T., Falla, S., Moreno, P., Tafur-Sequera, M., & Papagayo, D. (2017). Hacia la medición del impacto de las prácticas de apropiación social de la ciencia y la tecnología: Propuesta de una batería de indicadores. *Historia, Ciencias, Saude - Manguinhos, 24(1),* 145–164. https://doi.org/10.1590/S0104-59702017000100004

Demers, C., Moreau, A. C., & Lefebvre, P. (2020). Élaboration et premiers pas de validation de questionnaires pour évaluer la fidélité du modèle de réponse à l'intervention en littératie dans les écoles primaires francophones québécoises. *Canadian Journal of Education, 43*(1), 60–85.

Escobedo, M. T., Hernández, J. A., Estebané, V., & Martínez, G. (2016). Modelos de ecuaciones estructurales: Características, fases, construcción, aplicación y resultados. *Ciencia & Trabajo, 18*(55), 16–22. https://www.scielo.cl/pdf/cyt/v18n55/art04.pdf

European Commission. (2021). *European research area policy agenda – Overview of actions for 2022-2024.* https://commission.europa.eu/system/files/2021-11/ec_rtd_era-policy-agenda-2021.pdf

Fives, A. (2016). Modeling the interaction of academic self-beliefs, frequency of reading at home, emotional support, and reading achievement: An RCT study of at-risk early readers in first and second grade. *Reading Psychology, 37(3),* 339–370. https://doi.org/10.1080/02702711.2015.1055870

Frías-Navarro, D. (2022). *Apuntes de estimación de la fiabilidad de consistencia interna de los ítems de un instrumento de medida.* Universidad de Valencia. España. https://www.uv.es/friasnav/AlfaCronbach.pdf

Gall, M. D., Borg, W. R., & Gall, J. P. (1996). Educational research: An introduction. Longman Publishing.

Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B., & Vermeersch, C. M. J. (2016). *Impact evaluation in practice.* The World Bank.

Howlett, M. (2019). *The policy design primer: Choosing the right tools for the job*. Routledge.

Jiang, Y., Ma, L., & Gao, L. (2016). Assessing teachers' metacognition in teaching: The teacher metacognition inventory. *Teaching and Teacher Education, 59,* 403–413. https://doi.org/10.1016/j.tate.2016.07.014

Ley, T., Tammets, K., Sarmiento-Márquez, E. M., Leoste, J., Hallik, M., & Poom-Valickis, K. (2022). Adopting technology in schools: Modelling, measuring and supporting knowledge appropriation. *European Journal of Teacher Education, 45*(4), 548–571. https://doi.org/10.1080/02619768.2021.1937113

Li, W., Kang, S., & Shao, Y. (2023). Development of the reading literacy questionnaire for EFL learners at primary schools. *Frontiers in Psychology, 14.* https://doi.org/10.3389/fpsyg.2023.1154076

Lloret-Segura, S., Ferreres-Traver, A., Hernández-Baeza, A., & Tomás-Marco, I. (2014). El análisis factorial exploratorio de los ítems: Una guía práctica, revisada y actualizada. *Anales de Psicología/Annals of Psychology, 30*(3), 1151–1169. https://dx.doi.org/10.6018/analesps.30.3.199361

Mahardhani, A. J. (2023). The role of public policy in fostering technological innovation and sustainability. *Journal of Contemporary Administration and Management (ADMAN), 1(2),* 47–53. https://doi.org/10.61100/adman.v1i2.22

Matsumoto-Royo, K., Ramírez-Montoya, M. S., & Conget, P. (2021). Opportunities to develop lifelong learning tendencies in practice-based teacher education: Getting ready for education 4.0. *Future Internet, 13*(11), 292. https://doi.org/10.3390/fi13110292

McKenna, M., & Stahl, S. (2020). *Assessment for reading instruction*. The Guilford Press.

McNeish, D. (2018). Thanks, coefficient alpha; we will take it from here. *Psychological Methods, 23*(3), 412. https://doi.org/10.1037/met0000144

Medina, L., Valdivia, A., & Martín, E. S. (2014). Prácticas pedagógicas para la enseñanza de la lectura inicial: Un estudio en el contexto de la evaluación docente chilena. *Psykhe, 23(2).* https://doi.org/10.7764/psykhe.23.2.734

Middleton, F. (2022). Reliability vs validity in research: Differences, types and examples. Scribbr. https://www.scribbr.co.uk/research-methods/reliability-or-validity/

Ministerio de Ciencias, Tecnología e Innovación. (2020). *Lineamientos para una política nacional de apropiación social del conocimiento*. https://minciencias.gov.co/sites/default/files/documento_de_lineamientos_para_la_politica_nacional_de_apropiacion_social_del_conocimiento_1.pdf

National Reading Panel. (2000). *Teaching children to read: An evidence-based assessment of the scientific research literature on reading and its implications for reading instruction.* National Institute of Child Health and Human Development.

National Institute of Standards and Technology. (2024). *Bartlett's test.* https://www.itl.nist.gov/div898/handbook/eda/section3/eda3

Organization for Economic Co-operation and Development. (2023). *Toolkit - Impact by design - Effective results frameworks for sustainable development.* https://one.oecd.org/document/DCD(2023)17/en/pdf

Orellana-García, P., Valenzuela, M.F., & Muños, K. (2018). Impacto de la lectura repetida interactiva en las habilidades verbales de preescolares de contextos vulnerables. *Educación y Educadores, 21(3),* 409–432. https://doi.org/10.5294/edu.2018.21.3.3

Pew Research Center. (2023). *Writing survey questions.* https://www.pewresearch.org/methods/u-s-survey-research/questionnaire-design/

Quevedo-Pinzón, E., & Franco-Avellaneda, M. (2022). Creencias de docentes de preescolar sobre ciencia y tecnología: Desafíos para la apropiación social del conocimiento en la infancia. *Revista Colombiana de Educación, 1*(84), 1–22. https://doi.org/10.17227/rce.num84-11365

Riaño, C. E., & Palomino, M. (2015). Diseño y elaboración de un cuestionario acorde con el método de Delphi para seleccionar laboratorios virtuales (LV). *Sophia, 11*(2), 129–141. https://www.redalyc.org/pdf/4137/413740778002.pdf

Romero-Rodríguez, J. M., Ramírez-Montoya, M. S., Aznar-Díaz, I., & Hinojo-Lucena, F. J. (2020). Social appropriation of knowledge as a key factor for local development and open innovation: A systematic review. *Journal of Open Innovation: Technology, Market, and Complexity, 6(2),* 44. https://doi.org/10.3390/joitmc6020044

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5-12 (BETA). *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Şahin, M., Akbaşlı, S., & Yanpar Yelken, T. (2010). Key competencies for lifelong learning: The case of prospective teachers. *Educational Research and Reviews, 5(10), 545–556.*

Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities: Evidence, theory, and practice. In S. Neuman & D. Dickinson (Eds.), *Handbook for Research in Early Literacy*. Guilford Press.

Sireci, S., & Benitez Baena, I. (2023). Evidence for test validation: A guide for practitioners. *Psicothema, 35(2),* 217–226. https://doi.org/10.7334/psicothema2022.477

Tuomi, I., Cachia, R., & Villar Onrubia, D. (2023). On the futures of technology in education: Emerging trends and policy implications. *Publications Office of the European Union.* https://doi.org/10.2760/079734

UNESCO. (2023). *Global Education Monitoring Report 2023: Technology in Education – A tool on whose terms?* UNESCO.

Vaske, J. J., Beaman, J., & Sponarski, C. C. (2017). Rethinking internal consistency in Cronbach's alpha. *Leisure Sciences, 39*(2), 163–173. http://dx.doi.org/10.1080/01490400.2015.1127189

World Intellectual Property Organization. *GII 2023 at a glance. The Global Innovation Index 2023 captures the innovation ecosystem performance of 132 economies and tracks the most recent global innovation trends.* https://www.wipo.int/edocs/pubdocs/en/wipo-pub-2000-2023-section1-en-gii-2023-at-a-glance-global-innovation-index-2023.pdf

Zambrano, S. E., Martínez, C. J. A., del Río, C. Y., Dehesa, M. M., Sánchez, C. N. A., Villalpando, H. J., et al. (2023). Design, construct, and validate the emotional skills questionnaire for school children (QESSC) Pilot test. *Invest Discapacidad, 9*(3), 115–128. https://dx.doi.org/10.35366/112697

Appendix

Appendix 1

Survey Evaluation of the Impact of Plan Leo Primero – Teachers

Good morning/afternoon, my name is _____ (SURVEYOR'S NAME), a surveyor from the company conducting the Impact Evaluation of the Chilean Ministry of Education's (MINEDUC) Plan Leo Primero.
This survey aims to collect information regarding the evaluation, experiences, and assessments of the Plan "Leo Primero," aimed at gaining insights into the program's implementation or understanding your experience with other educational strategies for first-grade students. This information will be highly relevant for the Ministry of Education in enhancing Plan Leo Primero. As a primary public policy, this initiative fosters the development of reading comprehension skills among students nationwide.

The survey is estimated to take approximately 15 minutes. Please rest assured that all the information you provide is confidential and anonymous, protected by law, and will solely be used for this study. Your responses will not influence or affect your relationship with MINEDUC in any manner.

Introduction

In your teaching experience, have you ever employed the Plan Leo Primero as an educational method for first-grade pupils?

1	Yes, I have used the Plan Leo Primero in my teaching experience.	Move to P2
2	I have never used the Plan Leo Primero in my teaching experience.	Move to P3
99	Do not know/No response	

During which years did you implement the Plan Leo Primero as an educational approach for first-grade students?

1	2018
2	2019
3	2020
4	2021
99	Do not know/No response

In 2021, which language education program or strategy did you utilize for first-grade students?

2	SM textbook
3	Matte method
4	Other, please specify
99	Do not know/No response

Knowledge of Plan Leo Primero (Only if you answered 2 in P1)
Why haven't you employed the Plan Leo Primero as an educational strategy for first grade in your teaching practice yet?

1	Because I am not familiar with it in depth	
2	Because I have been working with another methodology for a while	
3	Because it has not yielded good results for me	
4	By instruction of the institution	
5	By instructions from the provider	
6	Other, please specify	
99	Need to know/No response.	

Leo Primero 2021 experience (Only if you answered 4 in P2)
Participation in Activities
On a scale of 1 to 5, where 1 represents Strongly Disagree, 2: Disagree, 3: Neither Agree nor Disagree, 4: Agree, and 5: Strongly Agree, to what extent do you agree with the given statements?

	1	2	3	4	5
Students possess the ability to identify the shape and sound of letters.					

Approximately 80% of students demonstrate proficiency in recognizing, isolating, and blending phonemes and syllables, showcasing phonological awareness.					
Students exhibit fluent reading abilities when reading short texts (3 lines), pronouncing each word accurately, with occasional self-correction, and adhering to punctuation marks.					
After reading a familiar text, the majority of students can appropriately utilize all vocabulary words from the Plan Leo Primero cards and identify the central ideas or events of the text.					
Following the reading of a familiar text, the majority of students demonstrate the ability to identify the central ideas or events discussed in class.					
Students are capable of independently composing a simple and concise text on a topic of their choice.					
The implementation of Plan Leo Primero effectively supported students in their journey to acquire reading skills.					
Plan Leo Primero contributed significantly to students' enhancement in reading comprehension skills.					
I have enhanced my capacity to search for information in digital resources for both my teacher training and class planning compared to the start of this school year.					
The insights gathered from the exit tickets on the Leo and Sumo Primero monitoring app have proven valuable in identifying students' reading strengths and areas needing improvement.					
Class content adheres to the sequential structure outlined in the Plan Leo Primero textbook.					
The educational institution has promoted the utilization of Plan Leo Primero's digital resources, including public class and strategy videos, to the teaching team. These resources have proven beneficial for me in monitoring students' levels of reading comprehension.					
Plan Leo Primero's resources have proven valuable in aiding me to assess students' levels of reading comprehension.					
The monitoring platform and unit tests have proved useful for me in tracking the content to focus on with the students.					
My proficiency in utilizing platforms such as WhatsApp, the website, webinars, and video streaming of Plan Leo Primero has improved since the beginning of this school year.					
I intend to employ reading teaching strategies developed or enhanced during the year 2021 in the upcoming years.					
I enthusiastically engaged in and contributed to initiatives concerning Leo Primero and reading. I have a keen interest in sharing knowledge, sharing anecdotes, connecting with current events, commenting on news and progress, and celebrating students' achievements when they reach milestones.					
Learning strategies and Plan Leo Primero were adapted to meet the individual needs of each student.					
The entire academic team is at a consensus regarding pedagogical decisions.					

On a scale of 1 to 5, where 1 represents Never, 2: once every 3 months, 3: 1 or 2 times per month, 4: 1 or 2 times per week, and 5: more than 3 times per week, how frequently do you engage in the following activities?

	1	2	3	4	5
I focused on strengthening and enhancing students' knowledge of the alphabetic principle.					
I engaged in activities aimed at practicing and strengthening students' word-decoding skills.					
I conducted exercises to develop phonological awareness, separate syllables, recognize words, and increase awareness of reading pauses.					
I organized activities to bolster reading comprehension by asking questions to students about the content of the texts during class.					
I organized activities to improve writing skills, incorporating dictations and free writing exercises.					
I promoted student engagement by encouraging them to alternate reading a portion of the text to their classmates.					
I monitored the time students required to arrange the class for each activity.					
I integrated content from Plan Leo Primero's digital resources (website, public classes, digital library, Leo Primero app, audiobooks) into activities aimed at enhancing student reading skills.					
I utilized information from videos, websites, blogs, scientific publications, and books to prepare my lessons or to acquire new educational strategies.					
I incorporated content from Plan Leo Primero's digital resources (website, public classes, digital library, Leo Primero app, audiobooks, strategy videos, and class videos) into my professional development efforts.					
I kept myself informed about Plan Leo Primero through various channels such as the website, social media, email, or WhatsApp.					

I used Plan Leo Primero's didactic guide to plan my lessons.					
I used Plan Leo Primero's resources (textbook, shared reading sheets, public classes, and others) to prepare lessons for first-grade students.					
I participated in webinars and watched video streams of the Leo Primero Plan.					
I documented students' progress on the class progress tracking platform.					
I reviewed the video capsules and public classes of the Leo Primero Plan to enhance my reading teaching strategies.					
I took part in school meetings where I received information regarding the Leo Primero Plan, including usage guidelines, videos, documents, and conference dates.					
I facilitate group reading sessions with the students and offer feedback on their reading abilities.					

Health Crisis Covid-19

What teaching strategy or strategies did the school implement for second-grade students in your subject during the quarantine period prompted by the COVID-19 health crisis?

1	Send physical learning materials (guides, texts, and tests)	Pass to Question 12
2	Send learning materials through email (guides, texts, and tests)	
3	Online classes taught by schoolteachers	
4	Send teaching guidelines to parents so that they can teach children how to read.	
5	Other (please specify):	
6	The establishment did not implement teaching strategies during the quarantine period.	Pass to Question 14
7	I did not belong to the establishment during this period	
99	Need to know/No response.	

On a scale of 1 to 4, with 1 being wholly insufficient and 4 being entirely sufficient, how effective do you find the educational strategies implemented by the establishment for students' reading learning?

1	2	3	4

Did the planning of the additional strategies implemented during the quarantine align with the objectives of the Leo Primero Plan or the plan used in 2020?

1	Yes
2	No
99	Do not know/No response

Respondent Characterization

Sex of respondent:

1	Male
2	Female

What is your age?

99	Do not know/No response

What is the highest level of education you have attained?

1	Never attended
2	Incomplete primary education
3	Complete primary education
4	Incomplete secondary education
5	Complete secondary education
6	Incomplete higher-level technical education (1 to 3-year programs)
7	Complete higher-level technical education (1 to 3-year programs)
8	Incomplete higher education (4 or more-year programs)
9	Complete higher education (4 or more-year programs)
10	Incomplete postgraduate education

11	Complete postgraduate education
99	Do not know/No response

Do you have any specialization studies?

1	Yes	Pass to Question 18
2	No	Pass to Question 19
99	Do not know/No response	

What is the level of this specialization?

1	Course
2	Diploma
3	Master's
4	Doctorate
5	Other, please specify:
99	Do not know/No response

Did you take the Inicia Test?

1	Yes
2	No
99	Do not know/No response

As a teacher, have you participated in other programs offered by the Ministry of Education?

1	Yes (please specify):
2	No
99	Do not know/No response

What other roles or positions do you hold within this educational institution besides being a teacher?

99	Need to know/No response.

Is there a classroom support professional present in your classroom?

1	Yes
2	No
99	Do not know/No response

For how many years have you been employed as a teacher?

99	Need to know/No response.

For how many years have you been teaching at this establishment?

99	Need to know/No response.

Thank you for your participation in this survey! If you have any further questions or need assistance, feel free to ask.

Appendix 2

Items of the Instrument

1	Development of Comprehension Reading in 1st-Grade Text Strategies	Scale
1.1	I focused on strengthening and reinforcing students' understanding of the alphabetic principle.	Frequency Scale
1.2	I engaged in activities aimed at practicing and reinforcing students' reading decoding skills.	Frequency Scale
1.3	I organized activities to enhance phonological awareness, encompassing syllable separation, word recognition, and awareness of reading pauses.	Frequency Scale
1.4	During class, I facilitated comprehension reading practice by posing questions to students about the content of the texts they read.	Frequency Scale
1.5	I organized activities aimed at improving writing skills, such as dictations and free writing exercises.	Frequency Scale
1.6	I fostered student participation by encouraging them to alternate reading text excerpts with their classmates.	Frequency Scale
1.7	Students demonstrated the ability to identify both the shape and sound of letters.	Agreement Level Scale
1.8	Approximately 80% of students exhibit proficiency in identifying the sounds comprising words (phonological awareness), recognizing, separating, and combining phonemes and syllables.	Agreement Level Scale
1.9	Students demonstrated fluent reading skills, accurately pronouncing each word in a short text (3 lines), with occasional self-correction and proper respect for punctuation.	Agreement Level Scale
1.10	Students adeptly utilize all vocabulary words from the flashcards provided in the Leo Primero Plan.	Agreement Level Scale
1.11	After reading a familiar text, students display the ability to identify the central ideas or events presented in the text read during class.	Agreement Level Scale
1.12	Students can independently compose simple and concise texts.	Agreement Level Scale
1.13	I paid attention to students' time on each activity to organize the class.	Frequency Scale
1.14	The implementation of the Leo Primero Plan effectively supported students in their reading-learning process.	Agreement Level Scale
1.15	The Leo Primero Plan significantly contributed to students' enhancement in reading comprehension.	Agreement Level Scale
2	The Impact of Technological Innovation on Actors and Learning	
2.1	I incorporated content from the digital resources of the Leo Primero Plan (website, public classes, digital library, Leo Primero app, audiobooks) into activities designed to enhance students' reading learning.	Frequency Scale
2.2	I used information from videos, websites, blogs, scientific publications, and books to design my classes and acquire new educational strategies.	Frequency Scale
2.3	I used information from the digital resources of the Leo Primero Plan (website, public classes, digital library, Leo Primero app, audiobooks, strategy videos, and class videos) for my professional development.	Frequency Scale
2.4	I stayed updated on Leo Primero Plan updates through channels such as the website, social media, email, or WhatsApp.	Frequency Scale
2.5	My ability to locate information in digital resources for my professional development and class planning has improved since the beginning of this school year.	Agreement Level Scale
2.6	Insights gleaned from the exit tickets in the Leo and Sumo Primero monitoring app have been instrumental in pinpointing students' reading strengths and weaknesses.	Agreement Level Scale
3	The Social Adoption of the Program by Stakeholders	
3.1	I utilized the didactic guide provided by the Leo Primero Plan for planning my classes.	Frequency Scale
3.2	I employed resources from the Leo Primero Plan (school textbooks, shared reading sheets, public classes, etc.) to prepare lessons for first-grade students.	Frequency Scale
3.3	Class content follows the sequence outlined in the Leo Primero Plan school textbook.	Agreement Level Scale

3.4	The educational institution has promoted the utilization of Leo Primero Plan digital resources (such as public class videos and strategy videos) among the teaching team.	Agreement Level Scale
3.5	The resources provided by the Leo Primero Plan have aided in monitoring students' reading comprehension levels.	Agreement Level Scale
3.6	I participated in webinars and accessed video streams provided by the Leo Primero Plan.	Frequency Scale
3.7	I documented students' progress on the class progress tracking platform.	Frequency Scale
3.8	I revisited the Leo Primero Plan's video capsules and public classes to improve my reading teaching strategies.	Frequency Scale
3.9	The monitoring platform and unit tests have assisted in tracking the content to be taught to students.	Agreement Level Scale
3.10	My proficiency with platforms such as WhatsApp, the website, webinars, and video streaming under the Leo Primero Plan has noticeably improved since the start of this school year.	Agreement Level Scale
3.11	I actively engaged in school meetings, where I received comprehensive information regarding the Leo Primero Plan, including usage guidelines, access to videos, documents, and conference schedules.	Frequency Scale
3.13	In the upcoming years, I plan to utilize reading teaching strategies that were either developed or reinforced during the year 2021.	Agreement Level Scale
3.14	I collaborated with fellow members of the educational community, including headteachers and other teachers, to reflect on reading teaching strategies and enhance pedagogical practices.	Frequency Scale
3.15	I actively engaged in and contributed to initiatives related to Leo Primero and reading. I demonstrated a keen interest in acquiring knowledge, shared anecdotes, linked concepts to current events, offered commentary on news and developments, and celebrated students' accomplishments.	Agreement Level Scale
3.16	Learning strategies and the implementation of the Leo Primero Plan were customized to accommodate the unique needs of each student.	Agreement Level Scale
3.17	I organized group reading sessions with students and offered constructive feedback on their reading skills.	Frequency Scale
3.18	Pedagogical decisions were collaboratively made by the academic team.	Agreement Level Scale

Appendix 3

Experts

Expert	Education	Experience
Andrea Rolla	Primary Education Teacher from Princeton University. Ph.D. in Education from Harvard University, USA.	Associate Researcher at the Center on the Developing Child and the Faculty of Education at Harvard University and General Director of the Un Buen Comienzo project
Carmen Sotomayor	High School Teacher in Spanish, University of Chile. Ph.D. in Mother Tongue Pedagogy from the Catholic University of Louvain, Belgium.	Researcher in Teaching and Learning for CIAE-UChile.
Fernando Salamanca	University of Chile. Master's in planning and Regional Development, PUC. Ph.D. in Planning Studies from the University of London (England).	Expert in the Evaluation of Public Programs and Institutions. Academic at the University of Chile.
María Elsa Porta	Speech Therapist, Universidad de Aconcagua, Argentina.	Independent Researcher for the National Council for Scientific and Technical Research (CONICET). Researcher at the Institute of Linguistics (Universidad Nacional de Cuyo).
Andrea Cocio	Degree in Spanish Language Pedagogy, PUC. Ph.D. in Education Sciences, PUC.	Researcher in Reading Learning for the Department of Languages and Translation at the Catholic University of Temuco.
Marcela Bertoglio	Early Childhood Educator; Master's in Curriculum and Educational Administration. Ph.D. in Education.	Director of the School of Early Childhood Education at the Catholic University of Maule.
Adriana Gaete	High School Teacher in History, Geography, and Civic Education, PUC. Postgraduate degree in Guidance, Family, and Education (Universidad de los Andes).	Executive Vice President of JUNJI. Former Director of the Belén Educa Foundation.
Bárbara Eyzaguirre	Psychologist, PUC.	Researcher at CEP. Co-founder of Soy Astoreca and Advisor to the Board of the Foundation.
Cristina Sepúlveda	Psychologist (University of the Frontier). Diploma in School Psychology for Teacher Management.	Professional in Curriculum and Evaluation, Ministry of Education.
Pelusa Orellana	English Teacher, PUC. Master in Instructional Leadership, The University of Alabama, USA. Ph.D. in Education: Early Childhood, Families, and Literacy, The University of North Carolina at Chapel Hill, USA.	Vice Dean of Research, Faculty of Education, University of Los Andes.
Roberto Araya	Electrical Engineer, University of California, USA. Master's in science with a focus on Applied Mathematics, University of Chile. Ph.D. in Electrical Engineering, University of California, Los Angeles, USA.	Full Professor, Institute of Advanced Studies in Education (CIAE), University of Chile.