



This electronic thesis or dissertation has been downloaded from the University of Bristol Research Portal, <http://research-information.bristol.ac.uk>

Author:
Cepeda, Tamara

Title:
Assessment "Made in Chile"

An Exploratory Study on School Based Assessment Practices in one Chilean Secondary School.

General rights

Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode>. This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy

Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited on the University of Bristol Research Portal. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

- Your contact details
- Bibliographic details for the item, including a URL
- An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.



This electronic thesis or dissertation has been downloaded from the University of Bristol Research Portal, <http://research-information.bristol.ac.uk>

Author:

Cepeda, Tamara

Title:

Assessment "Made in Chile": An Exploratory Study on School Based Assessment Practices in one Chilean Secondary School.

General rights

Access to the thesis is subject to the Creative Commons Attribution - NonCommercial-No Derivatives 4.0 International Public License. A copy of this may be found at <https://creativecommons.org/licenses/by-nc-nd/4.0/legalcode> This license sets out your rights and the restrictions that apply to your access to the thesis so it is important you read this before proceeding.

Take down policy

Some pages of this thesis may have been removed for copyright restrictions prior to having it been deposited on the University of Bristol Research Portal. However, if you have discovered material within the thesis that you consider to be unlawful e.g. breaches of copyright (either yours or that of a third party) or any other law, including but not limited to those relating to patent, trademark, confidentiality, data protection, obscenity, defamation, libel, then please contact collections-metadata@bristol.ac.uk and include the following information in your message:

- Your contact details
- Bibliographic details for the item, including a URL
- An outline nature of the complaint

Your claim will be investigated and, where appropriate, the item in question will be removed from public view as soon as possible.



**Assessment “Made in Chile”:
An Exploratory Study on School Based Assessment Practices in one Chilean
Secondary School.**

By Tamara Cepeda Acevedo

A dissertation submitted to the University of Bristol in accordance with the
requirements of the Degree of Doctor of Philosophy in the Faculty of Social
Sciences and Law

School of Education, University of Bristol.

2024

Word count: 84,003.

Abstract

This study is a detailed account of School Based Assessment practices in one secondary school in Chile, which explores and questions the purposes, methods, results and fairness underpinning current assessment practices. The research design of the study is informed by a conceptual framework based around the international literature on key features of the nature, purposes and outcomes of SBA, and then tailored to the specific nuances found within the Chilean system, policy and contextual characteristics. The study is underpinned by a constructivist ontology and an interpretivist epistemology. This supports an exploratory qualitative study using documentary analysis and semi-structured interviews as methodology, with ten teacher and ten student interviews from one subsidised school in San Bernardo, Santiago, that were analysed using thematic analysis.

The findings of this study show that assessment purposes are highly impacted by the need to cover curriculum contents, with classic paper and pencil assessment methods, such as tests, that suit teachers' lack of time. Teachers also make summative use of formative strategies, as grades production is a main objective of SBA. Following that fact, feedback is scarce for students. Finally, a general perspective of an unfair system is depicted by teachers and students, mainly because of the impact of the high stakes, the inequality of the country, and also historical structural conditions within the school. The considerable impact that grades have on students' lives in terms of being promoted from one level to another, university entrance opportunities are high stakes and some research has been highly critical, emphasising the instrumentalist logic which reduces the educational process to numbers.

This study although small-scale contributes new knowledge by informing policy debate with a much-needed bottom-up approach which by emphasising stakeholders and the school's reality, can potentially support and serve as a basis for policy changes in the future.

Acknowledgments

*Because I wrote, because I wrote, I am alive.
(Enrique Lihn, Chilean poet).*

Firstly, I want to thank my daughter, Romina. She has been benefited and victim, partner and witness of these years of study at the University of Bristol, so far away from home in our family of two. Thanks for your company, patience, infinite love and the wonderful human being you became in the process. I could not be prouder of who you are. I love you infinitely.

Even when COVID-19 appeared so early in my PhD, we were not alone because incredible people came across our journey. Thanks to Pilar Wenzel for her friendship, where I can always talk about everything, freely, and for your constant encouragement in difficult times.

To Milagros Terrones and Ximena Ibinarriaga, for creating a Latin family with us full of laugh, dance, and good food but, above all, a friendship that endured despite illness and distance.

To Hugo Parra and his amazing wife Catalina. Their family always embraced us with love, kindness, and family warmth. Hugo is a philosophy teacher (like me) and a colleague who was always there to discuss my ideas and to support my progress (and lack of it) with the humility of those, like us, who never had anything and made it from the lower socioeconomic backgrounds.

To Marcela Ramos and Manuel who grew up with Romina and even today, from Glasgow and Santiago, keep in touch. Marcela has always supported every phase encouraging me with her experience and advise, she never doubted that I could make it. I did it several times.

To Mark Biram, a beautiful British soul who happened to speak Spanish and had the difficult task of proofreading my work along these years. I am far smarter and wittier in Spanish, sorry Mark.

To my supervisors, Professor Sally Thomas and Professor Shelley McKeown-Jones, for being by my side and helping me to finish this work despite all the challenges that I faced, in the good and bad times. Thanks for your patience and constant encouragement.

To Dr. Raúl Velozo Farías, who believed fiercely in me during my undergraduate studies, and Dr. Shawanda Stockfelt, for labelling me as a “prime candidate” for a PhD during my Master’s degree. Both of you made me believe.

Thanks to ANID (National Association of Research and Development, Chilean Government) for funding my research during these years, making possible each page and word.

Finally, I want to thank in my mother tongue to those who cannot read this thesis, because we live in a country where English is a privilege: a mis padres, que siempre se sacrificaron por mí, para que no me faltara nada aún en los tiempos más duros y difíciles. Por toda la ayuda que me dan, por creer en mi talento desde siempre y no dudar de mí. A mi hermana Romina que, aunque nos haya dejado tan joven, me hace siempre querer tener una vida de la que ella hubiese estado orgullosa. Lo poco que he podido escribir desde tu muerte, siempre es para ti.

A mis tíos Manuela y Jorge y primos Ariel y Fer, por recibirnos en su hogar en nuestras idas y venidas, siempre con amor y risas.

A mi amiga Ximena Mellado, por cuidar de mi hija mientras me encerraba a escribir. A mis amigas, Melissa Salgado y Yohana San Martín. Las tres siempre se han sentido como una extensión de la hermana que la vida me quitó. Por quererme, así como soy y soportar mis distancias y viajes, para recibirme siempre de vuelta con un abrazo (de esos que sí me gustan).

Para terminar, gracias a todos los y las profesores y estudiantes que han nutrido mi vida y experiencia durante 20 años en distintos colegios y lugares de Chile. ¡Gracias totales! I would not have made it without you all.

Author's declaration

I declare that the work in this dissertation was carried out in accordance with the requirements of the University's *Regulations and Code of Practice for Research Degree Programmes* and that it has not been submitted for any other academic award. Except where indicated by specific reference in the text, the work is the candidate's own work. Work done in collaboration with, or with the assistance of, others, is indicated as such. Any views expressed in the dissertation are those of the author.

SIGNED: Tamara Fabiola Cepeda Acevedo DATE: November 25, 2024.

Table of contents

Contenido

Abstract	2
Acknowledgments	3
Author's declaration	4
Table of contents	5
List of abbreviations	10
List of figures	11
List of tables	12
Chapter 1: Introduction	13
1.1 Research problem and background	13
1.2 Rationale	18
1.2.1 Academic rationale.....	18
1.2.2 Local rationale	21
1.2.3 Personal rationale	24
1.3 Research aim and research questions.....	25
1.3.1 Research Aim.....	25
1.3.2 Research objectives.....	25
1.3.3 Research questions	26
1.4 Overview of conceptual framework.....	26
1.5 Overview of Methodology	29
1.6 Overview of the thesis.....	30
Chapter 2: International literature review on SBA	32
2.1 Introduction.....	32
2.2 Approach to Conceptual Framework	33
2.3 Socio-culturally responsive assessment as a response to cultural scripts.....	34
2.4 International experiences on SBA exhibiting the relevance of cultural scripts.	37
2.5 School based assessment: characteristics.....	39
2.5.1 Definition, strengths, and weaknesses of summative assessment.....	41
2.5.2 Definition, strengths and weaknesses of formative assessment	42
2.5.3 Problematizing formative and summative	43
2.5.4 Assessment purposes: why assess students?	46
2.6 The relevance of teachers' and students' perspectives on SBA	48

2.7 Assessment methods (RQ1)	51
2.7.1 Assessment methods and external testing	54
2.8 Information provided by the assessment: perspectives on grading and feedback (RQ2)	56
2.8.1 Grades on SBA	57
2.8.2 Formative Feedback on SBA.....	60
2.9 Fairness of the assessment process (RQ3)	62
Chapter 3: School Based Assessment in Latin America and Chile	68
3.1 Introduction.....	68
3.2 Evaluation in Latin America in its own terms.....	68
3.3 Chilean educational system	74
3.3.1 External assessment: SIMCE test (National System of Learning Results Evaluation)	76
3.3.2 PAES test for Secondary School leavers (Admission Test for Higher Education)	76
3.3.3 National curriculum in Chile: subjects' plans and programs.....	78
3.3.4 The assessment policy issued by Mineduc.....	79
3.3.5 Critics on literature around the national curriculum and its contents	82
3.3.6 Problematics around the assessment decree and assessment practice.....	84
3.3.7 Teachers working conditions in Chile: work overload and lack of time.....	89
3.4 Topics identified in the literature review that guided the research design and methodology	91
Chapter 4: Research design and methodology	93
4.1 Introduction.....	93
4.2 Philosophical Approach to Research Design	93
4.3 Research design.....	95
4.4 Methods: documents analysis and interviews.....	96
4.4.1 Documents	97
4.4.2 Interviews.....	97
4.5 Preliminary study and research design	98
4.5.1 Preliminary study documents and decision-making	98
4.5.2 Interviews in preliminary study and decisions made for the main study	99
4.6 Sampling strategy.....	104
4.6.1 School district: San Bernardo	104
4.6.2 The School	106
4.6.3 Documents sample.....	108
4.6.4 Sample for interviews' participants	110
4.7 Reflexive account of research design.....	112

4.8 Data collection for the main study.....	113
4.8.1 Data collection for documents (RQ1).....	113
4.8.2 Interviews for main study (RQ1, RQ2 and RQ3).....	114
4.9 Introduction to Data Analysis.....	116
4.9.1 Documents analysis (RQ1).....	118
4.9.2 Interviews’ transcripts analysis (RQ1, 2 and 3)	121
4.9.3 Data Triangulation.....	123
4.10 Analysis of complete data set codes: themes and subthemes.	124
4.11 Themes and subthemes from data analysis.....	126
4.11.1 General Themes/Subthemes from Documents	126
4.11.2 General themes/Subthemes from Interviews’ transcripts.....	127
4.12 Overall themes map.....	128
4.13 Themes and sub-themes overview	129
4.14 Ethical issues	131
4.15 Reflexivity and positionality	133
4.16 Findings chapter’s structure.....	136
Chapter 5: Assessment purposes and methods.	138
5.1 Introduction.....	138
5.2 Assessment purposes on the school assessment decree	138
5.3 Assessment methods on the school assessment decree	140
5.4 Assessment methods on schoolbooks registers (PD01 to PD08).....	141
5.5 Assessment purposes following teachers’ and students’ interviews	147
5.5.1 The role of memory, writing, and notebooks to accomplish the assessment purposes through students’ voices.....	151
5.6 Assessment methods in interviews.....	152
5.6.1 Graded formative work as an assessment method	154
5.6.2 Use of standardised test formats as an assessment method in interviews.....	156
5.7 Irregular use of formative assessment: not for everyone, not always.	157
5.7.1 Summative use of formative: rubrics for grading activities	160
5.7.2 Accumulative grades: formative made in Chile	163
Chapter 6: Assessment information: grades and feedback	166
6.1 Introduction.....	166
6.2 Grades on the school documents as background information	166
6.3 Summative and grading culture through teachers’ and students’ voices	168
6.3.1 Negative consequences of grading	170

6.3.2 Grades differences and achievement: motivation, expectations, and cultural capital .	173
6.3.3 High stakes: repeating/promotion and university entrance	176
6.3.4 Grades manipulation through score pollution	184
6.4 Feedback on documents and through teachers' and students' voices.....	185
Chapter 7: Assessment and Fairness.....	191
7.1 Introduction.....	191
7.2 Social inequality and the impact of the high stakes of the educational system	192
7.3 A socio-culturally responsive assessment: the fairness of asking teachers for formative and summative practices	194
7.4 A Structural condition of teachers' lack of time and work overload: the constrained freedom	198
Chapter 8: Discussion.....	202
8.1 Introduction.....	202
8.2 Addressing RQ1 on assessment purposes and methods: policy and practice disconnection on SBA due to lack of attention to the school conditions and needs.	202
8.2.1 RQ1a. What are the assessment purposes used by teachers following the school's documents and teachers' and students' perspectives?.....	203
8.2.2 RQ1b What are the assessment methods used, following school documents and teachers' and students' perspectives?.....	206
8.3 Addressing RQ2 on the information provided by assessment: numbers, humans, creativity, and words. 208	
8.3.1 SBA and grades: The Power of numbers vs. formative spaces	208
8.3.2 The creativity of grades manipulation	209
8.3.3 SBA and feedback: numbers vs. words	211
8.4 Addressing RQ3. What aspects of the school assessment are linked by teachers and students with the fairness concept in one Chilean secondary school? The overwhelming sense of unfairness	212
8.4.1 An unfair society, an unfair school	213
8.4.3 A socio-culturally responsive assessment made in Chile	214
Chapter 9: Conclusion	216
9.1 Introduction.....	216
9.2 General conclusion	216
9.3 Contributions to knowledge.....	218
9.3.1 The relevance of teachers' and students' voices within educational research	219
9.4 Implications	219
9.4.1 Implications for research.....	219
9.4.2 Implications for Schools	220

9.4.3 Implications for Policy	221
9.5 Limitations and Reflexive Account	223
9.6 Recommendations for future research on SBA.....	225
References	226
Appendixes	242
Appendix A: Some examples of countries using SBA and some of its characteristics.	242
Appendix B: Documents reviewed during preliminary study	245
Appendix C: Students' interview schedule (showing changes implemented after preliminary study) 246	
Appendix D: Students' final interview guide for main study	248
Appendix E: Teachers' interview schedule (showing changes implemented after preliminary study) 249	
Appendix F: Teachers' interview guide for fieldwork	250
Appendix G: SoE Research Ethics form	251
Appendix H: SoE Research Ethics form (updated version).....	262
Appendix I: Access to School consent	265
Appendix J: Information sheets	266
Appendix K: Interview Consent forms.....	283
Appendix L: Ethics form approval	291
Appendix M: Themes, subthemes and topics from interviews' transcriptions (RQ1, 2 and 3) ..	292
Appendix N: Excerpt of teacher's interview transcription (original version in Spanish).	293
Appendix O: Excerpt of student's interview transcription (original version in Spanish).	299

List of abbreviations

ACACA	The Australasian Curriculum, Assessment and Certification Authorities.
ACER	Australian Council for Educational Research.
ARG	Assessment Reform Group (England).
BERA	British Educational Research Association.
BCN	Biblioteca del Congreso Nacional, Chile.
CIAE	Centre for Advanced Research on Education, University of Chile. Centro de Investigación Avanzada en Educación, Universidad de Chile.
CPEIP	Centre for Improvement, Experimentation and Pedagogical Research. Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas, Chile.
DEMRE	Department of Assessment, Measurement and Educational Register, Chile. Departamento de Evaluación, Medición y Registro Escolar, Chile.
HKEAA	Hong Kong Examinations and Assessment Authority.
NEM	Secondary School Grades. Notas de Enseñanza Media.
OECD	Organization for Economic Cooperation and Development.
PAES	Admission Test for Higher Education. Prueba de Admisión a la Educación Superior.
SEP	Preferential School Subsidy. Subvención Escolar Preferencial.
SIMCE	National System of Learning Results Evaluation. Sistema Nacional de Evaluación de Resultados de Aprendizaje.
PSU	University Selection Test, 2019. Prueba de Selección Universitaria.
PTU	Transitory University Test. Prueba de Transición Universitaria.
SBA	School Based Assessment.
SWAAP	System Wide Analysis of Assessment Practices.
WHO	World Health Organization.

List of figures

- Figure 1: Image of preliminary study work notes analysing the schoolbooks.
- Figure 2: Printed interview transcripts folder.
- Figure 3: Example of NVivo coding for documents.
- Figure 4: Example of first transcription analysis by hand identifying patterns (original version in Spanish).
- Figure 5: Example of NVivo coding for interview's transcriptions.
- Figure 6: Overall themes and subthemes map obtained from the complete data set.
- Figure 7: Detail of assessment methods used by subject and their type (following word patterns).
- Figure 8: Example of multiple-choice test question (original version in Spanish).
- Figure 9: Example of multiple-choice online test question, used during 2020 (original version in Spanish).
- Figure 10: Example of in-class activity in Mathematics (original version in Spanish).
- Figure 11: Example of points-based and graded rubric used in the school.

List of tables

- Table 1: Overview of topics from the literature review that guided the research design and methodology.
- Table 2: Sample for preliminary study interviews.
- Table 3: Overview of sample for documents (RQ1).
- Table 4: Sample for students in the School secondary level participating in the study.
- Table 5: Sample for teachers' interviews in the School secondary school level.
- Table 6: Data set for analysis.
- Table 7: Labelling of files based on the schoolbooks registers.
- Table 8: Original documents analysed.
- Table 9: Example of data analysis and triangulation.
- Table 10: Main themes and subthemes for document analysis.
- Table 11: Main themes and subthemes for interviews' transcripts analysis.
- Table 12: Analysis of schoolbooks registers: themes, subthemes, description and coded words for assessment methods.
- Table 13: Grades registered in schoolbooks per class and subject in 2019 academic year.

Chapter 1: Introduction

“In an ideal world, each and every aspect of educational assessment practice would operate in harmony with the ultimate educational objective: the cultivation of valuable learning. Inevitably, though, the real world is far from idea, and disharmonies abound”.
Newton, 2014.

1.1 Research problem and background

The present time is labelled as an era of strong assessment focus (Hirsh, 2020), where School Based Assessment (SBA hereafter) is regarded as a ‘worldwide movement’ to promote students’ learning (Tong & Adamson, 2015). SBA provides a practical concept implying an integrative perspective, incorporating both formative and summative aspects of assessment undertaken internally within schools. Thus, students are assessed by their teachers in each learning subject (HKEAA, 2013). SBA involves the idea that regular classroom-based assessment (understood simply as occurring in the classroom space) provides a reliable and important indicator of student learning, while also contributing feedback for future learning (Brown & Hirschfeld, 2008). Therefore, the main idea of SBA is the respect and trust in teachers’ abilities in judging students’ learning (Cumming & Maxwell, 2004). Teachers use the curriculum guidelines to generate a variety of classroom-based assessment procedures to judge and report student progress and achievement (Cumming & Maxwell, 2004). This variety of procedures used in SBA is greater than the procedures used for external assessment strategies (Yates & Johnston, 2018). This is important because alternative national or external examinations provide judgements that are not based on day-to-day life within the classroom, where relevant information that appears when the student’s learning is happening, can be obtained (Pellegrino et al., 2001). SBA considers that centralised summative examination is only partially useful because strategies that function to help students’ learning are also needed for it to yield real progress (Othman, 2019). This study seeks to explore teachers’ and students’ perspectives on SBA in one secondary school in Chile.

Students are assessed to know how their learning process is going on. In this vein, a close analysis of assessment practices embedded within an educational system is considered to be an effective contributor to the improvement of student’s learning outcomes and achievement (System Wide Analysis of Assessment Practices, SWAAP report, 2017). Following this, the

importance of looking at assessment practices is related with the students' achievement, which seems to be a recurrent concept when talking about assessment. At this point, fundamentally, what does achievement mean? It can be understood as performance and learning objectives mastered (McMillan, 2001). As such, achievement is the knowledge obtained or skills mastered in school subjects, evidenced through assessment activities that teachers apply. This process allows teachers to build a holistic picture of students' achievement across a full range of activities and assessment procedures (Harlen, 2004). Achievement requires performance or demonstration that is analysed by different practices and methods that teachers apply but especially, during the assessment. Consequently, the students' assessment made inside the school by teachers can be understood as a complex process of collection, analysis and assessment of evidence about the teaching and learning performance and process (Zhao et al., 2017).

When the characteristics and purposes of the assessment process are questioned, a dichotomy between 'summative' and 'formative' is often invoked (Davies et al.; 2012). But what do these concepts mean? Summative assessment appears to be focused on reporting in a systematic way in order to draw inferences about students' learning and achievement and also is generally associated in Europe with the 'assessment of learning' (Assessment Reform Group, ARG, 2006; Harlen, 2004). This assessment could be summarised as 'checking up' on students' learning, for example via externally set and marked examinations or internally generating an overview of achievements derived from information gathered over a period of time during the teaching process (Harlen et al., 1994). Students' knowledge or skills are usually designated by scores or marks that are assigned by external markers or teachers, allowing them to rank and certificate as well, typically having high-stakes consequences (Cizek, 1996). However, despite the traditional prevalence of summative assessment, in recent years, there has been a considerable shift towards explicitly incorporating formative classroom assessment practices. This was true especially in Western countries (Kennedy et al., 2008), but more recently it has been widely adopted in different countries worldwide. The formative assessment concept is presented by some as being full of virtues, such as being a strategy to improve and adjust the educational process by making changes in order to help students to learn, to accomplish the objectives' plans as part of everyday practice (ARG, 2002; Black & Wiliam, 1998; Klenowski, 2009; Lambert & Lines, 2000; Mineduc, 2018b; Penuel & Shepard, 2016). It can also serve a knowledge monitoring purpose, which aims to identify errors in cognitive processing which prevent students from attaining learning progression (Heritage, 2008). Formative assessment has also been linked with the concept of 'assessment for

learning' and with sociocultural theory, reinforcing the relevance of the learning process as a collective task (Baird et al., 2004; Elwood & Murphy, 2015; Harlen, 2004).

So, what does SBA imply? Remesal (2011) states that assessment within the classroom fulfils two principal functions: firstly, it promotes reflection and change by monitoring teaching and learning; and secondly, it allows certification that serves accountability purposes. Within this logic, the formative component is intertwined with a low-stakes system or no further consequences. However, the summative component is often linked with high stakes, with the implication of punitive consequences for non-achievement or lower achievement. This means, educational or occupational: university selection, promotion, and certification (Perassi, 2008; Stobart & Eggen, 2012). Therefore, SBA involves information given to students, in the form of feedback and grades, implying important consequences for students' lives, especially within high-stake spaces. Harlen (2004), the ARG in England (2006), and Smyth and Banks (2012) showed that usually high-stakes systems and external testing can harm students' motivation for learning and also the breadth and depth of curriculum and pedagogy, promoting shallow and superficial learning. Therefore, the omnipresence of grading practices within the SBA process is also an important element to look at.

As suggested by a systematic review of the evidence of assessment reliability and validity used for summative purposes, 20 years ago, Harlen (2004) stated that more studies are needed to understand how teachers utilise assessment for different purposes, considering the evidence that they use and its interpretation. Nevertheless, literature examining specifically students' classroom experience on assessment is still regarded as less prominent (Arnold, 2022). This refers to the need to look at the day-to-day assessment of students made inside the classroom over the expanded and common study of external standardised tests used for accountability purposes (Stiggins & Conklin, 1992). In this regard, studies like Harlen's (2004), recognised that attempts to develop and implement classroom-based assessment systems have not received even a small part of the resources used to run external examinations in the form of standardised tests.

The omnipresent concepts of the dichotomy of 'formative' and 'summative' assessment are considered to critically reflect upon the interaction between them. Their purposes and the contexts where are used, why to prefer one over the other, and why they need to be used at all are relevant topics to be discussed (Kennedy et al., 2008). Formative assessment appears as a complex educational and social process but so is SBA in general. As it is impacted by different ideas within distinct contexts, this will inevitably have an impact on decisions on what the assessment focus should be in each school. The complexity of dichotomising the assessment as to

being formative and summative has been highlighted by several studies, although sometimes it is difficult to identify clear distinctions between them (Davies et al.; 2012). For that reason, Harlen (2013) proposed that both concepts are more useful when seen as dimensions. The literature also recognises that tensions arise from using both approaches and from the possibility of being used interchangeably – that is to say, summative used for formative and vice versa- (Yates & Johnston, 2018). These complexities will be further explored in Chapter 2.

The intensified focus on assessment in conjunction with policy changes over the last decade has led teachers to constantly adjust their pedagogical and ethical practices within the classroom space (Hirsh, 2020). It must be noted that how the characteristics of the assessment systems interact with each other and are presented within an educational system, implies different consequences that should be taken into account. Consequently, when achievement is considered, ethical principles should guide the behaviour of teachers and students in their everyday lives, in the sense that it could be a practice that entails results that could bring about ‘positive’ or ‘negative’ consequences in terms of students’ growth (Biesta, 2004). Teachers’ ethos when gathering information about students’ achievement should empower an assessment design and implementation based on the assessment purpose and students’ needs (Cizek, 1996). Ethical practice in the assessment dynamic created within schools should be conceived as protecting students’ basic rights not to be harmed (socially, psychologically, emotionally) and not to be coerced or manipulated (Estaji, 2011; Green et al., 2007; Pope et al., 2009). Several studies report the use of fear of assessment, which can certainly harm students’ motivation (Flitcroft & Woods, 2008). However, little research has studied the concept from teachers’ and students’ perspectives within the classroom reality when the fairness of the assessment is considered (Rasooli et al., 2023).

The assessment practice, then, should be also proposed with fairness in mind, a notion that is highly present and well-developed in the Latin American Spanish-language literature on evaluation (Santos, 2003). The system is linked, in this regard, with a broader goal of encouraging ethics and fairness ‘to ensure that no harm is done to individuals or groups as a result of an assessment and that no participants are unfairly disadvantaged’ (SWAAP, 2017, p.9). In this sense, the consequences of the assessment procedures inside the school need to be acknowledged as multi-faceted (Kožuh, 2019). Within high-stakes systems, it is abundantly clear that promotion, university entrance, and achievement stereotypes are related to the assessment consequences and outcomes. As Estaji (2011) states, the higher the stakes, the greater the likelihood of cheating

and unethical practices taking hold. This further underlines the importance of assessment practices being both fair and equitable. In this light, the differences in terms of outcomes take on greater relevance, as the grades resulting from assessment practices would reveal a pattern of differences among students' outcomes which are worth analysing as they reveal a culture that both rejects and punishes errors, without necessarily offering constructive suggestions vis-à-vis how to improve (Sánchez & Jara, 2019). Currently in some contexts, and despite an increasing body of research on formative assessment, assessment and testing strategies still tend to reflect the old primacy attached to assigning grades (Haertel et al., 1984; Hirsh, 2020; Fleer, 2015; Rasooli et al., 2023). Results and grades not only matter for students, parents, teachers and schools, but they also become the sole objective and language in a profoundly instrumentalised education system (Baird et al., 2014; Flórez-Petour, 2014). To ignore the relevance of assigning grades in this high-stakes context is to disregard a crucial factor, and thus could even be regarded as irresponsible (Brooks, 2002). As Santos (2003) stated, our understanding of the role of education cannot and must not be reduced to simply numbers and statistics. This also brings to the fore the deeper question about what matters, prompting distortion within the educational setting: are grades all that matter?

Back in the 1970s Bloom and colleagues already included in their studies the notion that low achievement, evidenced by assessment, has been generally linked with individualistic logics, for example, the students' rank position when compared to their classmates (Bloom et al., 1971). Phye (1996) established that this achievement monitoring through assessment activities implies that teachers acknowledge that the information resulting could have the unethical impact of embarrassing and labelling students. There was also evidence of teachers interacting differently with students that they perceive to be 'high' or 'low' achievers (Stiggins & Conklin, 1992), which is reinforced by contemporary Latin-American inquiry (Avendaño & Parada, 2012; Sánchez & Jara, 2019; Vera et al., 2017). Moreover, in this vein, Stiggins (2002) also highlights that the focus on formative aspects of assessment leaves the summative unquestioned, leaving some students to give up in hopelessness. The literature also presents considerable evidence of the negative impact of grades on students' school engagement at the secondary level (Poorthuis et al., 2015). Shepard (2019) and also Latin-American studies such as Londoño-Restrepo (2015) stated that grades could be used as rewards to motivate, moderate, and ultimately control students' behaviour, punishing them. This practice is far removed from the objectives that the assessment process is supposed to have, which should be to evaluate students' learning. Taken altogether, these distinct arguments

invite us to reflect on the positive and negative effects of assessment both on teaching and in the students' learning process.

The rationale for this study, then, is to explore in detail what SBA looks like in one Chilean secondary school. The focus will be placed on three main themes in regards to teachers' and students' perspectives on: assessment purposes and methods used and the reasons supporting their design. Information provided by the assessment process, in terms of grades and feedback, will also be discussed; and so will be the link between assessment practice with fairness. These areas of inquiry have already been highlighted by authors like Beets (2012), who have stressed the relevance of understanding the assessment purposes, methods, and the effects of the assessment procedures within schools to improve them. Therefore, this study will contribute to the generation of new knowledge about i) current school assessment practices in Chile and the concepts used by teachers and students to refer to the process, ii) a reflection on grades and feedback in one school within a high stake system, including concepts of promotion and university entrance in Chile, and finally iii) to promote a reflection based on teachers' and students' perspectives about the link between the assessment system and its omnipresent ethical aspects, mainly in terms of fairness.

1.2 Rationale

This section will outline the academic, local, and personal rationale for researching SBA in Chile, highlighting the relevance of teachers' and student's perspectives on the topic.

1.2.1 Academic rationale

Within the school context, teachers and students are those who embody the school assessment procedure. Therefore, they are the most important actors that should be studied to understand SBA. Several authors on assessment research have stressed the importance of considering teachers' and students' voices in relation to SBA, despite this being usually done focusing on only one of these groups (Baird et al., 2014; Brookhart, 2001; Brown, et al., 2008; Carless, 2005; Cheng et al., 2011; Marshal & Drummond, 2006; Monteiro et al., 2021; Sivenbring, 2019; Tapan, 2001; Torrance & Pryor, 2001; Willis, 2008). The lack of attention placed on teachers' and students' voices in assessment is remarkable, especially considering the worldwide tendency

to reform and change educational systems (Sivenbring, 2019). Studies within the classroom context are constantly regarded as remaining under-researched area (Baird et al., 2014; Young & Kim, 2010; Lambert & Lines, 2000). Hirsch (2020) stated: "Research that reflects experiences of those who are the subject of school reforms and teachers' instruction and assessment is, however, quite scarce" (p. 90). After the worldwide lockdowns, our knowledge on assessment has been put as "one of those things that you think you know what it is until you start to think really hard about it", highlighting the constant need to reflect on the assessment practice and its purposes (Research Schools Network, 2021).

Consequently, this research will be grounded on teachers' and students' perspectives on SBA, considering the main characteristics and also challenges of this type of assessment through literature on the topic relating the experience of different countries. Teachers' perspectives play an important part not just on how educational reforms are implemented in schools and classrooms (Brown et al., 2011). As they are important for all the educational process as front-line workers, the impact of their perspectives on assessment on classroom decisions and practice is also considerable (Remesal, 2010; Bonner et al., 2018). The assessment that occurs within the classroom involves different levels of teacher autonomy (Bonner et al., 2018; Swanson & Stevenson, 2002). This is interesting, especially for studies within educational contexts where they decide how to assess; such as the SBA where they are responsible for planning and creating their daily lessons, assignments, and assessments, guided by their perspectives (Bonner et al., 2018; Tong & Adamson, 2015). On the other hand, and as it has been stated in the previous section, students are the principal actors who live the process, including how their learning and results imply consequences for their future. The assessment perspectives then, will represent a central focus of this study, taking into account and questioning the profoundly top-down paradigm which has tended to exclude the voices of teachers and mainly students from studies on assessment (Cheng et al., 2011; De Luca et al., 2018; Sivenbring, 2019). This occurs especially when research on assessment tends to be primarily focused on technical, instrumental, and political aspects (Perassi, 2008). In these cases, the particularities of schools and their specific and perhaps even multiple realities inevitably leave an incomplete and thus deficient and unsatisfactory analysis in many respects.

On the other hand, the literature on students' perspectives on assessment has shown that these perspectives influence students' behaviour, learning, and achievement respectively (Brown, 2011). These elements are key within the educational world and often absent from research on

SBA. Existing literature on students' responses to formative assessment suggests that students' reactions and perspectives differ depending on the assessment purposes and methods, and also upon students' general perceptions on assessment, and the classroom culture towards that assessment (De Luca et al., 2018). This gives a glance at the relevant aspects to analyse when studying school assessment and also the relevance of considering students' perspectives. As an example of this interest, Remesal (2011) conducts interviews with teachers about their perceptions on assessment and its effect on teaching-learning, on students' accreditation of learning, and also on teachers' accountability. Similarly, it is worth noting a study of secondary school teachers' and students' perspectives on assessment in Finland, conducted by Mäkipää and Ouakrim-Soivio (2019). They asked about assessment practices that secondary teachers use, the assessment practices that students appreciate, the improvement that students propose to the assessment practices, and students' perspectives in understanding teachers' feedback. Expanding this focus, studies such as Pope et al (2009) go into depth by addressing the ethical implications of teachers' assessment, focusing on the difficult balance between grading activities and the improvement of students' learning. This study adds to research on SBA a more integrated view of key aspects of this form of assessment, in the particular context of a country with high levels of inequality and marketization of the educational setting (see local rationale in the following section). Therefore, following ideas from studies on teachers' and students' perspectives on SBA (Mäkipää & Ouakrim-Soivio, 2019; Remesal, 2011), this study will explore three main themes: assessment methods used by teachers and the reasons supporting their design (Bota & Tulbure, 2015; James & Pedder, 2006); perspectives on the information provided by the assessment, namely, grades and feedback (Black & William, 1998; Harris & Brown, 2009; Yates & Johnston, 2018), and the fairness of the assessment process (Brown, 2006, 2011; Garner et al., 2009; Mäkipää & Ouakrim-Soivio, 2019; Remesal, 2011; Yates & Johnston, 2018). The research conducted in this study will be located within international and local literature, placed within the context of educational policy that in some sense agrees but also comes into tension with the general thrust of this literature, and emphasises the importance of the regional and national specificities and events that configures the reality within schools (Brown, 2011; Kennedy et al., 2008). There is a significant difference pertaining to the number of studies on school assessment practices when compared to standardised and external assessment studies, that this study addresses. The relevance of research in this area and its resultant knowledge has been reinforced by studies such as the SWAAP report (2017), prepared by the Australian Council for Educational

Research (ACER), which states the urgent need for a process of analysis of how assessment practices are integrated into education systems. These practices, in the form of school assessment, determine the learning, achievement, and results of millions of students worldwide. This research brings an in-depth study of one of those schools, characterising SBA and how it is framed by school policies as it is lived day-to-day by teachers and students from their perspectives and also registered in different teacher-created documents. In this way, the study will contribute to new knowledge on SBA in three main areas: assessment purposes and methods, information provided by the assessment, and assessment fairness.

1.2.2 Local rationale

Chile is a particular example of a country with no external examinations applied to students to be promoted or to finish secondary school. The external examination is used only for accountability purposes and university entrance, which will be further explained in Chapter 3 (see p. 78). It is of utmost importance to mention that Chile is described by Stromquist and Sanyal (2013) as:

“Despite this apparent well-being, however, Chile faces a major problem regarding income distribution, as the richest 20% of Chileans capture 62.2% of the national income while the poorest 20% take 3.3% (Zibechi, 2012); expressed in other terms, the average income of the richest 20% of Chileans is 12 times that of the poorest 20% (World Bank, 2011)” (p.153).

This inequality is also contextualised within the educational reality, characterised by higher-than-average inequality internationally, neoliberal policies, marketization and a strong correlation between economic background and academic results, where better family income means better academic results and poorer family income equals lower academic results, as in most country contexts (Bellei & Vanni, 2015; Cabalin, 2012; García-Huidobro, 2007; Gysling, 2017; Oliva, 2008; Oliva & Gascón, 2016; Puga, 2011; Zancajo, 2019). This will be further explored in Chapter 3. Suffice to say here that the national reality will configure an interesting educational system to look at, especially for the use of SBA.

In the case of Chile, there is a knowledge gap as very few studies describe and explain the reality and lived experience of SBA in the schooling context. Only 7 studies were found that included teachers’ and/or students’ perspectives and practices on SBA in the country (Castillo et

al., 2023; Maldonado-Fuentes, 2021; Martínez, 2019; Mineduc, 2016; Peters & Contreras, 2019; Ravela et al., 2014; Vera et al., 2017). Ravela et al (2014) focused on a comparative view of assessment in four countries in Latin America (Chile, Colombia, Peru, and Uruguay). Prieto and Contreras, (2008) also conducted a study on teachers' conceptions that inform their assessment practices, however in a literature review style. Studies such as Sanchez and Jara (2019) are based on pre-service teachers' perspectives on learning assessment, thus, previous to the real experience within the classroom. More recently (2023), Castillo, Mejías, and Vásquez conducted an exploratory-quantitative study aiming to identify conceptions and strategies of formative assessment applied by teachers during the COVID-19 pandemic.

The grave difficulties in proposing constructive changes to tackle the educational situation in Chile are inextricably linked with and symptomatic of the dearth of literature related to assessment practices used with Chilean students in the school space, and considering teachers' and students' voices. Raczynski and Salinas (2008) stated that it was urgent to qualitatively research the educational process, especially with its involved actors inside the classroom. Thus, this thesis begins the urgent task of filling this lacuna and aims to foment further debate among Chilean scholars and beyond, as several countries worldwide are implementing systems that use formative and summative assessment, namely SBA, in different contexts but with similar aims, challenges, and difficulties.

Regarding the school system, this literature also evidences a significant increase in inequality of outcomes in the transition from primary to secondary education in Chile, a phenomenon that is inextricably linked to the market-orientated transformation of the system (Torché, 2005). Foregrounded by the omnipresence of summative aspects of assessment and numerical grades, students 'learn', 'pass', 'fail', repeat the year, finish secondary school, and obtain grades that allow them (or not) to enter university. Moreover, grades may also reflect a worrying and worldwide phenomenon: the low achievement that is implied by 'failing' a subject, repeating the academic year, and being labelled as non-capable (Bellei et al., 2003; Sánchez & Jara, 2019). Therefore, and despite the recent formative intentions of the new national assessment policy, the assessment is, in the country, framed as high-stakes by its summative features, and this study seeks to provide new evidence on this issue.

In addition, practices inside the Chilean classroom are regarded sometimes as not considering students' academic problems or weaknesses (García & Paredes, 2010). There is literature that describes that some teachers unreflectively register grades without any further

reflection of their meaning and only rarely do they take responsibility for the learning problems or low achievement of their students (García & Paredes, 2010). Commonly, Chilean teachers tend to attribute the cause of students' low results to external factors in the school, such as poor family support or lack of student effort, ignoring the ethical implications of that attitude (García & Paredes, 2010; Mizala & Romaguera, 2000; Sánchez & Jara, 2019). For instance, Sánchez and Jara (2019) stated when conducting 60 interviews with pre-service teachers, that teachers do not accept that children commit mistakes and attribute to them exclusive responsibility if learning is not attained. In this regard, it is clear that the school cannot change the family's socioeconomic background, yet other literature on the topic states that the school can mitigate against its impact on students through effective and ethical educational practice (Brunner & Elacqua, 2003).

The revised national assessment policy, created by the Ministry of Education (Mineduc, 2018a), attempts to expand and look beyond the Chilean classic reductionist vision of assessment as simply 'summative' in the form of grading and numbers. This has been done with the promotion of new guidelines which place a strong focus on not restricting assessment practices and promoting the use of formative assessment (Mineduc 2018a, Article 18, f and g). Formative assessment is defined by Mineduc as being used to support and monitor students' respective learning processes. That means focusing on making decisions about assessment practices that are necessary for the students' progress (Mineduc, 2018a). Despite Mineduc's interest in incorporating variation into assessment strategies within the schools, formative assessment is regarded by Shepard (2019) as not being possible to separate from the summative process where tests and accountability-focused instruments are commonly used. The difficulties between the formative and summative assessment interaction (Black & William, 2009; Earle, 2014; Newton, 2007; Remesal, 2011; Yates & Johnston, 2018) will be further explored in Chapter 2. Evidence of summative assessment or summative use of formative rather than formative itself remains more likely to be found in the context of Chilean SBA and this study seeks to explore this issue further by exploring the complexities of the use of formative and summative assessment in SBA. Thus, this study contributes to the comprehension of the main features and complexities of SBA for other countries using the same system.

1.2.3 Personal rationale

I was born and raised in the district where this research was conducted, in the southern area of Santiago, the Chilean capital city. I grew up in a compact area of the city, cheek by jowl with a dangerous barrio where crime, drugs and violence were common, even during the dictatorial political regime under which the country lived at that time. I was raised in a low-income family with three children and my father worked as a bus driver, so I studied with the assistance of scholarships throughout my entire life. As a teacher at the secondary level I have worked in 13 schools during my career, and I have witnessed first-hand the reality of my students within a school where, as was stated before, some of them fail, repeat and see their future possibilities diminished, reduced or often entirely extinguished. For me, it is clear that without my grades and the scholarships I obtained, I would not be here today. That is why the predominance of grades and the assessment strategies that prompted them are clearly relevant to analyse, and thus far have been understudied. School assessment, grades and achievement tell a history of differences among students and a gap between the curriculum's objectives and students' level of performance, as evidenced by assessment (Heritage, 2008; Young and Kim, 2010; McMillan, 2013; Scott, 2018). Consequently, there is a clear gap between successful students and those who are failing who deserve attention. Differences in achievement beg the question of what is being done inside classrooms in terms of teaching and what the assessment procedures results are telling us about achievement and learning. So, why assessment? Because it is linked with students' learning and future because it is the key for a better life. Additionally, the new assessment decree issued by the Ministry of Education challenged, once again, the school communities by promoting the use of formative assessment. However, there is a lack of studies describing how teachers assess their students inside the classroom in the country. For instance, the only study that Mineduc conducted before the publication of the assessment decree N° 67 during 2018, did not include any subsidised school within the medium or medium-low socioeconomic groups of the country (Mineduc, 2016). Both facts, the lack of research on SBA and also the lack of focus on the most usual schools that the country has, subsidised, and the most common people that the country have medium and medium-low class, made me think about a study on SBA in this segment. I was deeply moved by the question: if we do not have updated information about how teachers are implementing assessment in reality, how are we supposed to implement a new policy? Where does this policy come from? In my personal experience as a teacher, I have witnessed a strong summative culture

where grades are a main component of the whole assessment process. In that sense, I was really interested in exploring how teachers were reacting to this new decree that followed the line of Mineduc in relation to formative assessment promotion. My interest on the topic is also based on the relevance of the assessment process for secondary school students, in terms of the impact of grades on promotion, certification and university entrance, keeping always in mind the neoliberal educational policies implemented in the country in the last decades.

1.3 Research aim and research questions

1.3.1 Research Aim

The research seeks to better understand and describe SBA practices in one Chilean secondary school through teachers' and students' voices, with an emphasis placed on assessment purposes and methods used, grading and feedback reflecting students' results and, finally, the fairness of the overall assessment system.

1.3.2 Research objectives

The research objectives are as follows:

- a. To conduct an international literature review of SBA practices.
- b. To review the Chilean research context in terms of the current policy and practice of SBA.
- c. To conduct an empirical study in one secondary school by analysing and describing:
 - i) how and why teachers conduct specific SBA practices and teachers' and students' perspectives based on them;
 - ii) students' and teachers' perspectives about the grades and feedback resulting from the SBA and
 - iii) students' and teachers' perspectives on assessment fairness.
- d. To describe the implications of the findings for improving SBA practices for teachers' and students' within the school context in Chile and similar countries contexts.

1.3.3 Research questions

In order to accomplish the objectives proposed for this study, the specific questions that will guide it are:

RQ1. What are the SBA purposes and methods used in one secondary school in Chile and what are teachers' and students' perspectives about them?

1a. What are the assessment purposes used by teachers, as indicated by the school's documents and teachers' and students' perspectives?

1b. What are the assessment methods used by teachers, as indicated by school documents and teachers' and students' perspectives?

RQ2. What are teachers' and students' perspectives about the information provided by the assessment, in terms of grades and feedback, in one Chilean secondary school?

RQ3. What are teachers' and students' perspectives about fairness of the assessment system in one Chilean secondary school?

1.4 Overview of conceptual framework

The conceptual framework of this study is based on an inductive process where different ideas based on experiences on SBA throughout the world are integrated. This integration provides a broader understanding of SBA and shows relevant relationships derived from those concepts. Thus, the concept of School Based Assessment sets the 'research parameters' (Ravitch & Riggan 2016). In this way, the conceptual framework allows us to connect all the aspects of the empirical study (Shields & Tajalli, 2006). Consequently, three of SBA's key features were researched: assessment purposes and methods (RQ1), information provided by the assessment, in terms of summative grading and formative feedback (RQ2), and ethics of the assessment process (RQ3). As Liehr and Smith (1999) stated, "each of these terms refers to a structure" (p. 12) that guided the researcher during the process. Additionally, this study has considered teachers' and students' voices as a key concept and a common frame from where to look and to understand the SBA reality inside the school. Teachers' and students' perspectives on these three themes have been guided by studies such as Brown's inventories of teachers' and students' beliefs on assessment (2011, 2016). The study will be also framed by international literature in SBA studies and by experiences in a variety of locations, such as Hong Kong, Ireland, Norway, Scotland, and Finland

(Cheng et al., 2011; Cheng, 2019; Cumming & Maxwell, 2004; Mäkipää & Ouakrim-Sovio, 2019; Murchan, 2018; Tong & Adamson, 2015). This allowed us to understand that SBA involves a combination of summative and formative assessment. In this way, it is focused on the assessment used with students occurring inside the school, decided by teachers, excluding standardised or external examinations, that in any case will have an impact inside the classroom, as will be detailed in Chapters 2 and 3. Moreover, in contexts where both external and internal assessments exist, this results in potential contradictions between what the schools want to assess and what is being asked because of the external assessments pressure (Martínez, 2019).

The lack of Chilean studies on school assessment methods described in previous sections also implies further research is needed and also prompted a revision of Latin American literature on SBA to characterise the features of the zone. There is evidence from countries such as Peru and Mexico that formative assessment, which is an important part of SBA, has not been correctly promoted in order to have a real application: it should be but is not (Falcón-Troncoso et al., 2021; Monzón, 2015). This implied a study design where the exploration of SBA practices was planned without assuming that the practices suggested by educational policies can be taken for granted as occurring in the educational space. Therefore, this study considered the “socio-culturally responsive assessment” concept. In practical terms, this study was impacted by the acknowledgment of cultural and language differences found between the Western and Southern cultures, the concept will be further explored in Chapter 2 (section 2.3). It is important to make clear how some of these concepts embedded within my subjectivity as a researcher, shaped the research and the collection and interpretation of data (Merriam & Tisdell, 2016). In this way, this study acknowledges global differences in understanding assessment/evaluation concepts, understanding them as a ‘script’ impacted by social and cultural mediation (Elwood & Murphy, 2015). In this regard, the conceptual complexity around the SBA concept allowed the identification of key components and topics that were acknowledged by the research questions (assessment purposes, methods, results, and fairness), and how these are inextricably linked to another (Hennink et al., 2011). The assessment concept is shown exhibiting several differences in fundamental purpose and international agendas which promote certain practices that interact in complex and differential ways (Kennedy et al., 2008). This is reinforced by a body of literature on the topic in Spanish, where the assessment topic is studied sometimes without reference to the formative or summative concepts but rather as an integral process that does not always clearly label or distinguish between them (Santos Guerra, 1988, 1993, 2003). Here, the influence of local

authors such as Santos Guerra (1988, 1993, 2003) and Condemarín and Medina (2000) is highlighted. For instance, a recent study by Muñoz and Solís (2021) shows the use of qualitative evaluation (formative) and quantitative evaluation (summative), describing how both can be part of formative. The study states the need to understand evaluation with a mixed view, integrating qualitative and quantitative aspects of the process (Muñoz & Solís, 2021). It must be noted here that 'School Based Assessment' is a concept that has no direct translation used or exists in Spanish. Thus, it is usual to find concepts such as 'mixed' evaluation or 'integrated' evaluation, as in Muñoz and Solís (2021). The Chilean researcher Gysling (2017) explained in her doctoral thesis (Leiden University) that in English there are three words to refer to different processes of assessment: assessment, appraisal and evaluation, which are not distinguished in Spanish. In fact, in Spanish, we only ever tend to use 'evaluación' (evaluation), and assessment, per se, does not exist as a term. Therefore, the conceptual plurality involving methodological and theoretical ideas of the assessment/evaluation concept needs to be taken into account, without assuming that they can be generally used interchangeably in the same way (Escudero, 2003). Therefore, throughout this study, assessment will be understood as a polysemous concept that can have a different conceptual basis (Maldonado-Fuentes, 2021). That means that the treatment of the SBA process will assume that practice does not always follow the same patterns worldwide despite policy intentions and that Western ideas reach some zone at slow speed. This reinforces the idea of 'culturally responsive modes of assessment', as Brown et al (2022) would have it. In this sense, this study will acknowledge the different ways in which SBA is conducted in school contexts, respecting their differences as they respond to different school realities. Consequently, this study maintains a critical and careful standpoint when using concepts such as formative or summative assessment, given some consider the 'formative' concept as a Western construct (Kennedy et al., 2008). The assessment concepts in this study (assessment/evaluation) are used to reflect on their real appearance within the school context, always considering that the main feature of SBA is the use of summative and formative assessment. Therefore, the concepts are used in their suitability to answer the RQs and also considering the culture where they are used, which is impacted by history, policy, and social features.

1.5 Overview of Methodology

Assessment and learning occur in localised, situated, dynamics and in partially unique contexts (Baird et al., 2014). Thus, the examination of the Chilean educational context will embrace these complexities by adopting a philosophy where reality (ontology) is socially constructed from people's experiences (Prasad, 2017); in this case, from teachers' and students' perspectives within the school context. At an epistemological level, this study understands that knowledge is co-constructed by people, following their cultural characteristics (Creswell, 2013). Once again, teachers' and students' voices will be the key to construct knowledge on SBA. Considering that schools represent unique contexts in open systems, an interpretive philosophical approach, and a qualitative methodology design will be applied to the exploratory study of one school. The qualitative research paradigm allows the examination of people's experiences in detail by using different methods, that provide an in-depth understanding of the perspectives of people (Hennink et al., 2011). The methods proposed in this research are document analysis together with semi-structured interviews. Documents comprising assessment purposes and methods were selected by their suitability to give information about how SBA is implemented within the school. Interviews were conducted with 10 teachers and 10 students at the secondary school level in one subsidised school in Santiago, focusing on their perspectives on SBA. These methods will allow us to understand the meanings and interpretations that teachers and students give to the school assessment phenomena. Therefore, the methodology assumes an interpretive approach, characterising a research process flexible and based on people's natural settings (Denzin & Lincoln, 2008; Hennink et al., 2011). The interpretive paradigm involves that reality is socially constructed as people's experiences occur within social, cultural, historical, and personal contexts (Hennink et al., 2011).

The study's philosophical approach and design have been considered, as suggested by Merriam and Tisdale (2016), as a comfortable match with the researcher's worldview, personality, and skills. Since approaches are based on a set of beliefs, namely, human constructions, they have been understood here as being unified and never in terms of being absolute truths (Denzin & Lincoln, 2008). The approach of 'critical interpretivism' uses features that are common to, critical and interpretivist approaches. Both highlight the importance of local and specific constructed realities that are shaped by social, political, and cultural values, subjectivist in its focus and dialectical in its methods and process (Denzin & Lincoln, 2008). Therefore, the findings are

presented considering the possibilities of different approaches generating “intellectual, theoretical and practical space for dialogue, consensus, and confluence to occur” (Denzin & Lincoln, 2008, p.167). In this sense, the presentation of findings agrees with the belief that criteria for judging reality or validity are not absolute (Bradley & Schaefer, 1998, cited by Denzin & Lincoln, 2008). Critical perspectives are associated with research that recognises the power dynamics that impact how some social conditions are taken for granted (Cannella & Lincoln, 2016). For this research, this perspective included the unequal educational opportunity in Chile, especially after secondary school, and high stakes that creates forms of oppression for young people that result in inequality prompted by economic status (Kremerman, 2007; Torché, 2005).

Following these ideas, an initial design that included a mixed methodology in order to study students’ grades in a quantitative sense was discarded, as numbers alone did not allow to understand them as human constructions and their real impact on students’ lives. Additionally, the study design did not include observation of classes as they were moved to online provision during 2020 and 2021 due to the COVID-19 pandemic.

1.6 Overview of the thesis

The study is presented across ten chapters. The present chapter, the Introduction, presents the research problem and background and details the general, local, and personal rationale for conducting this study. It also makes explicit the research aim by detailing the research objectives and research questions. This is contextualised by presenting an overview of the conceptual framework and research design. Chapter 2 includes a critical account of the international literature on SBA. It contains an analytical account of how the assessment topic is treated in several different country contexts and what elements are highlighted as more relevant. These elements were considered to design the research questions and methodology and especially look to close the gap in the national literature on the topic. Additionally, this chapter is presented highlighting the relevance of teachers’ and students’ perspectives on the topic. Chapter 3 presents a contextualised account of the Chilean educational reality and SBA in the country as reflected in the academic literature, government policy documents, and local literature on educational evaluation in Spanish. This chapter supports the reader’s comprehension of specific features of the country that are relevant to understanding the assessment context and the research findings. This chapter also reflects on the multiple difficulties involved in the assessment topic study, which is

particularly challenging depending on the cultural context where the assessment occurs. Chapter 4 details the methodology and research design of the study and its justification. It presents the results of a preliminary study conducted in Chile between November 2019 and January 2020, that guided the improvement process of the interview questions and also allowed the study to determine the documents needed to answer the RQs. The chapter also includes a detailed account of the data collection and analysis made and the themes and subthemes used to answer RQ1, RQ2, and RQ3. Themes and subthemes were organised in their suitability to answer each research question on assessment purposes and methods, information, and fairness. That allowed to make a clear transition to findings, especially for Chapter 5, which is the longest as two subordinated questions and two sources of data are involved (documents and interview transcripts). Therefore, Chapter 5 includes findings for RQ1 on assessment purposes and methods, Chapter 6 for RQ2 on information provided by the assessment, in terms of grades and feedback, and, Chapter 7 for RQ3 on assessment fairness. Chapter 8 presents the discussion of findings presented in Chapters 5, 6, and 7. The findings are discussed in contraposition with the conceptual framework used for the study and the literature review on the topic (Chapter 2), as well as the research context (Chapter 3). The final chapter, 9, presents the conclusion of the study. It also includes the study limitations and future recommendations for research on the assessment topic, especially for the Chilean educational context and policy.

Chapter 2: International literature review on SBA

“Educational systems are complex and contradictory. Teachers’ daily work is regulated and influenced by decisions at the policy/macro level. They are expected to fulfil their professional duties according to laws, policy guidelines, and other directives, while their work is inevitably also charged with an extensive responsibility for other humans”.
(Hirsh, 2020).

2.1 Introduction

The study considers two interrelated concepts of assessment that refer to the same process: SBA and educational evaluation. The latter, as there is no equivalent concept to ‘school based assessment’ in Spanish, therefore the ‘educational evaluation’ concept (*evaluación educacional* in Spanish) was adopted for research in this language. Both concepts involve the same basic features at the school level when related to students’ assessment, both internationally and also in the Latin American context. Therefore, SBA as internationally understood is what is described by the Chilean policy to be implemented in the schools. This is an assessment system that combines and integrates the use of ‘formative’ and ‘summative’, as understood internationally, even when considering assessment through the Latin American and Chilean literature, where usually qualitative (for formative) and quantitative (for summative) concepts are used. The literature in Spanish tends to use simply ‘educational evaluation’ and ‘school evaluation’ (or in Spanish *evaluación educacional* and *evaluación escolar*) as an integral process. However, there is also literature in Spanish that study specifically summative/quantitative and/or formative/qualitative assessment, which is not the focus of this study as SBA integrates the use of both, and the literature review was intended to focus on SBA.

It must be recognised that in English, the use of the term ‘evaluation’ seems to be generally linked to external evaluation, with a focus on elements such as curriculum/contents, school quality/accountability and students’ performance (Brown, 2011; Newton, 2007). Therefore, the literature review will be organised in two different chapters. Chapter 2 will cover international literature in English for the SBA topic and Chapter 3 will address the research and policy context by covering Latin American and Chilean literature on evaluation, adding some specific features of the Chilean school system that helps to contextualise the study. These chapters will allow the reader to identify some differences on the assessment meaning, depending on the context.

2.2 Approach to Conceptual Framework

A conceptual or theoretical framework details the specific perspective under which the researcher explores, interprets, and/or explains events (Imenda, 2014). It is defined by Hughes et al (2019) as “a system of concepts, assumptions, expectations, beliefs and theories that support and inform one’s research” (p. 28). It allows us to understand a set of key elements and characteristics of the main phenomenon being studied (Maxwell 2013; Creswell 2007). Even when theoretical frameworks address a theory, they involve a set of concepts and explanations drawn from the same theory (Imenda, 2014). However, this is not always the case for the assessment field, where a number of theories have been put forward to explain SBA. These theories include psychometrics on learning measurement (William, 2006); socio-constructivism and assessment considering the central role of the student (Mäkipää & Ouakrim-Soivio, 2019), and sociocultural theories, where the culture and specific context are key when shape the assessment meaning and practice (Elwood & Murphy, 2015; Kennedy et al., 2008; Young & Kim, 2010). Additionally, a considerable amount of literature does not address explicit theories on assessment, or what constitutes real and ethical teaching and learning and their assessment. Literature on SBA does not always follow an explicit theoretical perspective, and different concepts and ideas are used to explain the assessment process.

Therefore, in this study SBA did not encompass a single theory; instead, this approach will synthesize the existing experiences in the literature, from where a conceptual framework will be created. It integrates a perspective that focuses on one single problem (Imenda, 2014). This process allowed the identification of concepts and principles that support the research problem. Some of these concepts oppose each other, as no one theory can adequately direct the research process to sufficiently answer the research questions (Imenda, 2014).

Following Hughes et al (2019), “it is important for the conceptual framework to include the nature and source of the data and explicit the meanings attached to it” (p. 28). Therefore, this study follows an educational perspective that considers the following key concepts in order to consider the context where SBA takes place and its characteristics:

- i) socio-culturally responsive assessment and cultural scripts in order to contextualise and to understand the structural specificities of SBA. Both concepts have been taken here as concepts and not as theories, since SBA involves the use of formative and summative assessments that are usually linked to the concept, grounding the

approach to a consideration of the cultural context where SBA takes place. However, even when the cultural context is linked to the sociocultural theory of assessment (Elwood & Murphy, 2015; Kennedy et al., 2008; Young & Kim, 2010), the study will draw from several theoretical frameworks to address the research objectives. This because assessment is also linked with other ideas and theories such as learning measurement (Wiliam, 2006). This is especially true for summative assessment, where testing sometimes completely ignores the socio-cultural reality where learning is occurring and focuses only on results. This exemplifies the idea that a research problem cannot meaningfully be researched if it draws from only one theory (Imenda, 2014), which is the case with assessment. Additionally, this is the reason why this study has synthesized existing views on SBA in the literature in the following concepts:

- ii) summative and formative assessment, assessment purposes and methods, grades and formative feedback, fairness, and teachers' and students' voices, in order to answer the research questions of the study in the specific features of SBA in one school in Chile.

Therefore, the conceptual framework involves a consideration of the ideas of the theoretical perspective(s) from which one approaches the construct (Hughes et al., 2019). In this line, the conceptual framework has allowed to lead the literature review process, to select research methods appropriated to the studied phenomenon and also to interpret the results (Imenda, 2014). The literature review has been inductively conducted, this is by reading several individual theoretical and empirical perspectives and reports on SBA to draw general conclusions on the phenomenon. Additionally, the conceptual framework considers international similarities based on the literature on the topic, but also addresses differences when using SBA, dialoguing with both experiences to finally converge into some ideas such as the use of an interpretive approach where teachers' and students' voices are essential to understand the research topic. These ideas will be characterised in the upcoming sections.

2.3 Socio-culturally responsive assessment as a response to cultural scripts.

This section will recognise the relevance that culture has for assessment implementation at any level and country. It will introduce two concepts that will have an impact on how SBA is

understood in this study: cultural scripts and culturally responsive assessment. As has been previously stated, assessment inside schools should both reveal and respond to its specific cultural and social characteristics. Internationally, when students' assessment is analysed, the procedures used vary significantly between countries and even within them. However, it is characterised as a process directed towards different assessment purposes that seems to be impacted by cultural pressures which favour certain kinds of practices, such as testing cultures in some countries (Kennedy et al., 2008).

More than 20 years ago, Broadfoot (1999) recognised assessment as a social product that is impacted by the values and traditions of particular cultures where the interests of different actors are combined differently in order to define quality or merit. In this vein, Elwood and Murphy (2015) also emphasised that assessment practice needs to be understood within its social, historical, and cultural context: "Recognise the essential relationship between learning and assessment, but that is affected by the social, cultural and historical lives of students and teachers and the political and economic contexts within assessment happens" (p.187). Cultural beliefs are called 'scripts' by Elwood and Murphy (2015) and refer to what assessment is and what is used for which also contain ontological and epistemological beliefs. Citing Bruner (1996) the authors explain that these beliefs are conserved, elaborated, and passed on in the form of assessment purposes, uses, and practices. Consequently, as the assessment procedures inside the school are immersed into a culture, it can be expected that they are influenced by different elements from that culture, meaning they acquire new and different meanings. Citing Frierson et al (2010) culture is understood as "a cumulative body of learnt and shared behaviour, values, customs, and beliefs common to a particular group or society" (p.75). Consequently, assessment needs to consider alternatives and to understand the ways the process interacts differently within different contexts.

A cultural script, as it might be understood in SBA, refers to how the school (as a collective) does assessments within broader systems. These ways of assessing practice involve how the school interprets what assessment is (i.e., behaviours, customs, methods) and what it is used for (i.e., purposes, reasons) aligned with the beliefs and values supporting these decisions. Culturally responsive assessment, according to the OECD reflections on the concept, is a type of assessment that is respectful and cognisant of the diversity of the culture and experiences where assessment processes are conducted (Burns et al., 2019). It has already been stated that countries throughout the world present differing characteristics and boundaries that also impact the assessment style that they adopt, depending on the cultural context where the assessment takes place (Kennedy et

al., 2008). These values about what is important will define the complete assessment process at different levels, which range from worldwide trends on assessment to national policies, to the micro-level aspects taking place inside each school. The cultural script concept will be used in this study to understand that policy on assessment is implemented in different countries with different cultural contexts, diversity of people, and specificities of the local culture. This diversity is understood as containing a range or multiplicity of cultural affiliations and a variety of languages (Burns et al., 2019). This is relevant because the decisions taken regarding assessment purposes and methods used to collect data from students' learning "are bound by cultural values" (Burns et al., 2019, p. 7). Hence, within the school we can find cultural scripts about assessment, which can allow SBA to be conducted in a cognisant and responsive way.

An example of 'culturally responsive modes of assessment' has been deployed by Brown et al in a study in Ireland (2022). Following the study, this assessment can be understood simply as being based on what constitutes good classroom assessment for all students. They studied classroom assessments that can be summative or formative as responding to students from diverse cultural backgrounds to be fair. The study revealed teachers' wide awareness of their students' needs but a narrow base of teachers' culturally responsive assessment knowledge. Within these teachers' beliefs, it was noteworthy that they were aware of providing assessment activities that allowed students to show their knowledge and skills in different ways. Even when teachers tended towards using a test and grade after a unit, this was challenged by inviting them to use alternative methods, such as portfolios, case studies, or crosswords (Brown et al., 2022). Teachers also designed their own assessments and used online material for them to be more suitable for their students. This study highlights the relevance of assessment methods that truly respond to students' needs while also providing relevant information about their learning. For this study, a focus of interest is how teachers decide on assessment methods and which purposes are linked to those (RQ1), and how they respond to their students' needs.

The "responsive assessment" concept presented draws from socio-cultural aspects suggested by Elwood and Murphy (2015) and puts forward the idea that assessment is a social practice incorporated within broader systems that are inevitably impacted by socioeconomic principles. Therefore, the concept of "socio-culturally responsive assessment" will be used in this study from now on, as it provides a more complete account of the Chilean inequalities introduced in Chapter 1 and will be further explored in Chapter 3.

Recently, Bennet (2023) highlighted the concept of “socio-culturally responsive assessment”. The author stresses the relevance of socioeconomic positioning impacting the products of the assessment, such as tests, as well as the wider cultural setting. This positioning reflects the structural inequalities affecting our society. However, Bennett (2023) used the concept towards a theory of sociocultural responsive assessment for standardised testing. Standardised testing is not the focus of this study, nor the spirit of the Chilean policy, that promotes the use of a variety of assessment methods and formative assessment (SBA). Therefore, this study will focus on the integral socio-cultural aspect of the assessment practice as proposed previously by Elwood and Murphy (2015).

Kennedy et al (2008) also considered a raft of cultural issues, especially for the summative and formative concepts within a range of different contexts. The authors explain that assessment can be directed toward different purposes and practices that are dictated by cultural pressures (Kennedy et al., 2008). While some countries support the worldwide spread idea of formative assessment, it is clear that the importance of summative cannot be obscured or put to one side by thinking that formative solves all the problems in the assessment field (Kennedy et al., 2008). Thinking of the negative consequences that summative assessment can bring for students, it is especially relevant to consider the culture where these negative consequences could occur. The study highlights the relevance of assessment for Asian and Eastern contexts where public examinations tend to play a key role for students and societies. Considering the relevance of summative assessment within these cultures, it is understandable that formative assessments are considered to be a Western construct that can be taken up or not, depending on the society’s characteristics and demands. Hence, the concept must be adapted to local needs and specific local priorities.

2.4 International experiences on SBA exhibiting the relevance of cultural scripts.

SBA is widely different in subjects and ways in places that serve as examples, such as Australia, Hong Kong, Indonesia, Ireland, Norway, Scotland, and Finland (Cheng, 2019; Cumming & Maxwell, 2004; Mäkipää & Ouakrim-Soivio, 2019; Murchan, 2018). A key orientation of SBA is to bring some balance to the students’ assessment beyond an examination system (Murchan, 2018). A main SBA objective, then, is that teachers are the main students’ assessors within high-stakes systems where, as it has been stated, promotion, certification and university entrance are often

involved. Thus, teachers are still asked to prepare their students for other external assessments, which can have an impact on their autonomy as creative professionals (Murchan, 2018). External, high-stakes testing has been labelled as cannibalising and fragmenting the existing curriculum, eliminating activities and arbitrarily leaving other areas of the curriculum unattended, privileging contents and knowledge over processes and critical judgement (Kohn, 2001; Sondel, 2016). Contrarily, SBA intends to promote higher-order thinking skills over shallow learning or test/driven cultures, which has been the case in several European and Asian countries.

As stated by the ARG in England (2006), summative assessment (involved in high-stakes systems) is confronted and comes into tension with the ethos behind formative assessment. This is because the information that results from the summative assessment can be used to classify students, creating a class hierarchy and, in doing so, labelling some of them in a negative manner. In this sense, the assessment can also reflect the students' ability to adapt to the school discipline without necessarily demonstrating knowledge and skills (Kožuh, 2019). These difficulties can be addressed with formative assessment according to the literature, especially for the presence of feedback supporting continuously the students' progress. SBA richness can support and enhance students' learning experience (Murchan, 2018).

The use of SBA tends to appear when a traditionally summative assessment system starts to promote the use of formative assessment. However, this inclusion, usually promoted by educational policies, requires profound changes in the perceived role of the school, as it should become a place of support for students' development and promote constructive criticism (Kožuh, 2019). These changes are not made without resistance, as SBA is a widespread practice that is not always preferred by all practitioners, such as in the case of Ireland where teachers seem to prefer external and centralised assessment instead of SBA (Murchan, 2018). This is because teachers' responsibility and freedom to create assessment instances within SBA requires different elements to ensure objectivity in the process. This has been a main concern in Ireland and Hong Kong as well (Morris & Adamson, 2010; Murchan, 2018). In a study on SBA in South Africa, Reyneke (2016) highlighted several practical issues about SBA implementation in the literature, such as the access to appropriate assessment resources, a lack of practical support within the school, a lack of time to implement and administer feedback to students, issues with large classes, limited resources, and teachers' heavy workload.

Kennedy et al reinforce the importance of asking if students will learn from the tasks and methods that are used in specific socio-cultural contexts rather than simply taking Western

assessment contexts to be the norm (Kennedy et al., 2008). This is particularly true when considering that assessment-constructed artefacts follow practices and processes which are socially and culturally impacted (Elwood & Murphy, 2015). Schools and their practices are associated with broader systems that shape educational policies and structures through government policies and directives (Elwood & Murphy, 2015). In that way, they impact in what it is understood by educational purposes and achievement and success: learners and assessors produce, reproduce, and transform the collective life of society through assessment (Roth & Radford, 2011). For this reason, culture influences people's beliefs about how they know and what is legitimated and valued as worth knowing in a given society (Elwood & Murphy, 2015). Then, it is understandable that forms of knowing and showing knowledge will differ greatly worldwide. These differences cannot be ignored easily as assessment procedures can have specific features that must be understood within the culture where they are created and acquire meaning (Elwood & Murphy, 2015). Dominant cultural scripts must seek to understand these complex interactions of assessment within the school, looking for teacher-student relationships (Elwood & Murphy, 2015). For instance, Puad and Asthon (2021) reported the difficulties of non-Western countries when implementing formative assessment policies, as summative is a dominant script. Places such as Indonesia, Hong Kong, China, Taiwan, and Saudi Arabia especially faced challenges, as they are cultures where learning tends to be very teacher centred, with learners, by extension, labelled as passive (Puad & Ashton, 2021). These elements, proper and unique to each country or culture cannot be ignored as they impact the characteristics of the assessment, not only in practice but in their meaning as well. Moreover, their meaning, beyond policy guidelines, can be found by considering how assessment is studied and what its features are across different countries and cultures (a table summarising specific features of SBA by countries can be found in Appendix A).

2.5 School based assessment: characteristics

As was stated in Chapter 1, SBA combines summative and formative assessment strategies that provide important information about students' present and future learning and it is conducted regularly by teachers inside the classrooms (Brown & Hirschfeld, 2008; HKEAA, 2013). SBA was first introduced in Hong Kong in 2000 and defined as an assessment carried out by schools with students being assessed by their own subject teachers (HKEAA, 2013). It is featured as helping students to understand their strengths and weaknesses through quality formative

feedback received from teachers. Following the HKEAA definitions (Hong Kong Examinations and Assessment Authority), SBA reduces dependence on the results of public examinations while also boosting students' confidence and motivation to learn, enhancing autonomous learning (HKEAA, 2021).

In time, SBA has become essential in numerous countries because it provides important information about students' learning in the form of grades, which are prompted by teachers' deciding on a variety of tasks in the natural setting of the classroom that can be combined (or not) with external examinations (Cheng, 2019; HKEAA, 2021; Murchan, 2018). However, it is claimed to also include more detailed feedback, in a formative sense, which encourages students to work in areas where improvement is needed (Cheng, 2019). From this perspective, the SBA process highlights the inclusion of purposeful assessment tasks which are decided on by teachers in daily teaching settings that aim to monitor and improve students' learning (Cheng, 2019). That is why, SBA is regarded as serving both, summative and formative purposes (Cheng et al., 2011; Cheng, 2019). The literature suggests that formative assessment occurs when the learning process is happening (Black & Wiliam, 2009). On the other hand, summative assessment is used after a certain period, for instance at the end of a unit, a course, or a term (Mäkipää & Ouakrim-Soivio, 2019). Therefore, SBA combines these elements allowing teachers to create assessment instances to grade students but also to improve their learning by making students reflect on their improvement needs and providing them with clear feedback and expectations (Cheng, 2019). In this vein, Stiggins (2002) stated that the crucial distinction is that summative helps to determine the status of learning, whereas formative helps to determine the status for learning.

Historically, summative and formative concepts go right back to Scriven's work (1967). Scriven is often regarded as creating the concepts concerning curriculum/program evaluation, which was extended later to students (Kennedy et al., 2008; Mäkipää & Ouakrim-Soivio, 2019; Newton, 2007). Subsequently, Bloom, Hastings and Madaus (1971) used this distinction for formative assessment occurring while learning is happening and summative assessment as being used after the learning has taken place (Newton, 2007). Beyond this, Black and Wiliam (2003) drew the distinction, then, followed the function that the assessment fundamentally has.

In more recent times, assessment purposes are often described as either formative or summative (school-based), which in turn also include external summative purposes (Irving et al., 2011). Following this, it is highly probable that in SBA systems, teachers will refer to at least these three purposes, as formative and summative as being usually impacted by external assessments,

especially during secondary school. The international literature reports complex teachers' conceptions of assessment, used for different purposes, highlighting that within educational spaces the balance between students, school, and society needs should be considered (Harris & Brown, 2009; Remesal, 2011). In this manner, Remesal (2011) highlights the importance of studying teachers' beliefs and conceptions of assessment as a first step before even trying to challenge their practices from summative to formative. The results reinforce the complexity of the assessment processes inside the school, whilst also showing the limitations of the formative/summative dichotomy (Remesal, 2011). At the same time, these complexities have an impact on how the literature on SBA, despite highlighting the good practices associated with feedback provision, is presented. This is because the literature tends to highlight the negative aspects of SBA implementation. These difficulties will be sometimes the tone of the literature, highlighting conflictive points of SBA implementation and use through the international literature on the topic. This literature review will also highlight several different, relevant, and complex elements of SBA.

2.5.1 Definition, strengths, and weaknesses of summative assessment

Summative assessment refers to final assessment processes (at the end of a learning unit or learning objective) and overall achievement in terms of concrete and valuable products used to assign grades, certify student achievement, and report (Briggs and Ellis, 2013; Meng et al., 2014; Shepard, 2019; Stiggins, 2002).

Especially for secondary students, the summative process also can affect students' decisions and post-graduation opportunities (Race et al., 2005). Following summative goals, the relationship between assessment and achievement is the focus on classroom day-to-day life as it has the potential for providing summative information about students' achievements, in the sense that it allows teachers to build up a picture of students' attainments and learning across different activities and goals (Harlen, 2004). However, adding SBA in a context where students have final examinations can increase fatigue among them given that they are additionally tested throughout the year by SBA, as has been reported in Singapore (Tan, 2017). Tan (2017) reported that SBA also generated extra work for teachers, producing an assessment system that is not used formatively but rather to record scores and to give an overall mark at the end of the course.

Summative assessment is used to compare learning outcomes against the objectives set by an educational system and the knowledge standards established by the syllabus or curriculum (Kožuh, 2019). This is particularly relevant in educational contexts where summative high-stakes assessment is still used to certify, promote, and access to university. In this vein, Cabral and Baldino (2019) refer to this situation as “two human beings meet(ing) in a real-world situation where one has the power of judgement about the future social situation of the other” (p. 278), describing the implications of SBA on high stakes. The high-stakes system can entail differences between countries, as in some of them large-scale standardised testing is enacted and grades are also used that result from SBA. In any case, the assessment carries significant consequences for the students, teachers, and schools (Herman et al., 2005).

2.5.2 Definition, strengths and weaknesses of formative assessment

Formative assessment is focused on students’ participation in the assessment process through different activities, such as their goal setting, their self-assessment or also peer-assessment and feedback, which are key for the monitoring of learning (Black & Wiliam, 1998; Cheng, 2019; Tong & Adamson, 2015). Within formative assessment strategies, the results are used continuously to improve teaching processes and by extension students’ learning outcomes (Black & Wiliam, 1998; Briggs & Ellis, 2013; De Luca et al., 2018; Shepard, 2019). In a formative sense, several authors posit that assessment activities allow teachers and students to work together to create, reflect, and evaluate students’ strengths and weaknesses in terms of the gap between the desired levels of achievement and current students’ performances and learning by providing formative feedback (Heritage, 2008; Johnson, Sondergeld & Walton, 2019; Meng et al., 2014; McMillan, 2013; Scott, 2018; Young & Kim, 2010).

The formative aspect of assessment includes activities like self-assessment and peer assessment, assessment influence on planning and sharing criteria, and clarification of learning goals with students (Black & Wiliam, 1998; James & Pedder, 2006). Learning goals and outcomes are also useful for checking students’ success with self-assessment, where they can independently check and self-monitor their processes (De Luca et al., 2018). These have been labelled as key elements of the formative purposes of assessment. However, they are also reported as being challenging to implement, especially within high-stakes systems, where examinations and grades are central (Tan, 2017). Results of the study conducted by Zhao et al (2017) suggest that

assessment results are scarcely used to adapt and improve instruction, which is a main objective within the formative assessment. Other aspects relevant in this regard, are that students might not understand the language used in the assessment process; for instance, when teachers share learning goals and success criteria (Sivenbring, 2019). Sivenbring study (2019), conducted with Swedish students shows a high level of difficulty for students to make sense of the assessment procedures language or jargon used to express it.

It is relevant to state that formative assessment is also linked with an ethos focused on “equitable and collaborative learning” (Shepard, 2019, p.193). However, some key elements of formative assessment can also be challenging, such as activities that involve communication of curricular knowledge or goals where the professional language used is regarded as remote and inaccessible by students (Sivenbring, 2019). Additionally, research reports that students’ reactions to formative strategies differ greatly according to their intentions and actions, their perceptions of assessment, and also classroom and school and country culture when implementing it (De Luca et al., 2018; Kennedy et al., 2008). For instance, Puad and Ashton study (2021) on classroom assessment in Indonesia shows that teachers tend to believe that peer assessment is not possible because students do not have the knowledge or objectivity needed to implement it, nor the honesty to do self-assessment. This reinforces the idea that formative assessment is not a simple process but rather a complex social and educational process (Kennedy et al., 2008). Therefore, one that needs exploring from teachers’ and students’ perspectives.

2.5.3 Problematizing formative and summative

The previous section has indicated that the assessment process is usually ‘dichotomised’ following two purposes or intentions: summative and formative. The tendency to isolate assessment in this way is problematic, as the concepts are separated and even shown to be incompatible (Taras, 2005; Kennedy et al., 2008). However, it cannot be avoided as the concepts are highly present within the SBA literature. The term ‘purpose’ itself can be interpreted in several ways and the distinction has turned incredibly problematic in the last thirty years (Newton, 2007). This is because the assessment field is already a complex process of collection, analysis and evaluation of evidence about the teaching and learning process (Remesal, 2011). Elwood and Murphy (2015) stated that the dichotomy was deficient and unproductive and, in any case, ought to be considered within social and cultural contexts. In this vein, Stiggins (2002) proffered the

crucial question: how can we be sure that our assessment instruments, procedures, and scores serve to help learners want to learn and feel able to learn? (p. 759). This raises the fundamental question of the meaning of the assessment in its relation to students' learning. Additionally, it invites us to keep a focus beyond distinctions that can be complex and that certainly obscure the assessment discussion, even when they are considered because they are part of common usage within the literature on the topic and also assessment policies in different countries.

Newton (2007) made efforts to clarify the purposes of the assessment in terms of the ones that serve for judgement, decision, or their impact level. While the judgement level helps grading achievement, the decision level allows to decide, act, or conduct a process, such as university entrance (Newton, 2007). The impact level here is related to the assessment design itself and the result produced on learners, such as their motivation (Newton, 2007). Brown (2011) simplifies the list of assessment purposes suggested by Newton (2007) avoiding the use of 'formative' and 'summative' and referring to purposes of improvement, school accountability, student accountability, and anti-purpose (associated to assessment with no purpose, irrelevant or detrimental to students' and teachers' work and life). Authors such as Harlen (2007) and Newton (2007) recognise that confusion about formative and summative has hindered the development of the assessment field by taken away the discussion from students' learning. Additionally, discussions about the purposes of assessment tend to diverge according to context. Therefore, the technical explanations about purposes found in one country cannot necessarily be transferred to another country's context, and could not be known in others at all.

The distinction between summative and formative is unclear in terms of the same tasks being used for both purposes: formative use of summative tests, for instance by talking about results and what can be improved; summative use of formative instruments such as rubrics, where a grade is obtained by the assignation of points to each aspect of the rubric. A rubric is understood as a coherent set of criteria that help students work with a description of levels of performance quality (Brookhart, 2013). As an assessment tool, rubrics set evaluative criteria, quality definitions, and a scoring strategy (Popham, 2000).

Therefore, differentiating between formative and summative in practice is far from simple (Mäkipää & Ouakrim-Soivio, 2019). Despite teachers' intentions to use school summative assessment formatively, this is to guide the learning process, tensions arise between that pedagogical aim and the school's requirements for reporting (Yates and Johnston, 2018). This is why the formative process is impacted to some extent by the need for grading and reporting. In

this vein, Boud (2000) stated that summative assessment becomes pointless if it inhibits the very learning that it seeks to certify. Thus, in this way, several studies reported the difficulties, limitations or reconciliation of the dichotomy summative/formative (Black & Wiliam, 2009; Earle, 2014; Newton, 2007; Remesal, 2011; Yates & Johnston, 2018) and the need of balance and complementarity between them as not being mutually exclusive (Flitcroft & Woods, 2018; Kožuh, 2019; Mäkipää & Ouakrim-Soivio, 2019).

Harlen (2007) states that rather than constituting a dichotomy, the concepts should be seen as different dimensions. In the same line, Black (1998) considered the concepts as being two ends of the same spectrum. However, the literature also recognises that if formative and summative are carried out concurrently, giving the student a score and also comments for improvement, the latter will most likely be ignored (Wiliam, 2011). In any case, how teachers understand both concepts carries fundamental importance, particularly in times of change concerning assessment policy (Brill & Twist, 2013), such as in the case of Chile in 2020. Additionally, there is acknowledgment of the relevance of teachers' beliefs, as they tend to be stable (even in error), impacting considerably their practice (Bonner et al., 2018). This could mean teachers' resistance to change, even when policies are asking them so.

Regarding this, Mäkipää and Ouakrim-Soivio (2019) stated that the literature shows that teachers' knowledge of formative assessment is deficient, a limited range of assessment practices, evidencing several studies on teachers' assessment literacy. Therefore, the assessment implemented by teachers is incongruent with the recommendations that can be found internationally in the literature (Mäkipää & Ouakrim-Soivio, 2019) and also in assessment policies. This should be especially relevant for SBA, which intends to combine its best elements to enrich the assessment process (Kožuh, 2019). The literature reports a number of different understandings and tensions between teachers related to the formative aspects of assessment (Johnson et al., 2019). The complexities of formative classroom assessment practices within cultures dominated by high-stakes psychometric testing have also been highlighted (Wiliam, 2006). These tensions also result from disagreements between teachers about what constitutes formative and summative uses of assessment in terms of whether a grade is awarded or not (Irving et al., 2011).

The study carried out by Mäkipää and Ouakrim-Soivio (2019) shows generally weak correlations and tensions between the use of formative and summative assessment purposes. Even when teachers indicate a preference for formative assessment, tensions arise between

formative purposes and the requirements of reporting results which are present within SBA (Irving et al., 2011). This also has a key impact on teachers' motives to use the feedback in order to pass students rather than focus on their learning improvement, which is related to their concerns about certification and qualifications (Pope et al., 2009).

James and Pedder (2006) showed, in their study on 558 teachers in England, that the assessment of learning (summative) reflects a rigid organization of learning through the national curriculum that usually involves testing. Again, the paradox of educational frameworks that prioritize measured performance (summative) over the quality of learning (formative) appears frequently within teacher's everyday practice. This paradox also impacts on students depending on teachers and rigid programs rather than being considered as independent and flexible learners, who are able to adapt their learning to new situations and challenges (James and Pedder, 2006). The need for the rehabilitation of summative not outweigh the positive aspects of formative has been long highlighted by authors such as Biggs in 1998.

Overall, the literature on the topic has assumed that there is a distinction to be made and that both terms are widely used but without any clear agreement on definitions. Although it was helpful to promote the formative assessment concept, the problem seems to be associated with the distinct meanings that educational professionals attach to it (Newton, 2007). The concepts are indeed highly present in the literature on SBA and also in national assessment policies, despite the complexity of their use. On this, McMillan (2003) stated the need to widen the purpose of assessment beyond the traditional view. Kennedy et al (2008) highlighted that "what does seem to be clear is that any bifurcation between formative and summative assessment is no longer useful" (p. 205), as the assessment focus should always be on students' learning. Therefore, it would be expected that despite their difficulties, the concepts can be used in very different ways, especially when particular school context realities and cultures in different places of the world are taken into consideration. In any case, following Newton's (2007) ideas, these complexities should always be openly acknowledged rather than concealed when studying assessment systems.

2.5.4 Assessment purposes: why assess students?

Often, purposes can be seen as being more complex than being merely formative and summative. For example, Harris and Brown (2009) reported in their study with teachers that a more multifaceted meaning of the purposes of assessment, such as compliance, external reporting

(including parents), motivating students, facilitating group instruction, and individualising learning came to the fore. The study showed that teachers have complex conceptions of assessment, using them for different purposes as they follow their distinct interests and attempt to balance students', society and the school's needs respectively (Harris & Brown, 2009). A strong tension between teachers' feelings about what is best for their students and what is needed for the school's accountability purposes was reported.

Reinforcing and expanding upon this comprehension of assessment purposes, Mäkipää and Ouakrim-Soivio (2019) also cited different assessment purposes in their literature reviews, such as to direct students' learning towards expected outcomes (Gronlund, 2003), to reveal students' learning (William, 2011), to motivate students and to help the future planning of teaching. Literature on the topic also focuses on assessment practices and values, teachers' professional learning, and school management practices and systems (James & Pedder, 2006). Past studies on teachers' beliefs on assessment, such as Delandshere and Jones (1999) identified the purposes and functions of the assessment as important factors. These include the distribution of students according to their achievement levels, teachers' perception of the curriculum and their sense of self-efficacy feeling, and finally, their beliefs about the teaching and learning process considering the students as learners.

Another insightful study on assessment purposes and methods, that supports the idea of its complex nature, is the one conducted by Zhao et al (2017). They analysed 266 articles on classroom assessment practices in Mathematics, where teachers reflected on the assessment purposes, assessment methods, and who they assess, putting a special emphasis on the feedback provided. In the purpose categories, they are summarised as allowing checks of the level of students' understanding, knowledge and/or skills, stimulating students to learn, informing instructional decision-making, establishing a harmonious classroom environment, and promoting students' confidence. As it has been stated, purposes are seen in terms of categories that go beyond and do not always include the overarching terms summative or formative. An element that will be relevant to consider for RQ1 about assessment purposes and methods.

Despite the form evaluation currently adopts (in the sense of being summative, formative, or both), assessment is the key to obtaining information about what students know, understand, and can do and, therefore, to know where they are and how they can reach their planned learning objectives (Briggs & Ellis, 2013; Towne & Shavelson, 2002). Beyond the purpose of the assessment, be it formative or summative, the urgent need for research on assessment inside the schools is

highlighted by authors such as MacMillan (2013) as it influences students' learning and effective instruction and teaching. In this vein, Harlen (2004) states that assessment conducted by teachers contains professional judgements that are used to draw inferences, make judgements of evidence, and also to gather evidence of students' learning. Therefore, it serves many different functions such as allowing the evaluation of the curriculum and gauging students' achievements (Kim and Young, 2010). In this way, the context where assessment occurs takes on a vital relevance, as information about students' achievements and learning activities must be situated within each school's circumstances (ARG, 2002). Therefore, school self-evaluation guidelines should provide key information about the curriculum, teaching, and assessment methods while also taking into account students' backgrounds and histories (ARG, 2002). The literature on SBA shows that it contains different and complex elements and challenges, although the relevance of the assessment process to obtain information about students' learning is not doubted. However, the distinct ways in which this information is obtained and used to respond to specific cultural and policy characteristics, will be further explored in the following section.

2.6 The relevance of teachers' and students' perspectives on SBA

Teachers' perspectives on SBA and school assessment procedures are a fundamental part of their labour, as they shape their classroom decisions and are tightly linked to practice (Pizarro & Gómez, 2019; Remesal, 2011). On the one hand, teachers' perspectives, and beliefs regarding assessment constitute a central aspect of the assessment practice inside schools (Othman, 2019; Pizarro & Gómez, 2019). The mind-sets within examination-driven cultures can shape them to test their students as they were tested as learners at school (Vogt & Tsagari, 2014). For instance, Remesal (2011) noticed that important studies of assessment, such as Stiggins (1999) or MacMillan (2001) failed to address the teachers' beliefs and conceptions that guide their practice. Remesal's (2011) studies on the topic show that assessment procedures have a relevant effect on teaching, learning, students' accreditation of learning, and teachers' accountability. On the other hand, students' position about the assessment procedures can help to improve our understanding to make them a valuable and potent tool (Sivenbring, 2019). Students' voices are defined in this study as "the views that students express regarding their learning experiences" (González et al., 2017, p. 995). Cook-Sather (2006) stated that students' voices position them to speak and act as critics and creators of educational practice, that is, through their voices they have the power to

influence the analysis, decisions and practices of the school. However, in the assessment field, De Luca et al. (2018) and Sivenbring (2019) state that only a small number of studies have examined students' general conceptions and perceptions of school assessment. Even less have considered the utility of the assessment process as a facilitator of learning. In that sense, Sivenbring (2019) highlights the thesis of Torkildsen (2016), which states that if assessment procedures are to be effective, students' and teachers' practices should be a central starting point to generate discussions. In turn, Tong and Adamson (2015) also stress the relevance of valuing students' voices on SBA in secondary schools, as they are under researched compared to other stakeholders involved in assessment procedures. As it can be seen, the literature is regarded as being more 'teacher centred', presenting a one-way flow of information or being one-sided with an emphasis on summative procedures and considering students as secondary and passive actors (Mäkipää & Ouakrim-Soivio, 2019; Remesal, 2011; Sivenbring, 2019). Students seem to be considered and viewed as recipients of other decisions rather than as key participants in the assessment process (Cheng et al., 2011). Assessment being teacher-led also brings consequences for the use of peer or self-assessment. This also goes against the social constructivist approach to learning, where students' active role is considered to be an important part of the assessment process (Mäkipää & Ouakrim-Soivio, 2019). These elements around students' perspectives are relevant also in their link to ethical aspects of the assessment process (RQ3).

Despite studies on perspectives on assessment have been done with both qualitative and quantitative methods, an important researcher on the topic is Brown (2006, 2019). Brown's work proposed a 'teachers' and students' conceptions of assessment inventory' or 'SCoA' where a questionnaire allowed to argue that teachers' practice on assessment is consistent with their conceptions (2006, 2011). The first trial of this inventory was conducted in New Zealand in 2003 by Brown and Hirschfeld (Brown, 2011). These inventories, then, are based on the identification of four purposes and practices associated with assessment (Brown, 2011). The items of the study are summarised in topics such as accountability, teaching and learning, anti-purposes of assessment (negative or detrimental aspects), and assessment usefulness (Brown, 2006, 2011; Yates & Johnston 2018). Some of these ideas were used to create the interview questions for this study (see Chapter 4, section 4.5.2).

Contrarily, Remesal (2011) could not confirm Brown's model on studies with Spanish teachers, demonstrating the complex nature of teachers' and students' conceptions of assessment and the relevance of the context, stressing as well the relevance of further research on the topic.

Her study shows tendencies that might be attributed to intrinsic tensions between pedagogic and societal functions of the assessment procedures in the school and difficulties in the implementation of assessment for learning practices. These elements have been considered in the design of this study.

Brown (2011) stated that the inventory was also used by Gao (2010) in Hong Kong to analyse students' perspectives on SBA within high-stakes systems. The study shows an increased awareness of the unfair, inaccurate, accountability grading of SBA, resulting in anxiety among students (Brown, 2011). In this sense, Brown (2011) argued that teachers' conceptions of assessment reflect the legal, cultural, and social priorities and elements coming from their environments that should be considered. Other studies also emphasise this interpretation by reporting that teachers' conceptions of assessment could be related to the structures and characteristics of the educational system and also to policy (Remesal, 2011; Yates & Johnston, 2018). Once again, assessment practices are regarded by the literature as responding to specific contextual realities.

The SBA literature shows that students know that their performance assessment is related to teaching and learning improvement (Brown, 2011). Students tend to attribute school success (limited to grades) to hard work, giving the impression that improvement hinges on students doing better, working harder, and being less lazy (Sivenbring, 2019). Secondly, assessment processes have an impact on students' emotional wellbeing and on the quality of the relationships they have with classmates. As has been stated before, previous literature suggests that assessment processes duly affect students' learning but also that the self-concept (as it is linked with labels of 'good' or 'bad' students), and perceived differences among classmates and students, subjects, or social group (Stiggins & Conklin, 1992). Finally, students become aware that assessment can be unfair, negative, subjective, or even irrelevant in their lives (Brown, 2011). This ethical aspect will be further analysed in the following sections of this study (section 2.7, page 64).

In the present research, the interest in teachers' and students' perspectives will also be focused and channelled through the possible changes and challenges to the assessment system, with its concomitant implications of students repeating the year and university entrance and the fairness of the system. These topics that have been introduced recurrently covered in Chapter 1 and will be revisited transversally in the three main questions that will guide this study- they are: (1) teachers' and students' perspectives on assessment purposes and methods (2) teachers' and students' perspectives on results, in terms of grades and feedback, and (3) teachers' and students'

perspectives on fairness of the assessment system. It must be noted that the topics for teachers' and students' perspectives have been complemented with document analysis in RQ1.

2.7 Assessment methods (RQ1)

This section will summarise studies on assessment methods, highlighting different topics that are relevant for its study and understanding. Assessment methods are key to SBA as they allow the collection of multiple sources and types of evidence about students' learning within the natural classroom setting during the school year (Davidson, 2009). Literature on assessment methods was also relevant to decide on which elements of SBA this study would focus and guided the research questions and interview questions. Therefore, assessment activities to foster learning must be appropriate, allowing even the development of materials to suit students (Leung & Andrews, 2012). Studies of assessment typically incorporate teachers' perspectives of assessment practices and methods, as they have an impact on what they do in the classroom (Bota & Tulbure, 2015; Earl, 2014; Hartell & Strimel, 2019; James & Pedder, 2006; Young & Kim, 2010). The assessment process helps teachers to consider their different roles and tasks inside the classroom (Mäkipää & Ouakrim-Soivio, 2019). As SBA allows teachers to decide about the activities that the students perform, the methods used are a central point of interest. In that sense, SBA should allow for a wider range and variety of tasks, integrating assessment during the complete learning process (ARG, 2002; Yates & Johnston, 2018). However, the literature shows consistently that even when teachers report the use of a variety of assessment methods, in practice traditional assessment styles are found (such as paper and pencil tests), which has an impact on students' results (Bonner et al., 2018; Bota & Tulbure, 2015; Mäkipää & Ouakrim-Soivio, 2019; Puad & Ashton, 2021; Smyth & Banks, 2012). There are critics especially focused on test dominance as having a negative effect on students' learning. There is also evidence of a higher use of multiple-choice tests and recall activities in subjects such as Mathematics, which contrast greatly with the methods used in other subjects such as English or Arts (Duncan & Noonan, 2007). This highlights the appropriateness of each assessment method, following the nature of what is being taught. In contrast, Kennedy and colleagues claim the design of the assessment should not include multiple-choice questions that encourage surface learning in the form of factual recall-type questions (Kennedy et al., 2008). Summative work used to grade, such as tests, appears to be forcing teachers to adopt assessment practices that go against formative assessment (Black et al., 2010).

As an additional characteristic of the assessment methods used within SBA, the dynamic implies a one-way flow where teachers present assessment information to students, without any further negotiation, placing the responsibility of achievement on students alone (Sivenbring, 2019). Within this dynamic, students follow guidelines assuming that assessments are inevitable and conforming, while teachers reward or punish students due to their power to assign grades (Sivenbring, 2019). This classic form of assessment where grades have a central and key role, will be further explored in the following section.

Bota and Tulbure (2015) studied teachers' assessment styles and students' results from Romanian secondary schools. Their results show that teachers use traditional assessment styles associated with quantitative aspects of evaluation, namely, grading. Interestingly, they also use the concept 'evaluation' for the classroom process instead of 'assessment', as it occurs in the Latin American research context. These summative evaluations are centred on knowledge and the authors express their concerns about the lack of accounting for students' attitudes, personality traits or behaviours (Bota & Tulbure, 2015). The study characterises the alternative assessment as including teachers considering the quality of students' performance and including a focus on improvement and supporting students' learning activities by the use of methods such as projects, portfolios, or investigation. Bota and Tulbure (2015) highlight how alternative evaluation can be adequate in certain contexts and that teachers who prefer it are inclined towards a focus on evaluation adjustment, optimization, and improvement of the process, offering feedback, establishing clear evaluation criteria, and respecting students' needs and learning potential. As occurring in Latin American studies, the authors do not use the term 'formative assessment' to refer to these characteristics, even when their views concur with the literature on the topic. This reinforces the relevance of understanding and being aware of how similar formative concepts apply in different countries despite the different terminology used. According to the results of the study, students had better results using alternative evaluation, as the diverse methods allowed them to show their strengths, to apply knowledge to real-life contexts, also encouraging group activities. Alternative evaluation similarly appears to be allowing teachers to observe students' progress and the evaluation of superior cognitive capacities (Bota & Tulbure, 2015).

Along the same lines, studies like Pope et al (2009) reported a narrow range of classroom assessment practices for reading and literacy in 35 countries. As it has been described in prior sections, narrow assessment methods usually imply classic paper and pencil tests that do not

always reflect the best way students' learning. This highlights the need for a variety of assessment methods used in school contexts (Flitcroft & Woods, 2018).

In their study on Finnish students' perceptions of assessment practices in secondary school, Mäkipää and Ouakrim-Soivio (2019) analysed which assessment practices are used and how the different roles and tasks of assessment are understood by students and their teachers, including possible improvements. The study also included items on students' understanding of their feedback and grades received. The Finnish SBA contains formative and summative assessments, such as feedback based on test results, but the policy underlines the relevance of the formative assessment practice. Therefore, the current requirements for secondary school assessment emphasizes the use of versatile assessment methods and feedback to students deepen their learning. However, students reported the use of traditional strategies with low student participation (Mäkipää & Ouakrim-Soivio, 2019). They reported that 54% of the students did not experience versatile assessment methods and did not receive any feedback.

A further study by Zhao et al (2017) also studied assessment methods. They were categorised as observation, questioning/discussion, in-class assignments, after-class assignments, presentations, portfolios, quizzes, and written tests. They can be used after a period of learning, and be based on one lesson, topics or learning units. These categories result in an interesting frame to guide the analysis of assessment purposes and methods elsewhere. However, they also talk about the high variety of assessment methods which can be found inside the classroom.

SBA allows teachers to exercise absolute control and to have the power to decide how and what to assess (Baird et al., 2014). This will depend on the specific policy and guidelines for each country. However, as shown, teachers are regarded by the literature as not taking full advantage of the freedom available to them for their assessment-products creation (Harlen, 2004). It must be understood that teachers' freedom depends on the level of control of SBA that each school has, following national policies. As exposed in the aforementioned sections, they tend to emulate the form and scope of external tests or make narrow interpretations of assessment products (Harlen, 2004). The extensive use of tests can have detrimental effects on learning as students simply memorise what is needed to pass without allowing them to apply their learning to wider contexts. Tests also restrict methods to only what is necessary to pass them. Hartell and Strimel (2019) identified teacher-created testing as being a common way of assessing students' knowledge in Sweden, defining this as "a form of written assignment consisting of questions and/or problems to be answered/solved by pupils individually during a limited period of time" (p. 783). Testing indeed

shows several consequences, such as tests' outcomes being more valued than the learning in and of itself. This certainly generates a focus on numerical results that associated to failure demoralises students, generating anxiety, especially among girls (ARG, 2002). The frequency of the summative assessment can also have a negative impact on formative, especially if it occurs more often than the necessary for reporting aims and generating judgemental feedback that promotes comparison among students (ARG, 2002). Classic tests used as assessment method will also be restraining students from using higher order thinking skills, not being able to adapt to changing circumstances and not understanding how to learn, and, impacting in this manner their ability to work and learn independently and also collaboratively (ARG, 2002). Therefore, there is a complex dynamic regarding assessment methods when they are used which is they leave many unintended consequences and present important challenges. The available literature shows them with no variety and being impacted by the high stakes where they are regularly implemented.

2.7.1 Assessment methods and external testing

The literature reports incongruence between assessment practices valued by teachers and their real practice, possibly for the impact of external testing as they want to prepare their students for them (Bonner et al., 2018; Harris & Brown, 2009; James & Pedder, 2006). The relationship between SBA and high-stakes examinations is critiqued, as it might reduce the breadth of the curriculum, emphasising the subjects tested (Smyth & Banks, 2012). This also contributes to teachers being encouraged to teach the test routines, motivating students to learn test-taking skills rather than contents (Smyth & Banks, 2012). Particularly, it encourages them to prepare students for high-stakes tests against the enactment of formative assessment, generating student resistance to changes within the classroom (Box et al., 2015).

Additionally, it is relevant to consider that teachers succumb to social conditions, changing their practice to fit with accountability requirements or being flexible between their personal conceptions and social mandates, such as to help their students to prepare for national testing. This is especially true when the students come from lower socio-economic backgrounds and cannot afford preparatory classes (Rasooli et al., 2023). Additionally, it can have consequences in a reduced instructional time for non-tested subjects and increasing time for test practice (Smyth & Banks, 2012). Hong Kong is an example of how the high-stakes system has impacted teachers' understanding of the purposes of the assessment, which is sensitive to policy priorities (Brown et

al., 2011). In the study, students agreed more with items related to external examinations than related to SBA (Cheng, et al., 2011). That is, students prefer external examinations rather than formative assessment or their teachers assessing them summatively. This is a consequence of teachers considering it more important to prepare their students for examinations rather than following government suggestions on assessment for learning changes (Chan, 2007). Not surprisingly, Chan's (2007) paper is titled "We have various forms of assessments but only summative assessments count", which seems to perfectly summarise the focus of the assessment methods usually used by teachers in different countries. In Hong Kong's context, students use private tutorial schools to compete successfully for university places (Kennedy et al., 2008).

Studies on assessment methods also talk about restricted teaching assessment and methods, as teachers end up emphasising practice for exams, which is required by the high stakes system (Murchan, 2018). For instance, in the USA a body of research has documented restrictive pedagogies in urban schools in the country, associated to test-centric practices (Sondel, 2016). However, research also reports that teachers considered standardised tests as not providing accurate measurements of students' skills and abilities (Bonner et al., 2018). Conversely, teachers use these test formats within school subjects in order to help students but invite them to question about the surface rather than deep learning that they promote (Kennedy et al., 2008). Teachers understand that this practice can benefit students as they are prepared for assessments with university formats. Therefore, even when they express dislike for this practice, they show an understanding of students' disadvantages if not exposed to them (Harris & Brown, 2009). Therefore, the strong impact of external examinations and testing on SBA cannot be easily dismissed. Smyth and Banks (2012) reported that high-stakes examination strongly shaped teaching and learning in secondary schools, reducing activities that are more engaging and student-centred. Nevertheless, students showed that they understood the 'rules of the game', focusing on their success and following them but reporting high levels of stress and pressure in doing so. As with other studies on the topic, students from middle-class backgrounds increased the amount of time that was used for study and taking private tuition to be prepared for their exams. This resource mobilisation allows them to invest in learning activities that help them succeed (Smyth & Banks, 2012). In this way, exams shape how assessment is conducted inside the classroom and the main consequence seems to be that teachers use practices and methods that they do not necessarily believe in (Bonner et al., 2018). This situation generates reports of discrepancies and inconsistencies between what teachers say and what they will do in the

classroom (values) and what they actually do in practice (James & Pedder, 2006; Monteiro et al., 2021). This discrepancy is extended to what policy, or the national curriculum suggests about versatile assessment methods (Mäkipää & Ouakrim-Soivio, 2019). Teachers' conceptions of SBA are impacted by their juggling to satisfy multiple pressures and interests, especially about how they assess their students and also about how they use the data resulting from this (Harris & Brown, 2009). Pressures are identified by teachers as protecting their students from harm and thus helping them to learn, but also complying with employer obligations and seeking school improvement (Harris & Brown, 2009). Consequently, in Harris and Brown's study teachers' tensions between balancing the school's institutional needs and student's needs were omnipresent.

2.8 Information provided by the assessment: perspectives on grading and feedback (RQ2)

As it has been stated before, assessment procedures inside the classroom provide valuable information about students learning. Internationally, some authors relate the process of 'learning' with feedback and 'achievement' with reporting (Stiggins, 2002), evidencing that the dichotomy between formative/summative aspects is extended to the information provided by the assessment. Understanding that the feedback discussion involves that grades could be considered a form of feedback, this study will follow the Chilean policy on SBA that identifies students' grades as part of summative assessment strategies and feedback as part of formative assessment strategies (Mineduc, 2018a). Grades are the most common form of report even when summaries can adopt the form of grades or narratives that are sent to parents and/or school members/inspectors (Earle, 2014). On the other hand, formative feedback is used to improve and adjust instruction (Black & Wiliam, 2009). The quality and type of feedback provided to students also plays a key role in their learning (Hattie & Timperley, 2007). Logically, the measurement result in grades (marks) not only should have an impact on the teaching and learning process but also on students' motivation for learning (ARG, 2002). However, whereas feedback can positively incentivise students' future work, grades can also demotivate students and reduce or eliminate entirely their level of participation in classes (Brookhart, 2001; Kožuh, 2019). Therefore, following the Chilean policy on assessment, the assessment information within SBA is studied considering

both of these elements, which will be defined and explored in the following sections: grades and feedback on SBA.

2.8.1 Grades on SBA

In relation to grades obtained in SBA, they are usually part of high-stake systems (Cheng et al., 2011). The results of these often have an impact on promotion, certification, and university entrance. Some examples of countries using grades as part of high stakes systems and examinations were presented in section 2.3. As grades are an important part of SBA, the symbol associated with each student at the end of an instructional period carries many interpretations across different situations and countries (Cabral & Baldino, 2019). Grades obtained by SBA represent a percentage of the final annual results, being the remaining percentage obtained by national examinations/tests. For instance, as 15% in English in Hong Kong (Tong & Adamson, 2015), 40% in Ireland (Murchan, 2018), variable among sectors of Australia (Cumming and Maxwell, 2004) or alarmingly 100% in the Chilean case (Mineduc, 2018a)¹. Therefore, it could be used in combination with other forms of external examination, especially for the certification purposes which can be found at Secondary School Level (Cumming & Maxwell, 2004). Grades are also related to students' accountability for their learning (Harris & Brown, 2009). This accountability is extended to their parents, who receive school reports describing if the student is learning what they need to learn or not (Puad & Ashton, 2021). In the case of Chile, grades become the only language that parents understand as they are the only information they tend to receive about their children's learning (Florez Petour, 2014). Results are an important element for parents, who need to know if the student is making progress and especially when thinking about further study or employment (Reyneke, 2016). In this way, what happens inside the classroom and the school with assessment results goes beyond this space and has consequences in the family context.

Teachers' and students' perspectives can help understand various elements concerning grading strategies: for example, how the grading methods are communicated to students, students' characteristics, their interests, their intellectual habits, and/or their misconceptions are

¹ Final results in Chile and therefore, promotion, are 100% based on grades assigned by teachers, with no external examination. However, for students applying to university entrance, the numerical average of grades obtained during the complete secondary school is considered (NEM). Further details to follow in Chapter 3.

integrated into grading methods (achievement, learning, ability, motivation, effort, interest, personality, socioeconomic expectations, and even cultural background), scales, point-accumulation, deciding on borderline/repeating year, reasons for students' success or failure and students' intellectual habits and misconceptions (Ball & Forzani, 2009; Stiggins & Conklin, 1992). Therefore, grades production implies several associated elements, with complex interrelations. This is especially true when the different understandings of the meaning and purposes of the grading strategies used by teachers are also taken into account (Brookhart, 2013). Among these lists of elements related to grading, achievement appears as one of the major factors that teachers consider when grading their students (McMillan, 2001). Students' performance is usually judged against prescribed programs and curriculums (Harris & Brown, 2009). For instance, in Australia, students' grades are reported against descriptive performance bands which give information about the student's knowledge, understanding, and what they can do (Cumming & Maxwell, 2004). However, teaching the curriculum can be also the main philosophy of work for teachers, linking students' work exclusively to contents (Rasooli et al., 2023).

A review made by the Swedish Research Council showed that there is little scientific evidence supporting the idea that grading or measuring students' attainment actually led to better results in final assessments (Sivenbring, 2019). Besides, grades are also regarded as encouraging high achievers and discouraging and overwhelming lower achievers (Harris & Brown, 2009; Sivenbring, 2019). For instance, in Cheng et al (2011), SBA is perceived differently by students, depending on their higher or lower competence. Grading, especially within high-stakes systems is also linked to problems of reliability and demands that increase teacher workload (Harlen, 2004; Murchan 2018). On the one hand, the lack of objectivity risks present in the SBA are also problematized, precisely because of their potential impact on university entrance (Tong & Adamson, 2015). On the other hand, placed within the SBA framework, a main concern shown by the research is the extra burden and pressures that it implies for teachers (Murchan, 2018). Another difficulty highlighted by Harlen (2004) is that the high-stakes system impacts teachers training students to pass rather than use the assessment to improve their learning. Similarly, grades within the high-stakes systems tend to prepare students to individually compete in the global market rather than to facilitate their education so as to participate as citizens in a diverse and interconnected world (Rodgers Gibson, 2019). In that sense, assessment serves to order and classify students to later decide their future and social position (Murillo & Hidalgo, 2015).

Grading procedures are complex as some subjects could give more importance to social behaviour, such as effort and attitude (Stiggins & Conklin, 1992). Measurement research on grades warns against the use of 'enabling factors' as they will distort the meaning of the grades as achievement representation. The recommendation of measurement for grading purposes is aimed solely at achievement, in terms of knowledge and skills over effort or other variables (Stiggins & Conklin, 1992). As reported by Shepard (2019), awarding points for effort, work habits, and collaboration (performance), lead to the normalisation of behaviours that end up having an impact on grades. In that sense, Shepard (2019) states that to grade students considering their personality or socio-emotional dimensions is neither necessary nor defensible. These practices, which are usually labelled as 'score pollution' imply that teachers usually consider different elements to take decisions upon grades, as they include informal observations about students on different aspects such as effort, behaviour or work (Remesal, 2011). Thus, the questioning of the reliability of teacher-assigned grades is always present, especially when teachers have the freedom to decide upon them and about if they are helping students to learn (Stiggins, 2002). Londoño-Restrepo (2015) stated that the elements of power and discipline within the testing culture inside schools were highlighted in the theoretical work of authors such as Foucault in 1984. Tests determine if a person possesses true knowledge, which is socially labelled as such. If the person does not possess that knowledge, there are mechanisms to sanction them. Grades work in that sense – they are punitive by nature. In any case, the literature shows conflicted views as to whether grades should reflect students' personal features, for instance when increasing a grade based on students' effort, which is especially relevant when considering the stakes of the system in several countries worldwide.

Research shows that the information obtained from grading could be given without providing guidelines, knowledge, or ideas about how to improve, rendering students powerless about how to change their outcomes (De Luca et al., 2018; Sivenbring, 2019). This gives the false impression that improvement depends only on students' efforts to do better, working harder, and being less lazy, with no relation or mention of how teachers' actions can also improve or damage learning. Students seem to simply conform themselves to this dynamic, as teachers have the power to assess and assign grades (Sivenbring, 2019). Perilously, the only information resulting from the complex teaching and learning process could be easily one grade and nothing more. Worse than this, grades results are regarded as not impacting teaching strategies, despite the information that assessment outcomes can provide about students' learning (Cheng, 2019). Very

often, students feel that solely being assessed with a grade is frustrating, as Mäkipää and Ouakrim-Soivio (2019) showed in their study on students' perspectives on assessment practices in Finland: "...a grade does not reflect all the work that has been done in the course. I would like teachers to give more oral/written feedback in which there are words, not only grades" (p. 31).

In a similar vein, Irving et al (2011) reported that when the grading process results in lower grades than expected, students' self-esteem, personal growth, and sense of achievement are also damaged. This results in teachers not using the results for formative purposes as they are related only to summative requirements. Overall, grading systems exhibit several complex aspects and difficulties, that once again respond to specific country-related realities that follow different policy styles. They are briefly summarised by the high stakes of SBA, how grades are produced and what is graded, achievement, grades' impact on motivation, their reliability/score pollution, and teachers' work overload by grading. In any case, the stakes of each school system seem to have a great impact on how grading works, leaving intended and unintended consequences on the process.

2.8.2 Formative Feedback on SBA

Grades by themselves (in terms of a number or letter) are regarded as providing information about students' learning which is insufficient for the purpose of enhancing and improving learning itself (Sivenbring, 2019). Thus, timely and constructive feedback appears as a key element of formative assessment, as it has a significant impact on students' learning, helping them to move forward (Black & Wiliam, 1998; Cheng, 2019; Stiggins, 2002; Tong & Adamson, 2015). It is defined as instructions for learning improvement given by teachers (Beaumont et al., 2011), which carry a formative aspect following the Chilean assessment policy (Mineduc, 2018a). The guidelines for the implementation of SBA in Chile defines feedback following Moss and Brookhart's (2009) ideas, as they regard instructions as actions that help students "to identify next steps that can be taken by them in order to improve" (Mineduc, 2018b). Internationally, feedback is supported as improving learning and providing more useful information than comparative feedback in the form of grades (De Luca et al., 2018). In that sense, if grades are important to parents, teachers, and students, feedback is more relevant for the students themselves to truly enhance their learning (Reyneke, 2016). A core idea behind feedback is that students need to understand their own

learning goals and progress through processes that are, in themselves, cognitive events (Dann, 2012).

The key aspect of feedback is that it is oriented towards reducing learning gaps by helping students modify their understanding as it is connected to different cognitive processes (Tong & Adamson, 2015). It is mostly related to the previous explanation and sharing of learning goals and successful criteria for students, generating classroom discussions by the use of questions and students' collaboration (Black & Wiliam, 1998; Wiliam & Thompson, 2007). Feedback, then, ought to guide the student towards missing or incomplete information given by them, confirm if their answers are correct, guide the direction that the student should follow in order to improve and indicate alternative ways to understand specific information (Tong & Adamson, 2015). As such, it is an activity where teachers and students are engaged in a way that requires time to make corrections and to be effective and impactful for students' learning (Black & Wiliam, 1998; Hattie & Timperley, 2007). For that aim, feedback should point towards task processing and encourage students to self-regulate their learning, while also promoting self-reflection. The activities related to feedback also present a possibility for students to collaborate among themselves as learning resources doing peer feedback.

In a similar vein, Johnson et al (2019) studied the enactment of formative assessment by surveying 1,097 participants in the United States. Specifically, about feedback, they reported that teachers provided limited individualized feedback to students, even when they were observed effectively using other formative strategies such as questioning, collaborative activities, instruction adjustment, and connecting the lesson to learning goals. However, if the feedback intends to be effective it needs to attend to several other characteristics, such as not being hurtful, being fair and making clear that aims to help the student (Tong & Adamson, 2015). In this regard, Adamson (2011) stresses the importance of feedback that matches the cultural context where is embedded, which is a key issue of interest in this study. In that sense, teachers' and students' beliefs and practices should be a key element to look at in relation to their context (Adamson, 2011).

Zhao et al (2017) analysed 266 articles on classroom assessment practice in mathematics including a special focus on the feedback provided by teachers. In the study, classroom assessment is considered in teachers' hands, with the purpose of informing teaching and learning by using feedback and grades, but within low-stakes contexts. The feedback category was analysed through its focus (task-related, process-related, or person-related), its nature (positive, negative, and what needs to be improved), its mode (verbal, written), the timing (immediate or delayed),

and the teaching adaptation prompted by feedback. Results indicated that teachers considered classroom assessment to be equivalent with feedback, despite other elements considered by formative assessment (Zhao et al., 2017). These categories were considered to analyse feedback in Chile in the research design of this study (see p.107). De Luca et al study (2018) shows that feedback is generally classified by students as positive or negative, depending on its relation to correct things or things to improve upon. The study also reports that most students identified the feedback as helping them to improve in different subjects, depending on their achievement. Thus, they received more feedback in the subjects where their achievement was lower. Contrarily, Mäkipää and Ouakrim-Soivio (2019) documented how Finnish students perceived that they did not receive any feedback during the course but only received grades, and some of them did not remember having received any type of feedback whatsoever. These different studies allow to observe that feedback is experienced in a variety of ways in different countries, but in any case, being mainly impacted by the presence of grading. This highlights the need to explore teachers' and students' perspectives on feedback within schools using SBA in order to know how the experience is lived within the classroom and support the learning process.

2.9 Fairness of the assessment process (RQ3)

The international literature on ethical assessment provides evidence of the need for an educational assessment engaged with values associated with the aims and purposes of education (Biesta, 2009). Authors such as Burns et al, (2019), recognise that the concept of fairness in relation to assessment needs to be explored. This understanding should also include the ethical and moral impact of the assessment choices not in relation with what should be done but rather what is rigorous and valued to be played within a specific educational system (Elwood & Murphy, 2015). Furthermore, Rasooli et al (2023) have stressed that fairness has been, historically, a central issue in assessment from teachers' and students' perspectives. This is especially true if as educators, we consider that what we understand by the term education is not always explicit, even less their aims and purposes. In this study, fairness will be aligned with Brown-Jeffy and Cooper's (2011) approach to equity, which implies giving students "what they need" Understandably, the comprehension of what students need is variegated and once again, these needs recognise the diversity of cultures where the assessment process is conducted (Burns et al.,

2019). In that sense, what is fair to students in relation to what they need will be connected with socio-culturally responsive assessment, as described in section 2.3.

Several studies mention the concept of 'fairness' and/or 'ethics' linked with the assessment procedures, despite not always fully defining or analysing the ethics of assessment. Thus, it is a complex concept as it is used in diverse ways to refer to different elements of SBA. This could be explained because these elements respond to specific aspects of the assessment that are relevant in one culture but could not be in another one.

SBA produces results that have consequences for students' lives, especially within high-stakes systems. In that sense, it is pertinent then to ask for the ethical consequences of the process. Colnerud (2006) explained that teachers balance between justice and care in their practice in a daily basis, therefore they are expected to take into account ethical standpoints explicitly when teaching. In the assessment scenario, teachers are expected to use grading procedures, discuss results, and also recognise and resolve ethical dilemmas or inappropriate assessment methods (Mäkipää & Ouakrim-Soivio, 2019). In this vein, Cheng et al (2011) explicitly recognised the challenge of the key components of SBA such as fairness, especially for its implications for teaching and learning. Thus, the need to articulate that language and ethical standards of practice by opening spaces within research appears as evident. The assessment procedure should always be open to questions about what is right or wrong for each student, from a moral point of view (Estaji, 2011; Pope et al., 2009). When studying assessment methods used by teachers, the traditional summative assessment seems to be used despite prompting concerns about unfairness and/or invalid measures of the students' capabilities. Some of the concerns present in the literature on the topic are that SBA is regarded as work loading teachers and presenting insufficient time for its development (Cheng, 2019). In terms of time, SBA requires a huge amount of teachers' time to design their assessments, prompting tasks and materials variety among schools pose challenges in terms of its value and fairness (Cheng, 2019). The ARG in England (2002) also stated that teachers' assessment is perceived and can be unreliable and biased by increasing teachers' workload. Consequently, moderation systems seem to be necessary to assure the quality of the assessment process (ARG, 2002). This moderation is especially relevant in contexts where results are used for external high-stakes purposes, such as university entrance in Chile.

Examples of different uses of the concept include work on fairness and authenticity, especially for qualifications (Yates & Johnston, 2018), the fairness of feedback (Tong & Adamson,

2015); ethics behind 'teaching the test' routines (Murchan, 2018); unfair assessment processes due to their design or context, especially in a sense of categorising and organising students more than support learning (Gardner et al., 2009); assessment results as unfair and as unreliable measures of students' capabilities, questioning its moral appropriateness (Beets, 2012; Black et al., 2010); fairness as a protection of students rights (Green et al., 2007); a need for teacher values which are committed within a moral framework which considers student learning (James & Pedder, 2006); critics of its negative effect on teacher autonomy, being deemed as unfair for students and not protecting them from harm (Estaji, 2011; Harris & Brown, 2009); for cultivating relations of irresponsibility (Malaby & Esh, 2014) and the need of a more consistent sense of ethical assessment practices (Green et al., 2007). The reduction of students' motivation and self-esteem by imposing stressful conditions that besides, promote shallow and superficial learning rather than deep understanding has also been identified in its relationship with assessment ethics (ARG, 2002).

This myriad of elements associated with the fairness of assessment concept implied for this study a free approach that asks generally about SBA's fairness in the Chilean system (see pages 107-108). In this way, ideas that teachers and students associate with fairness and fairness could be found without conditional answers to any specific possible aspect that is involved with the concept. This flexibility is important specially when the lack of teachers' explicit training on fairness in assessment is considered (Rasooli et al., 2023).

The study of Pope et al (2009) involves the analysis of the ethical implications of the assessment decisions that teachers make. They analysed how teachers attempt to keep a balance between students' learning and their grading duties. They made participants respond to a single question (n=103) describing situations where they found it difficult to know what was ethically/morally right or wrong when doing classroom assessments of students and/or ethical issues encountered. The findings showed teachers' concerns about the pollution of grades by assigning points for effort and harming students by results impacting their placement and promotion. These points of concern for teachers are interesting for exploring assessment information (RQ2), in terms of grades, in this study. This is because within the discussion on ethics behind SBA, the concept of score pollution comes to the fore. As was previously stated, this practice consists of including within scores other factors than achievement, such as effort, behaviour, or work (Remesal, 2011). This means that teachers must balance institutional versus and students' needs (Hirsh, 2020; Rasooli et al., 2023) and consider the 'no harm' principle (Green

et al., 2007; Pope et al., 2009). Harm could arise from multiple assessment situations: in measurements prompting stress and anxiety (especially within high stakes systems), and also in harming the teacher-student relationship and trust. Green et al (2007) considered that the 'no harm' principle is especially related to protecting the rights of the individuals affected by assessment, including principles related to fairness. However, the study also makes explicit that the consideration of what is harmful results problematic, and it is often a matter of judgement where specific contexts matter. Other studies, such as Brookhart et al (2016) also focus on fairness in grading practices. Grading practices should include multiple sources of information and should include to make clear to students what has been assessed and how it was graded. In this sense, fair grading should not only include arithmetical procedures but also variations according to the students' circumstances, teacher experience and perception of equity, consistency and accuracy. However, these elements are considered to constitute 'score pollution', which is regarded in the literature as being a practice that should be avoided (Green et al., 2007; Pope et al., 2009). All these elements invite us to the question of what grades reflect and under what sort of circumstances.

Generally speaking, conflicts were based on a clash of institutional requirements versus students' needs and also teachers' and students' needs (Pope et al., 2009). Different requirements promoted ethical dilemmas, making it clear that students' needs are a main source of concern for teachers (Pope et al., 2009). These dilemmatic spaces for teachers are brilliantly described by Hirsh (2020) as a situation "where their professional judgement told them one thing and policy directives something else" (p.99).

More recently, Rasooli et al (2023) interviewed 27 high school teachers in Iran to explore their conceptions of the fairness of classroom assessment. The study highlights the importance of social structures in the school context as contributing to shaping teachers' and students' conceptions of fairness. In this regard, teachers raised caring strategies that compensated students from disadvantaged backgrounds, considering it fair to adjust grades, ask questions, and consider students' attendance. Teachers' conceptions of fairness were also influenced by accountability and assessment mechanisms as they identified summative exams as unfair for students from disadvantaged schools. Overall, Iranian teachers' conceptions of assessment fairness were highly influenced by their students' cultural and socio-economic conditions.

Along the same lines, Brown (2011) acknowledged the negative effects of assessment by its 'irrelevance', a concept which is associated in the study with the assessment process being seen by

students as bad, subjective, or unfair. Using an abbreviated version of the inventory created by Brown (2007, 2011), Gao (2010) found an increased awareness of the unfairness of grading within high stakes SBA, leading to a greater level of anxiety among students in Hong Kong (Brown, 2011).

Elwood (2013) and Murillo and Hidalgo (2015) centre their ethical discussion on the impact of assessment (and grades) in terms of inequality and social justice, especially for the social consequences that are implied within high stakes spaces: without justice, assessment is helping to reproduce inequalities. Based on Messick's studies, Elwood (2013) stresses the relevance of focusing on equity, fairness, and bias on assessment. Assessment, within a high-stakes system is, in the study, a matter of ethics related to fairness and equity. The study also highlights the need for a code of ethics articulating basic ethical principles for assessment (similar to those followed through educational research): equal treatment, respect for persons, avoiding harm, and doing good. In the Chilean context, Gysling (2017) conducted multiple interviews with Chilean pre-service teachers and academics, identifying three principal ideas: the assessment is a process, to assess is not to grade and the assessment must be fair.

Along these lines, Smyth and Banks (2012) precisely state the importance of considering students' reflections about their learning experiences as research topic. They consider explicitly the relationship between high-stakes systems and educational inequality by considering the negative effects of tests on lower-income students (Smyth & Banks, 2012). Students familiar with the dominant culture, which they consider to be operating as cultural capital, exhibit better academic results, leaving working-class young people at a disadvantage. In Smyth and Banks (2012) study, external testing is shown to impact how knowledge is taught within the school, prompting excessive time from test preparation, narrowing the curriculum and also neglecting knowledge and subjects not included in the tests. At the school's internal level, there is also evidence of an increased focus on grading that also consolidates or exacerbates rather than decreases socioeconomic differences in teaching (Strandler, 2016). When studying assessment in secondary school and similarities between Nepal and England, Davies (2016) also highlights the accusations of bias in the role of teachers assessing students' coursework and also the impact of socioeconomic circumstances on successful rates.

As SBA can be inserted within high-stakes systems, fairness and authenticity are also seen to be considered important concerns. Fairness could be protected by elements such as time limit and providing reference materials, reducing cheating, and generating equal assessment conditions (Yates & Johnston, 2018). Gardner et al (2009) agree that assessment is made to categorise the

learner rather than to support their learning as it is related to entry to university or qualifications. Consequently, they also highlight the detrimental effects of national and standardised testing on students' motivation and self-esteem. The authors also suggest that tests and examinations may have increased negative consequences for students who enter the process at disadvantage. Assessment will, therefore, serve these students badly making the process unfair and revealing disadvantage or discrimination (Gardner et al., 2009). In addition, James and Pedder (2006) concluded that teachers are committed to fairness values and not just to methods of formative assessment within a moral framework that considers the quality of students' learning as important.

Overall, the concept of fairness is used with enormous variability and relates to elements like assessment process, purposes, methods, grading, feedback, and the inequality brought by its social consequences within high stakes. This is especially true in countries with well-reported high correlation between socioeconomic status and academic results, such as Chile (Honey & Carrasco, 2022). The Chilean researcher Gysling (2017) recognises that the relationship between assessment and justice is complex and involves dimensions where interrelated aspects are intertwined: equality, social equity, and the merit concept. The relationship between high-stakes assessment and educational inequality has been largely highlighted by a broad base of research, as students from lower income suffer disproportionately negative effects as score distributions consistently correlate to their socioeconomic status, family income and parental education (Grotsky et al., 2008; Reardon, 2021; Smyth & Banks, 2012). Reardon (2021) highlights the importance of studying the economic achievement gap between students from high and low-income families, as this gap is a measure of how (un)equally educational opportunities are distributed.

Consequently, the fairness of these elements that are inevitably present during the assessment process should always be considered when researching the topic, as they shape teachers' practices and most importantly, students' lives and opportunities. This will be especially researched throughout RQ3, in order to identify which aspects Chilean teachers and students link to the assessment 'fairness' concept.

Chapter 3: School Based Assessment in Latin America and Chile

*Chile, “the country that bought the neoliberal prescription and executed it faithfully”.
(González, Guell and Márquez; 2012).*

3.1 Introduction

This Chapter presents the specific characteristics of SBA in Latin America, and specifically Chile, given that, this country is the research context for the current study. The following sections will present relevant features of research on assessment in the Latin American context and also display the educational features of the school system in Chile that are worth understanding as this is the contextual reality where students’ assessments take place. In that way, specificities in which SBA is studied and understood in Latin America will be presented, highlighting authors that have an impact in how the assessment process has been historically understood. This section will be followed by a descriptive and reflective account of the Chilean educational system. In the first place, the impact of the segregation and inequality found in the country and linked to the implementation of neoliberal policies in the 80’s will be explained. These conditions will configure the structure and organization of the educational system and its external assessment policy: SIMCE and PAES. Both tests will have an influence on what students learn, based on the national curriculum, and how teachers assess, following the decree created by the Ministry of Education. However, curriculum and assessment are presented by the Chilean literature as containing critical points that will allow further reflection on the Chilean educational setting. Finally, this chapter will also expose critical elements of teachers’ working conditions in the country. Taken into account, the Chilean educational system features are of vital relevance to later understand how assessment purposes and methods (RQ1), assessment information, in terms of grades and feedback (RQ2) and the fairness of the system (RQ3) are presented in the country.

3.2 Evaluation in Latin America in its own terms

The Argentinian researcher Toranzos (2014) stated that worldwide changes in assessment have been slower for some countries, generating a gap between ideas with a renewed pedagogical

discourse and the concrete evaluative practices that are happening within the educational system. In this sense, it is unsurprising that often the tone of evaluation studies in Latin America has an entirely different focus and meaning, as the literature on assessment in English does not reach Latin America as easily as expected. Toranzos (2014) characterises the educational evaluation in Latin America as carrying strong meanings associated with external control, a penalty function, sanctioning, selection, inherent judgements about sufficiency or insufficiency of determined aspects, and the criteria for administration of rewards and punishments. All these different meanings imply that the evaluation process has a strong disparity between their criteria for analysis. Besides, these deeply rooted meanings attributed to the evaluation generate different stereotypes that are difficult to modify to install a renewed evaluative practice (Toranzos, 2014). Additionally, it is worth recalling that in Spanish, the word 'evaluation' (evaluación) is the only one that exists with no available translation for "assessment". Martínez (2019) wisely states that, when studying evaluation practices in Chile, "language about evaluation serves to understand and also to confuse us". In an attempt to clarify this, the Chilean researcher Gysling (2017) explains that in English three words can be used to refer to the various processes of assessment: assessment, appraisal, and evaluation, which are not distinguished at all in Spanish. Therefore, in the Spanish language, the topic is studied only by deploying that one concept, which can cause semantic confusion as it is not a univocal concept and involves several aspects of 'evaluation' (Santos Guerra, 1993). The author stated that in Spanish, we talk about evaluation diversely to refer to comparison processes applied to results' measurement; the process of accountability imposed by authorities to schools; the diagnosis that allows quantification of data, and also when students' learning is checked (Santos Guerra, 1993), an aspect where this study has been focused. Recognising this, Maldonado-Fuentes (2021) suggests that *evaluation* (in the Spanish language) must be understood as a polysemous and multidimensional concept that has a different conceptual basis, coming from different ideas and theories of evaluation. Understanding assessment purposes and methods can help identify this conceptual basis (RQ1). In this vein, Escudero (2003) and Maldonado-Fuentes (2021) have stressed the plurality of the concept, given by the multiple ideas about evaluation that exist and also by the wide theoretical-practical concepts generated. However, the aforementioned does not necessarily mean that evaluation concepts can be used interchangeably (Escudero, 2003). This is especially true if we consider different aspects of 'assessment', such as formative and summative, which are present in the international literature on the topic, worldwide promoted and known, (see pp. 36-38). However, it

has also been stated that the concepts are not widely used in the study of evaluation in Spanish, especially if the general process of evaluating students is considered. As an example in Chile, Martínez (2019) has researched evaluation practices in a school in Santiago, identifying that teachers' purposes in the evaluation process are varied, such as measurement and classification but also teaching and learning improvement, as part of the changes of new policies used in the country. However, those purposes are not labelled as summative or formative in her work; rather, they align with how those terms are understood in the international literature on the topic.

Concerning the influential literature, Santos Guerra's work provides a common conceptual framework that can be found throughout the studies of several researchers on evaluation whose work has been conducted in Spanish and has been highly influential on how we talk about evaluation in Latin America. He wrote a set of pioneering articles for educational evaluation in the Spanish language, such as: "General Pathology of Educational Evaluation" (1988), "The evaluation: a process of dialogue, understanding and improvement" (1993). A main idea based on his work, which still influences the evaluation study in Latin America and therefore, cannot be avoided, is that a mixed view is presented. This entails the idea that a number does not determine the human being is found (and defended), along with a technological-positivist dimension that seeks concrete data to make decisions on promotion, accreditation and failing (Santos Guerra, 1996). There is certainly a clash between different ideas and purposes, that have been described, for instance, within the formative/summative concepts. In his work, Santos Guerra (1988) describes different elements regarding these pathologies, such as; only the student, knowledge (contents), direct results, observable effects, and negative aspects are evaluated. Both, in a decontextualised isolated fashion and quantitatively, utilising inadequate instruments, assessing in a way that is incoherent with the teaching/learning process; with an evaluation that is competitive, stereotyped, and unethical. As noticed, the tone is based on criticisms of the summative evaluation process regarding its negative features. Students receive unquestionable results, usually in terms of grades, for which they are considered the only responsible party. Curriculum contents, even when necessary, should be thought of in terms of their selection, articulation, and significance for the thinking that they organise. The author also highlights that these features continuously repeat practices of memory evaluation. At the time of publication, Santos Guerra already stressed the need to consider within the evaluation processes the radical inequality in the natural and contextual conditions of the students. We should not and cannot assign meaning to students' performance stripped of its context. This is reinforced by researchers such as the

Argentinian Perassi (2008), who stated that to know the sociocultural contexts where students come from, contributes to obtaining an interpretation which is more linked to the complexities of each given reality. However, Santos Guerra went beyond this, by identifying the evaluation as a moral phenomenon and not merely technic: it is not important only to well-evaluate but also whom the evaluation is serving to. In that sense, evaluation should always consider specific values and a sense of fairness, favouring the teacher, the institution where it is done, and especially, the evaluated student (Santos Guerra, 2014). Fairness remains as a main relevant aspect in this study (RQ3).

The influence of Santos Guerra's work on the study of evaluation in Spanish is prevalent and significant, as research still considers the 'pathologies of the evaluation' concept. An example of what has just been stated is the Ecuadorian researcher Hernández-Nodarse (2017), who reflects on the reasons that make it difficult to transform the evaluative practices in Latin America. In his work, the key for transformations is to capacitate teachers to transform the evaluative practices, rethinking policies, strategies, implementations, and actions. Hernández-Nodarse (2017) states that the lack of change and improvement in the evaluation processes seems to be an epidemic without a cure, where the persistence of the traditional evaluative practices have not been organised and systematised in order to create actions that are coherent, efficient and formative. The author uses the formative concept as something which ought to happen but is 'torn apart' by the summative aspects of the evaluation in the region (Hernández-Nodarse, 2017). A main point in his work is to reject the positivist aspect of the evaluation, putting the students at the centre, following new educational and social demands. A key point of this study is to explore whether the introduction of a new assessment policy in Chile has, in any way, changed this situation on assessment purposes, methods (RQ1), results (RQ2) and the fairness involved in the process (RQ3).

Contrasting with this reality, Earle (2014) for instance, stated that the distinctions that are made between formative and summative purposes have received important attention in the UK in the previous 15 years. However, this dichotomy is seen as part of a paradigm of Western knowledge (Morín & Motta, 2002), as Latin American aim to see the evaluative practice with a holistic understanding (Pizarro & Gómez, 2019). Therefore, the assessment process is seen as the 'evaluation', which is a unique process happening in the classroom but involving several different aspects related to each sociocultural space, which usually lacks Western distinctions and language. Additionally, and despite the worldwide emphasis on the use of feedback as part of the assessment procedure, in Latin America this is still typically seen as mainly measurement and

grades (Pizarro & Gómez, 2019). This highlights the classic disjunction between quantity (grades) and quality (of learning) that has a pivotal role in the Latin American discussion (instead of “summative” and “formative”), which appears in teachers’ reflections (Pizarro & Gómez, 2019). In this vein, Pizarro and Gómez link back to Santos Guerra’s ideas (1996), especially on the technological-positivist dimension of education, which constantly requires and demands concrete data. Moreover, this feature is currently present in evaluative practices across the continent. Specifically, Pizarro and Gómez (2019) studied the evaluation practice in Colombian schools, identifying clear differences and disconnections in what teachers think, say, and actually do. They also highlight the lack of studies on teachers’ coherence between their evaluation conceptions and practice in the Latin American context. Some of these ideas on triangulation between what teachers say and do have been included in the research design of this study (see Chapter 4). Furthermore, the authors explain that it is common in Latin America to impulse reforms inspired by other countries’ models which are doomed to fail, as they do not consider the transformation or shifting of teacher beliefs, which is a slow construction process (Pizarro & Gómez, 2019). Usually, reforms are based on foreign ideas which fail to acknowledge the need to build new conceptions of evaluation from the teacher, and also take account of their interactions with the context (Pizarro & Gómez, 2019). This study and others have identified evidence of teachers’ strong resistance to changing their evaluative practices, maintaining them as traditional summative, mainly due to the effort and time required for the task to be carried out (Hernández-Nodarse, 2017; Nodarse et al., 2018; Pizarro & Gómez, 2019).

Agreeing with Toranzos (2014), the authors highlight the difficulty to ‘ground’ national policies according to schools’ reality, a process that can be slow and complex (Pizarro & Gómez, 2019). Therefore, there is a double challenge: on the one hand, to adapt policies that are usually taken from somewhere else and, on the other hand, to allow teachers to contextualise them following school reality. Internationally a good example is the work conducted by Brown et al (2011), which has identified the same challenge for East Asian societies, that are quite different to Western cultures in a number of aspects, such as history, language, ethnicity, and the like. This allows us to see how expectations of different cultures lead to differing policies and conceptions of practices on assessment (Brown et al., 2011). It is reasonable that each country and culture will adapt policies reflecting cultural differences, making those relevant to understanding the specific features of the various spaces where assessment is conducted. In Latin America, a traditional assessment practice has been corroborated in several countries: a qualitative study about

discourses and practices on classroom assessment in eight countries seeking to introduce SBA concluded that the process is mainly associated with grading and tends to be more summative than formative (Picaroni, 2009). Even when teachers value the formative aspects of assessment, there is no relationship between their efforts to help their students improve their learning and the assessment that is carried out, as arguably it does not guide students towards improvement (Picaroni, 2009). Another study conducted in Latin America by Hernández-Nodarse et al (2018) revealed that the main evaluation's objective is to verify and check students' knowledge, exhibiting a reduced vision of the evaluation objectives and also identifying that the evaluation contributed to their learning in a moderate to low/null way. The study also corroborated a low variety of evaluation techniques and stereotyped evaluations, where reproductive and memory questions were used.

In Chile, Gysling (2017) claims the Ministry of Education is always trying to promote an authentic evaluation system while criticising the evaluative practices that can be found in the school. This could be related to teachers' education, as in the case of Chile, pre-service teachers are still taught using ideas rooted in the 70's with general units of evaluation. It is important to mention that each university has the freedom and autonomy to design pedagogy programs, without alignment or regulation related to their curricular plans (Montecinos et al., 2015). However, teachers seem to adapt and nuance their beliefs on evaluation when they face the school reality, interacting with other educational actors and understanding the school evaluative culture, making the evaluation concept more complex (Maldonado-Fuentes, 2021).

Besides the ideas from Santos Guerra, in Chile, the work of the national researchers Condemarín and Medina (2000) has been guiding the evaluation topic, following their book "Evaluation of Learning" (Gysling, 2017). They stated that the traditional evaluation approach found in the Chilean system stymies educational change for a number of reasons, such as making teachers feel insecure about the validity of their evaluation in innovative practices and utilising tests as instruments for evaluation (Condemarín & Medina, 2000). The evaluation process aims for the authors to be 'authentic', and is based on the permanent integration of learning and evaluation by the student itself and their classmates, being an indispensable requirement of meaning construction and communication processes (Condemarín & Medina, 2000). At the centre of the evaluation discourse, rather than concepts of formative and summative, ideas about the authenticity of the process and its quality can be found, with constant criticisms about the presence of foreign ideas, which can prove difficult to implement. However, the Ministry of

Education does promote those foreign ideas with its evaluation decree and the use of summative and formative assessment. This will be further explored in the following sections, starting with a description of the Chilean educational system in order to understand the socio-cultural context where this study was carried out.

3.3 Chilean educational system

Chile has become an example of both school segregation and educational inequality due to the neoliberal policies which have increased the levels of social precariousness and impacted negatively in public education in the country (Bellei & Vanni, 2015; García-Huidobro, 2007; Gysling, 2017; Oliva, 2008; Oliva & Gascón, 2016; Puga, 2011; Zancajo, 2019). Despite Chilean economic growth in the last decades, the country presents profound social inequalities and has one of the highest levels of unequal income distribution in the world (Cabalin, 2012). Chile is shown to have the lowest levels of school social inclusion and one with the highest levels of school segregation among OECD countries (Zancajo, 2019). Not surprisingly, the educational topic in Chile is usually guided towards social justice objectives and tends to be framed considering the segregation and inequality promoted by the privatisation and marketization of education. The privatisation of education started in the 80's during the dictatorship occurring in the country, when educational reforms redefined the role of the state in education, privatising it and generating an educational marketplace (Gysling, 2016). Cabalin (2012) states that "this model has conserved the privileges of dominant classes, increased segregation, and caused inequality between a small elite and the majority of the population" (p. 221). Decades after the neoliberal policies commenced, the educational system has generated 'first class' students, who attend private and exclusive schools and have better performance on national tests, with increased opportunities to attend the best universities and for their future. 'Tourist' students attend poor public schools with lower results than the national average in tests and if they have an opportunity to access university, they must study with expensive loans (Cabalin, 2012).

An example of policy responding to this was the creation of a new voucher system known as 'SEP' back in 2008, which is focused on students who were identified as socio-economically vulnerable and disadvantaged (Subvención Escolar Preferencial in Spanish, Preferential School Subsidy in English) (Zancajo, 2019). Schools could receive additional funding if they accepted priority students and more than 70% of schools in the country have participated in this optional program

(Honey & Carrasco, 2022). However, the use of that money has been highly criticised, as several mayors of different districts (for public education) and private administrators have been accused of fund embezzlement. Sadly, this includes the district where I was born and where this study was conducted, San Bernardo, whose ex-major, Nora Cuevas, is being legally investigated by the fraud section for approximately 9 billion pounds (UChile newspaper, November 25, 2021).

Related to its structure, the Chilean educational system is based on four levels: preschool education (pre-kinder, 4 years old and kinder, 5 years old); basic education from Year 1 to Year 8 (6 to 13 years old), secondary education, from Year 9 to Year 12 (14 to 17 years old) and university education (from 18 onwards). The country has 13 years of compulsory education, from kindergarten to Year 12. Three types of schools can be found in the country, depending on their administrative and financial dependence: public (with state funding), subsidised (with private owners but receiving state funding), and private (privately owned, administrated and funded) (Honey & Carrasco, 2022; Roco, 2010). Most of the students in the country attend subsidised and private schools rather than public establishments (Contreras et al., 2011). This study focused on one subsidised school. Chile has 57% of its educational places in the subsidised and private sector. Within the segment of people in age for secondary school studies, 92% to 97% are enrolled in secondary schools (Mineduc, 2018b). This percentage is similar to other OECD countries but is higher than countries like Mexico and the United States (Mineduc, 2018b). However, there are 14% of young people who are not attending school, due to performance problems or expulsion (Mineduc, 2018b).

Promotion from one level to another is based entirely on school grades (SBA), with the possibility to repeat the year if the minimum requirements are not complied by students. SBA procedures involved in teachers' decided activities and students' performance represent 100% of grades that allow students to be promoted from one level to another, and be graduated from the secondary level at the end of Year 12 (Mineduc, 2018a). Students' average grades in the school system do not involve any compulsory external examination (Mineduc, 2018a). They are only based on the teacher's evaluation and the grades obtained in different school subjects, taken annually. However, for accountability purposes, the system uses the SIMCE test (National System of Learning Results Evaluation, for 'Sistema Nacional de Evaluación de Resultados de Aprendizaje' in Spanish). Specifically, for secondary school students, where this study is focused, the university examination has been voluntary for leavers, in a national test called PAES since 2022 (Admission Test for Higher Education, Prueba de Admisión a la Educación Superior in Spanish). Both tests will

be further explained in the following sections, as they could potentially have an impact on assessment purposes and methods (RQ1), and in the case of PAES it is linked to grades (RQ2). Furthermore, considering the inequality of the country and the marketization of the educational setting described in this section, both tests could also be related to the concept of fairness (RQ3).

3.3.1 External assessment: SIMCE test (National System of Learning Results Evaluation)

In time, the marketization of schools in the country promoted greater segmentation and inequality that was attempted to be controlled and regulated using national tests (SIMCE) to assess their quality (Gysling, 2016). SIMCE is an external examination during Years 4, 6, and 10, designed for school accountability in terms of the quality of education measurement. In fact, the last fifteen years in Chile have been characterised by testing times with accountability purposes, with the country taking part in international testing and expanding its SIMCE (Gysling, 2016). The results do not represent any percentage, grade, or score in the students' results at any level.

The impact of this accountability is that teachers must use their time to prepare students for such tests: the focus will be the learning measured by national tests (Gysling, 2016). Thus, the comprehensive curriculum contents are externally measured in the Secondary Schools by SIMCE in Year 10. SIMCE is externally designed and administered, therefore, not part of SBA.

3.3.2 PAES test for Secondary School leavers (Admission Test for Higher Education)

PAES is taken in the final year of secondary school, based on performance standards and on reporting levels of achievement. The measurement is reduced to language, mathematics, and social and natural sciences. However, 30% to 50% of the final score that is used to determine university entrance is based on the grades (NEM; Secondary School Grades, for 'Notas de Enseñanza Media' in Spanish) obtained during four years (Years 9 to 12) of secondary school (www.demre.cl). Divisively, the system also considers ranking by grades (since 2013), a score produced by the students' secondary grades average position among all the students graduating that year in the school (Larroucau et al., 2013). The inclusion of NEM and ranking has been thought as recognising students' effort during secondary school, regardless of their school type or socioeconomic situation (Larroucau et al., 2013). In that sense, they attempt to recognise socioeconomic inequality in the country, reacting against it.

During the process of carrying out my research, the test used in Chile to access the university went through several changes. During the social uprising occurring in the country in October 2019, the students protested against segregation and the inequality of the test. This was unprecedented in the country's history, once again reinforcing the urgent need to consider students' voices within educational research. Students' demonstrations highlight that they can step back from the system to critique it and that they are capable of active and critical reflections about their learning and how it is administered (Smyth & Banks, 2012). It must be noted that in the past, the country witnessed two massive student movements in 2006 and 2011 protesting against the neoliberal educational policies in the country (Cabalin, 2012). Chile has a strong history of student resistance to neoliberalism (Stromquist & Sanyal, 2013). One of the most relevant effects of these protests is the political and ethical discussion in relation to the competence logic and the market orientation of the educational system (Villalobos & Quaresma, 2015). During the 2019 social uprising, Piñera's right-hand government prosecuted student leaders of the 'No more PSU' movement. They were forbidden to give the test and faced charges under the 'State security law'.

Research between school grades and PAES scores (former PSU) shows a low correlation between them, where school grades as internal evaluation system shows a reduced predictive capacity for PSU scores (Rodríguez & Jarpa, 2015). The internal evaluation system of schools is labelled as poor and carrying little validity and reliability levels. The correlation is worst for people coming from the poorest sectors and public schools. This provides an interesting focus for new research looking at grades and how the internal evaluation system works, that this study covers (RQ2).

Data on university access shows that the higher percentage of young people entering university comes from upper-class backgrounds (Espinoza & González, 2015). This is brilliantly described by Stromquist and Sanyal (2013): "Despite having a strong economy and an educational system highly respected in Latin America, Chile shows persistent trajectories of inequality, with substantial differences in student achievement by type of school and considerably different life chances depending on one's family wealth" (p. 172). Therefore, students coming from other socioeconomic backgrounds are the victims of a type of double discrimination. First and foremost, their socioeconomic status and secondly, their inadequate preparation during secondary school to face the tests which determine entrance to university (Espinoza & González, 2015). These criteria have been identified as promoting competition and rivalry (Percherisa & Falabella, 2017) between

students while also reinforcing a school culture which privileges numerical results and individual success rather than any sense of solidarity or empathy (Arzola-Franco, 2017).

It is well known that evaluation practice, including teachers' and students' perspectives, are linked and framed within national policy mandates (Bonner et al., 2018). The Ministry of Education (Mineduc) is responsible for creating policies in order to promote the education in all its levels while also creating study plans and programmes and assessment policies. Both elements, curriculum and assessment policies, will be explored in the following sections.

3.3.3 National curriculum in Chile: subjects' plans and programs

Referring to national curriculum contents in schools in Chile, Mujica-Johnson (2020) explains that it has been created by the Ministry of Education for primary and secondary schools. These plans contain minimum compulsory contents that should be taught in each level/grade, guided by learning objectives that the teacher should develop with the students. The result is then summatively assessed, assigning one grade to students (Mujica-Johnson, 2020). Each plan contains curriculum contents and learning objectives for each subject and level, with some examples of activities and assessment strategies that can be used (Law N° 290.370, www.curriculumnacional.cl). There are compulsory subjects and elective subjects only for the secondary level, where students can choose between Musical Arts or Visual Arts, and in Year 11, among different elective subjects. Since 2020, schools have a 30% of the school hours for free disposition of elective subjects suggested by Mineduc (<https://www.curriculumnacional.cl/614/articles-37136bases.pdf>). The elective units for free disposition are school-decided and organised in three areas, where students can choose three in Year 11 and three in Year 12, following their interests.

The curricular bases are set and organised by Mineduc (www.curriculumnacional.cl) at five levels: curricular basis for preschool education, for Years 1 to Year 6, for Years 7 to Year 10, and curricular basis for Years 11 and 12. The last change in curriculum occurred in 2020, when some plans and programs were updated and others presented for the first time, especially for the new elective subjects structure.

As part of their autonomy, teachers decide how to cover learning objectives, a freedom that involves contents being treated differently at each school in terms of the information volume and the practices associated (Mujica-Johnson, 2020). This freedom is also extended to evaluation

instrument creation. The Ministry of Education does not provide the evaluation instruments for all the learning objectives and contents which are standardised at a national level; rather, they suggest only a few examples. Thus, teachers create their evaluative instruments following their own personal criteria, which inevitably involves their ideologies and beliefs (Mujica-Johnson, 2020). Generally, all the evaluation material created by teachers is reviewed by a coordinator, who is responsible for checking the quality of the products (tests, rubrics, questions, guidelines) before they are applied to students. This is not explicitly requested by the Mineduc assessment decree, and therefore, not compulsory. It is done this way as the coordinator orders the photocopies of the pedagogical material needed. The characteristics of the assessment policy in the country will be further explored in the following section, as it details how the learning promoted by the national curriculum is assessed.

3.3.4 The assessment policy issued by Mineduc

Mineduc has been historically responsible for creating policies that guide the assessment policy in the country. The document states that they attempt to “promote evaluation practices that propitiate students’ focus on their learning, finding meaning and relevance to it, and to be motivated to keep learning and ensuring that evaluation and grading practices are focused on the national curriculum” (Mineduc, 2018a, n/p). Evaluation in Chile is understood as actions that are led by professionals of education to obtain and interpret information about students’ learning (Mineduc, 2018a). The objective is to adopt decisions that allow learners to achieve learning progression and to receive feedback about their learning processes (Mineduc, 2018a). A main concern found in the guidelines is to promote students’ progression and to ameliorate their learning processes. The guidelines proposed the use of summative and formative assessment within the school, coincided with prevailing international understandings of SBA, and reinforced the interest shown by Mineduc (2017; n.d) before the publication of the decree currently in use, issued in 2018. Formative assessment was promoted in the Mineduc discourse before the new decree (Mineduc, 2018a). However, this study seeks to explore in detail whether the classroom dynamic remains largely dominated by quantitative evaluation (measurement) or whether the formative aspect is present in teachers’ purposes and methods (RQ1). In this regard, it is relevant to state that the new decree proposes only two new ideas for the assessment process: the intention to reduce/eliminate students’ repetition by implementing strategies to support students

at risk and allowing schools to choose how many grades students will have in each school subject, allowing for greater flexibility (Mineduc, 2018a).

In relation to school repeating, refers to students who do not pass their respective classes because they do not acquire the necessary competences to be promoted to the next educational level (Torres et al., 2015). School repetition is a multidimensional problem influenced by socio-demographical, academic, and socio-emotional elements (Meriño-Montero, 2021). In the case of Chile, the social inequality found in society highlights the segregation that families suffer. Therefore, the poorer, more vulnerable, and more excluded that students are, the higher the possibilities are for them not to obtain the learning needed, to have good performance, to fail a year, to have low school attendance, or finally, to quit school (Román, 2013). These types of students are categorised as 'priority students' and receive special funding from the State, as explained in the prior section 3.3. Related to academic elements, learning difficulties, low attendance, and personal characteristics (such as organization and autonomy) are stated as relevant for school repetition (Meriño-Montero, 2021). Finally, some socio-emotional elements that have an impact are self-esteem, motivation, behaviour, emotional stability, and parental support (Meriño-Montero, 2021).

The new Chilean assessment decree states that repeating, which has always been considered in the educational system, should be an exceptional measure, sufficiently justified where the school must provide support to the repeating student (Mineduc, 2018a). Therefore, it involves the idea that the educational system can help students by supporting their competencies, abilities, and knowledge that were not previously accomplished (Villalobos & Béjares, 2017). However, the document does not contain any theoretical or pedagogical justification for the strategy.

The national decree does impose some requirements about grading. Following Mineduc guidelines, grades are understood as learning gains represented using numbers, symbols, or concepts (Mineduc, 2018a). However, it makes a final year grading compulsory, prompted by summative assessment, in a numerical scale ranging from 1.0 to 7.0, considering one decimal and with a minimum passing grade of 4.0. The numerical average obtained by students in each subject generates an annual average which allows them to be promoted or otherwise. A student is promoted when all the academic subjects are passed; yet, when one subject is failed but the overall minimum average reaches 4.5, or when two subjects are failed but the overall minimum average is 5.0. Consequently, it can be understood that the evaluation system allows for the failing

of the academic year if the student does not comply with such requirements, which configure a high-stakes system, as cited before. Despite worldwide teacher in-classroom assessment has been related to formative role (Harlen, 2004), the opposite occurs in Chile where evaluative procedures have been historically reduced to grades, which reflects a very restricted activity of 'evaluation' (Córdoba, 2009). Grades are the only empirical data and indicator of students' learning, and this is the only data that parents receive and historically have understood (Flórez Petour, 2014). There is no register of students' learning expressed in language other than numbers. This generates quantitative data, expressed using numerical scales which allow the school community to make judgements about students' achievements (Córdoba, 2009).

During 2020, Chilean schools were supposed to be implementing a Ministry of Education decree on evaluation (Mineduc, 2018a). The international literature on the topic recognises that these periods of reform are seen as promoting reflection about the purposes of education and assessment procedures which situate this project in perfect time (Cizek, 1996). This reform promoted a holistic school review of evaluation practices to implement compulsory discussions and changes before March 2020, when the new academic year began, making it an important time to study evaluative practices.

The new assessment decree establishes minimal regulations for the process in the classroom as it states that each school must define its own internal decrees, determining the type and number of evaluations used (Mineduc, 2018a). This goes in line with the role assigned to the State in the 80's as an external regulator, minimally implied with processes occurring inside the classroom (Gysling, 2017). Martínez (2019) identified a contradiction in the sense that teachers are requested by Mineduc to use a diversity of assessment methods but they are also requested to use traditional measurement test formats that agree with external tests (SIMCE, PAES). In that sense, the Mineduc guidelines only request them to offer 'diversity' without explicitly explaining the ways this can be achieved, generating a disparity between teachers' strategies and making it easier and quicker to use written tests for evaluation, because these allow compliance with grades requirements on specific dates and at specific times (Martínez, 2019).

One interesting aspect of these internal school decrees is how the educational goals are proposed. Internationally, they can be understood as "the priorities that education systems set for improvement, to ensure that all learners are achieving the competencies, skills, and attributes with which the system aims to equip them" (SWAAP, 2017, p.7). However, as the national decree contains minimal regulations for internal evaluation occurring within the school, many of the

intentions of the decree, in terms of formative and summative assessment, are left up in the air as each school can decide how to enact it.

3.3.5 Critics on literature around the national curriculum and its contents

In this regard, Gysling (2017) highlights that Mineduc itself qualitatively researched and evidenced that teachers tend to evaluate and focus on curriculum contents without considering any of the learning objectives in which they are framed. This fact allows us to understand that teachers evaluate their students in an isolated and decontextualised way, focusing on data management and theoretical concepts without their proper integration within a wider understanding, reproducing an 'encyclopaedic' curriculum that does not form on thought abilities either the performance required in these days (Gysling, 2017). Despite Chilean teachers being responsible for designing their evaluative instruments and deciding on which learning objectives and contents to focus on, the reasons behind teachers' self-developed tests use in Chile are not explicit in any of these Mineduc decrees. This allows us to consider the fundamental collective failures of the system, the flawed logic of the subject curriculum, or the strategies adopted by teachers to attain learning outcomes (Sánchez & Jara, 2019).

Teachers' focus tends to be on curriculum coverage of contents, without balancing the completion of learning objectives and their students' abilities in the process, in what can be labelled as an evaluation process seen as a process of 'information acquisition'. By way of illustration, a study conducted by Hernández-Nodarse et al (2018) in Ecuador, stated that 53,3% of teachers identified the first objective of the evaluation to be obtaining information about the level of learning of their students about some contents in particular. Teachers (28%) also identified as an evaluation objective knowing what their students know and what they do not, always based on contents, thinking on new ways to teach them. In Chile, a study conducted by Martínez (2019) also showed that teachers understand the evaluation process as being a measurement, but some of them measure students' capabilities or abilities in relation to specific units or contents. The evaluation also is conceptualized following the curriculum objectives, which is declared by teachers in each class as requested by the school. In this way, evaluation is usually linked first and foremost to curriculum contents.

The role of the Ministry of Education in the evaluation phase is merely to suggest evaluation indicators and some activities (contained in the national plans and programs), which may then be used or not by teachers. This situation makes it relevant to recall that the evaluation creation is a

complex and time-consuming task for teachers (Cejas & Alvarez, 2006; Martínez, 2019). Therefore, the evaluation process has been labelled by some authors, such as Ávila and Paredes (2015), as controversial and is discussed only because it is usually reduced to instrumental practices.

Hernández-Nodarse et al (2018) studied the different purposes assigned to evaluation among teachers in a primary school in Ecuador. They concluded that teachers believe that the critical purpose of an evaluation is to provide information. This is explained by the authors by the instrumental rationality assigned to evaluation as a merely technical process of the certification of products or the learning results proposed by the official plans and programmes contained in the national curriculum. Once again, plans and programmes and their coverage certification are found to be at the centre of the evaluation practice.

Despite the curriculum being focused on comprehensive education, measurement “neither promotes nor seeks to ensure its implementation” (Gysling, 2016, p.20). The curriculum alignment with external measurements contains telling reflections about external results (SIMCE, PAES) in contraposition with internal results (grades) that are unquestioned. This phenomenon promotes evaluation procedures that are not connected, ‘sacrificing’ curriculum and internal classroom dynamics in terms of their external measurements (Gysling, 2017). Within this logic, in the country, it is only valid to teach what is assessed by those tests, known as ‘teach the test’ routines (Perassi, 2008). About external testing for accountability purposes, Flórez Petour and Rozas Assael (2019) state: “On the one hand, it reduces learning to language and mathematics in written formats and to mechanistic pedagogical approaches. On the other hand, this invisibilises cultural and socio-economic diversity by turning it into a problem to be solved rather than an opportunity for enrichment and social cohesion” (p. 174). Thus, a complex situation is promoted whereby, on the one hand, the assessment decree keeps evaluation anchored to grades with promotion purposes and, on the other hand, there is increasing pressure for results to be achieved in national measurements. The aforementioned reduces the curriculum and models the evaluation made through multiple choice tests and high-stakes university selection system that, in turn, increase the value of secondary school grades (Gysling, 2017). None of these features of the educational system are made to promote or even sustain new evaluative practices. Conversely, they are designed to promote an evaluative practice inside the classroom which maintains traditional features, privileging summative processes that end up in grades, mainly through tests where information-repeating is assessed (Gysling, 2017). Evaluation is, consequently, a coercive barrier which can give information, within different stakes, about teachers, schools, policies and the

educational system itself (Gysling, 2017). The same does not occur with the internal measurements in the form of grades: the results are not questioned in terms of the school's effectiveness.

3.3.6 Problematics around the assessment decree and assessment practice

The inequalities of the country must be necessarily linked to the evaluation criteria, as the school acts by defining and distributing knowledge within society (Bernstein, 2001). The evaluative practices determine what students can learn and how they will learn it; what teachers teach, and how they assess it; in other words, the products and processes of education itself (Briggs & Ellis, 2013; Córdoba, 2009). If this is linked to the inequality of the educational experience lived by students, following their socioeconomic level, it can be supposed that evaluation in between sectors will differ: it will be more demanding for the elite and one of less or scarce demand for popular sectors (Bernstein, 2001). Martínez (2019) recognises that in this regard in Chile, to conduct large research or analyse past rules is not necessary in order to understand what the evaluative culture ten years ago was: written tests (mostly multiple choice) with inflexible and individual results. However, these ideas are still the very basis of what teachers, directives, parents/carers, and students understand as being evaluation (Martínez, 2019). This deeply rooted understanding about what is assessment within the classroom will bring different consequences, in terms of summative assessment and grading predominance, score pollution, and top-down policies that fail to take into account teachers' and students' perspectives before suggesting changes. These elements will be further explored in the following subsections.

3.3.6.1 Summative assessment and grading predominance

Shepard (2019), when studying assessment for learning or formative assessment in the UK, explained that summative assessment procedures involving national examinations seem, in large part, to guide classroom assessment dynamics, despite this having been regarded as undermining the development of classroom cultural practices that support deep learning. Although the context is somewhat different in the case of Chile, the new guidelines do not explain how to protect the intentions of formative assessment practices from school-based grading and summative requirements, which has been highlighted as important by international literature on the topic (ARG, 2006; Shepard, 2019). Moreover, the new decree maintains the status-quo in terms of the

requirement of SBA numerical grades production and does not include any formal requirement for a formative assessment practice register (Mineduc, 2018a). The new policy suggests different activities as formative but always makes explicit that grades can be given for such activities (Mineduc, 2018a). From the multiple and different meanings that assessment procedures could potentially have, it is mainly focused on checking students' learning by employing numerical grades production. It is stated by the Chilean educational assessment researcher Flórez Petour (2014) and Flórez Petour and Rozas-Assael (2018), that the setting is deeply coloured by neo-liberal ideology and consequently, has been invaded by discourses of numbers and mechanical notions of learning. Those numbers are supposed to reflect students' learning, which represents a deeply reductionist and empiricist view of what constitutes learning (Barret, 2011). As numbers prompted by teachers grading procedures and achievement put a central focus on summative assessment in the Chilean context, it is unsurprising that measurements of achievement form the basis for what is understood to be quality education and are central to the activity of teaching (Briggs & Ellis, 2013; García & Paredes, 2010). The key element is that grades are teacher-produced, with no external examinations. This has been the historical assessment style used in the country.

With regard to this, Gysling (2017) recognised that both perspectives are not irreconcilable, nonetheless even quite recently the quantitative perspective is still arguably privileged, under which the qualitative is subordinated: the metric aspect is the common expression of performance. Consequently, any evaluation inside the classroom implies several meanings, including the sanctioning of the learning process and also operating as a control device (incentive or threat) for both learning and behaviour (Gysling, 2017). In the Latin American context, evaluation is highly focused on grades and numbers, namely, summative evaluation which comes into tension with formative purposes required as part of the new SBA policy. Indeed, the predominance is marked by evaluation as the final product rather than as a process (Arzola-Franco, 2017; Vera et al., 2017). The need for formative evaluation is strongly emphasised nowadays in the Chilean context, but few guidelines for formative evaluation implementation remain as optional and voluntary (and also unknown) on the Mineduc webpage and guidelines.

There are no changes to the new evaluation policy in the summative processes which consume schools, teachers, and students' time, rendering formative assessment implementation even more difficult. Therefore, the schools receive conflicting messages whereby simultaneously, their numerical results and formative assessment are demanded (Arzola-Franco, 2017; Vera et al.,

2017). As was set out in Chapter 2, internationally and nationally the literature indicates that high-stakes testing systems sometimes come into conflict with the priorities of formative assessment, and thus, there is an inherent tension between a formative approach and the persistence of these high-stakes systems where change is enacted (Arzola-Franco, 2017; Baird et al., 2014). According to this reductionism of evaluation to merely grades, teaching is seen as a lineal transference, where what is evaluated is the gain of curricular objectives, not the competencies development or any deep transversal understanding coming from the school experience (Gysling, 2017). Grades are also punitive and are associated with exerting control over students, it does not promote learning and barely represent it, but rather distort the pedagogical process as grades become an aim themselves (Gysling, 2017). The Colombian researcher Londoño-Restrepo (2015) stressed in her research that the relationships between education and evaluation has been marked by power relations where testing and punishment have been used to evaluate in a sanctioning and punitive fashion.

Classroom evaluation requires teachers to make judgements about students' achievements measured with standard scales expressed, as "it is important for teachers to follow agreed procedures if TA (teacher assessment) is to be sufficiently dependable to serve summative purposes" (Harlen, 2004, p. 6). For example, a 60% of correct responses can be required to obtain the passing grade, which is 4.0. This is the same as for evaluation instruments creation, points distribution on the evaluative instrument, which is teacher-decided. The details of these aspects could be explained in the internal school decrees created by each school, or not.

For example, Gysling (2017) states that everyone agrees that they are graded because the system requires it. Therefore, even when formative assessment implies that students' cooperation and participation are required, this is difficult to accomplish as they have been socialised in evaluative practices aimed at obtaining a grade. She explains that what is relevant is to pass, not to learn, leaving teachers who want to use formative assessment facing resistance from their own students and even from their families. Pizarro and Gómez (2019) recognised that teachers move towards a mixed evaluation precisely because they have the autonomy to act, but they always come back to the quantitative evaluation trying not to depart from the qualitative aspect. Morín and Motta (2002) identified this problem of separating quantity and quality as the disjunction paradigm, ideas that are also present in Santos Guerra's (1996) work, who stated that when the evaluation is focused on the value of knowledge, the relevance of it is to learn. Conversely, when the evaluation acts as an interchange where a grade is assigned, the relevance is simply to pass

(Santos Guerra, 2003). Therefore, the grades problem is a relevant part of the literature in Latin America, that identifies the evaluation found in the school as 'mixed' within a disjunction paradigm where different ideas on evaluation are placed altogether, which is a main feature of SBA.

Overall, the evaluative dynamic in Chile seems to be directed solely by summative purposes, the prerogatives of teachers, to be conducted in private, to be concerned mainly with measurement (grades) and failure and to be focused on being reported to others rather than students, as parents and guardians (Brooks, 2002; Flórez Petour, 2014; Harlen, 2004; Young & Kim, 2010). There are no formal requirements in the Chilean policy for evaluation results communication to students, just optional guidelines that, as explained, can be respected or not without any further revision.

3.3.6.2 Score pollution

The national decree suggests considering the gap between repeating students and their classmates, providing pedagogical support for low-level students at risk of repeating. However, there is no control regarding how each school enacts these principles in actual practice. These strategies are not further explained in the new guidelines for evaluation documents (Mineduc, 2018b). Therefore, grading is open to practices such as increases in grades, or assigning points for accomplishment, responsibility, or punctuality, to recognise work done (Gysling, 2017). Some teachers grade the work done in terms of effort, dedication or participation, without considering the quality of the product (Gysling, 2017). These practices reported in Chile have been labelled as 'score pollution', a strategy that has been described in Chapter 2 as increasing grades by considering elements other than academic performance. Nevertheless, the Latin American literature recognises that several teacher's practices can only be understood in reference to conceptions and intentions that guide them as well as the particular context where they are used (Prieto & Contreras, 2008). In Chile, Martínez's study (2019) revealed that teachers also used grade-improvement strategies to increase students' grades by modifying the numerical scale used and eliminating some of the questions from tests. A similar study is Peters and Contreras' (2019), who also found this practice in the form of extra 'tenths' which are used to improve a grade, even when in their study the practice is not found at all in private schools but rather subsidised. The grades increase is also related to personal criteria such as participation and good behaviour. In this regard, score pollution is insightful to be studied in the relationship with the reasons that teachers

and students identify as causing them, which is a relevant part of grades as assessment information (RQ2) and must be contextualised within the high stakes of the Chilean system.

3.3.6.3 Top-down assessment policies in Chile

Considering the aforementioned, the new policy presentation is a continuation of the historical top-down policy approaches, in which a highly centralised policy potentially fails to take account of the local contexts and obstacles for implementation (Cohen et al., 2011; Flórez Petour, 2014). These types of actions have been regarded as studies of evaluation reform processes which have given rise to fundamental inconsistencies and have even been described as being outright chaos (Flórez Petour, 2014). Mineduc attempts to reinforce the use of formative assessment mostly based upon ideas that are taken from other contexts without any due consideration of the particularities of the Chilean classroom reality in a bottom-up sense. That is to say, from students' and teachers' points of view, from the schools. Even when the national assessment decree document does not include any reference, the guidelines for its implementation for school principals and teachers do. They mention authors such as Bennett (2011), Black and Wiliam (1998, 2009), Brookhart (2008, 2009) Cauley and McMillan (2009), Hattie (2009), Heritage (2010), and Newton (2007). However, they also cited Chilean authors such as Condemarín and Medina, (2000) and Gysling (2017) (Mineduc, 2018b). This knowledge is rooted not in what is happening on the ground but ensures the approach which has been rightly criticised by Flórez Petour (2014). This refers to Chilean assessment policies based on foreign models that are attempted to be used in the Chilean system, in a "things going on elsewhere" fashion, with foreign experts being a common element. Similarly, this information coincides with the common use of "foreign criteria to report on what's not happening" or cannot happen in a specific context (Mitchell, 2017, p. 62). In this vein, as it is commonly used nowadays in social sciences research on public policy, this study aims to consider and highlight the importance and uniquely informed perspective of front-line workers (Nothdurfter, 2016). To accomplish this purpose, teachers were interviewed. They are somehow obliged to enact new policies and to deal with implementation conflicts on a day to day basis with their students, creating and adapting the solutions. Literature shows on numerous occasions that top-down approaches are broadly rejected and that the implementation of public policies from the bottom-up helps to decentralise power and the authoritarian nature of these approaches (Flórez Petour, 2014; Gilson et al., 2014). As previously stated, a number of Latin

American authors such as Pizarro and Gómez (2019) and Nodarse et al (2018) have identified teachers' resistance to changing their evaluation practices as reforms have failed to recognise their interactions with their specific context. However, there is also an identification of students' resistance to evaluative changes as they are used to certain practices (Martínez, 2019). In this light, this study will give due importance to teachers' and students' perspectives. Currently, there is an increasing body of research on bottom-up approaches to public policy implementation (Gilson et al., 2014; Nothdurfter, 2016; Watson, 2014). The main thrust of this new approach is to put the contextual knowledge of policy implementers to good use, seeking their active participation in adapting policies to local contexts and reality. This avoids any implementation gaps (or chasms), which are typical in top-down approaches (Gilson et al., 2014).

3.3.7 Teachers working conditions in Chile: work overload and lack of time

Chile presents concerning data as a country in the OECD reports on secondary level education, where this project has been conducted. First, secondary students' classes are organised into numerous groups and the number of students per teacher is one of the largest among the OECD countries, with an annual expenditure per student (5868 USD equivalent, ranking 36/39) that is also one of the lowest among OECD countries with available data (OECD, 2020). It is necessary, then, to mention here that resource allocation becomes a key aspect of assessment practices, as they secure "equitable access to high-quality learning opportunities" (SWAAT, 2017, p. 8). Therefore, within the school, it is expected to find large groups of students (usually 45) in a classroom with very basic material conditions. In this reality, Chilean teachers impart their lessons and assess the students' learning processes: in this reality, students are expected to learn effectively.

In terms of teachers' work, the full-time contract is 44 hours maximum per week, with 28-30 hours for teaching (lecture time) and the remaining for different non-lecture activities (Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas, CPEIP, 2019). Non-lecture time involves different tasks, such as planning activities and assessment, collaborative work, individual professional development activities, meetings with parents and students, staff meetings and activities proposed by the school's head teacher (CPEIP, 2019).

Internationally, the UK has been regarded as a nation wherein teachers have paperwork, administrative tasks, and chronic work overload which undermines class preparation (McNess et al., 2003). Early in 1994 Hargreaves identified teachers' time as being a subjectively defined

horizon of both, possibility and also a limitation. Time is still a main constraint in teachers' work. Consequently, evaluation could be impacted by 'economy principles' of reducing workload; for instance, using the classic paper-pencil tests with items that are quick to review, where the reasoning and mental processes used by students are never prioritised (Hernández-Nodarse, 2017). Therefore, the time issue is expected to influence negatively the variety of evaluations that is suggested by Mineduc in the country, given that in Chile and in other countries researchers have claimed "time constitutes a resource in multiple ways: scheduled time to learn and collaborate outside of the classroom, instructional time to implement different kind of assessments, some of which may be more time-consuming than others, and instructional time to act on data analysis; and the elapsed time over which instructional change occurs" (Young & Kim, 2019, p. 23). Time constraints and a need for time is a common code found when doing research with Chilean teachers, as the work of Torres-Olave (2022) can evidence when studying science teacher communities in Valparaíso. In her research, teachers declare that they did not have time to even think or reflect on what they do as part of their work (Torres-Olave, 2022). Martínez (2019) also declares that time is a component that has historically been a point of conflict in relation to Chilean teachers, their schedules, lecture and non-lecture time, planning, and even workplace harassment, conditioning their professional performance daily. In her research on evaluative practices, teachers declare that to diversify their evaluative instruments and design them, more time is required. It is a challenge for teachers to distribute their non-lecture time in order to plan a diversity of evaluations, at the same time grading and providing feedback (Martínez, 2019). In this line, the working conditions for teachers have generally been identified as an obstacle to introducing any kind of innovations, especially the lack of non-lecture time in teachers' contracts (Gysling, 2017) and also for the lack of collaborative work between them (Martínez, 2019). This also produces a clash with pre-service teachers' experience, as they acquire a critical vision about the traditional educational practices and develop a desire for innovation in the university, that clashes with reality when they face their professional practice in the schools, as they encounter a reality where they can only reproduce traditional evaluation practices (Gysling, 2017).

The complex reality that is described here has led to 20% of teachers leaving the profession before their fifth year of teaching due to acute work overload, low incomes, a lack of autonomy to teach, and a generally poor working environment (CIAE, 2021). After a decade, 30% of Chilean teachers leave the profession permanently (CIAE, 2021). Chile is among the countries with the highest teacher attrition rates (Gaete et al., 2017). The Chilean case is on the top of worldwide

countries with teachers leaving the profession, along with places such as the United States, with 40% and some regions of Canada with similar rates (González et al., 2020). Gaete et al (2017) also highlighted the lack of teachers' non-lecture time, as they do not have enough time for class planning and test correction. Additionally, standardised testing has also had an impact on teachers' disenchantment, leading them to leave the profession.

3.4 Topics identified in the literature review that guided the research design and methodology

The literature review presented in Chapters 2 and 3 helped to identify the main questions on SBA that were considered to design the research instruments for the study: they included the consideration of relevant sub-topics that guided the creation of the instruments for document selection and interview questions. Sub-topics were organised following their relation and suitability to answer the RQ's of the study: assessment purposes and methods (RQ1), assessment information (RQ2), and assessment fairness (RQ3).

<i>RQ's</i>	<i>Themes</i>	<i>Studies</i>
<i>RQ1: Assessment purposes and methods.</i>	- Variety of assessment methods.	Bota and Tulbure, 2015; Flitcroft and Woods, 2018; Mäkipää and Ouakrim-Soivio, 2019; Pope and al., 2009.
	- Impact of external testing on methods.	Bonner et al., 2018; Harris and Brown, 2009; James and Pedder, 2006; Tan, 2017.
	- Classification of assessment methods.	Earl, 2014; Zhao et al., 2017.
	- Multifaceted and complex assessment purposes.	Harris and Brown, 2009; James and Pedder, 2006; Mäkipää and Ouakrim-Soivio, 2019; Remesal, 2011; Zhao et al., 2017.
	- Difficulties with the formative and summative distinction; need for balance between concepts.	Black and Wiliam, 2009; Davies et al.; 2012; Earle, 2014; Flitcroft and Woods, 2018; Irving et al., 2011; Johnston, Sundergeld and Walton, 2019; Kennedy et al., 2008; Mäkipää and Ouakrim-Soivio, 2019; Remesal, 2011; Yates and Johnston, 2018.
<i>RQ2: Information provided by the assessment</i>	- Grading, grading impacting differently high and low achievers.	Cheng et. al., 2011; Harris and Brown, 2009; Sivenbring, 2019.
	- Summative and formative. - Feedback characteristics, focus, nature, mode, and timing.	Ball and Forzani, 2009; Black and Wiliam, 1998; De Luca el al., 2018; Stiggins, 2002; Zhao et al., 2017.
	- Results as determining students' future.	Sivenbring, 2019; Smyth and Banks, 2012.

(grades and feedback).	- Problems as results do not guide improvement.	De Luca et al., 2018; Pizarro and Gómez, 2019.
RQ3: Assessment fairness.	- Teaching and learning improvement.	Brown, 2011; Mäkipää and Ouakrim-Soivio, 2019.
	- Consistency, authenticity and teacher autonomy.	Green et al., 2007; Harris and Brown, 2009; Murchan, 2018; Yates and Johnston, 2018.
	- Feedback.	Tong and Adamson, 2015.
	- Design and context, categorising students.	Beets, 2012; Gardner et al., 2009.
	- Measurement, results and score pollution.	Black et al., 2010; Brookhart et al., 2016; Green et al., 2007.
	- Students' rights and no harm.	Green et al., 2017; Pope et al., 2009.
	- Teachers' values.	James and Pedder, 2006.
	- Appropriateness of teacher-student relationship.	Malaby and Ech, 2014.
	- Inequality and social justice.	Elwood, 2013; Smyth and Banks, 2012.
	- Culturally responsive assessment, cultural scripts.	Brown et al., 2022; Elwood and Murphy, 2015; Maldonado-Fuentes, 2021.

Table 1. Overview of topics from the literature review that guided the research design and methodology.

This table helped to consider and include different aspects of SBA that needed to be acknowledged when selecting documents, talking with teachers and students and afterwards, when analysing data and identifying themes for the presentation of the findings. It also helped to include the information found in the literature on SBA coming from the international context and also the Latin American and Chilean context, including in this way, main features but also difficulties and challenges.

Chapter 4: Research design and methodology

4.1 Introduction

The research aim is to better understand and describe SBA practices in one Chilean secondary school through teachers' and students' voices, with an emphasis placed on assessment purposes and methods used (RQ1), grading and feedback reflecting students' results (RQ2) and, finally, the fairness of the assessment system (RQ3). As the objective was to gain an understanding about SBA mainly focused on teachers' and students' perspectives, whose voices are often absent from educational research, a qualitative approach was selected as an appropriate method for the present research. The design included mainly semi-structured interviews with teachers and students. However, in order to identify assessment purposes and methods (RQ1) used in the school, some documents were selected for analysis. This is because in SBA, as seen in Chapters 2 and 3, assessment purposes and methods are usually linked to specific policies/guidelines that result on documents created within the school.

This chapter will describe an interpretivist philosophical approach to research design, a subsequent qualitative methodology and how methods were selected, designed, and used for data collection and data analysis made by thematic analysis. In the final sections, this chapter will describe ethical issues, a reflexive account of the research process, and positionality issues around the research process.

4.2 Philosophical Approach to Research Design

A research perspective is understood as an ideological frame that contains ontological, epistemological, and methodological assumptions that guide the research process (Denzin & Lincoln, 2008). Ontology entails beliefs about the nature of reality and epistemology entails the nature of knowledge (Merriam & Tisdell, 2016). The interpretive paradigm implies an ontology where reality is socially constructed as people experience the world within specific contexts (Prasad, 2017). Ontologically, reality in this study refers to subjective views and perspectives coming from people who share their experiences of the world. Therefore, there can be multiple perspectives about reality and interpretations of a single event (Merriam & Tisdell, 2016).

Epistemologically, this study considers that knowledge about the world is not factual but rather co-constructed by people and especially influenced by their perceived shared cultural characteristics (Creswell, 2013). In the research design, people's perspectives can be obtained not only by talking with them but also by analysing the documents created, that guide the assessment process following specific purposes, resulting in specific assessment methods used within the school.

Given the nature of the Chilean reality regarding the school system detailed in Chapter 3, it is necessary to consider a critical perspective. The school reality in the country is related to neoliberal educational policies, the inequality of the system and also to university access, power relations impacted by the high stakes of SBA, and even the complex use of Western concepts (Espinoza & González, 2015; Gysling, 2016; Kennedy et al., 2008; Londoño-Restrepo, 2015). This critical approach is not seen as incompatible with an interpretive one, as both utilise qualitative methodologies. Besides, critical research has become a broad concept that involves different orientations to research, with an aim to critique the way things are to bring about more justice as paramount (Merriam & Tisdell, 2016). Therefore, the ontology of this study considered that reality needs to be interpreted, but also transformed, a main claim of critical theory (Cannella & Lincoln, 2015). Regarding epistemology, this will be enriched by recognising that knowledge is co-constructed considering the context, structures, and also historical conditions where it takes place in order to be used as a tool for change (Merriam & Tisdell, 2016). Complementing this interpretive focus with some critical ideas will be especially analysed in the discussion section of this thesis, in light of the overall findings.

The design in the present research allows for an in-depth and nuanced understanding of people's opinions and experiences around SBA (Hennink et al., 2011). Specifically, the focus will be placed on teachers' and students' perspectives considering their specific educational contexts. This follows a worldview that considers that meanings are variegated, protean and not simply imprinted on people but impacted by the various cultural norms that wield influence upon them (Creswell, 2013). By using document analysis and semi-structured interviews, the present research also allowed the identification of key issues relating to SBA from the perspective of teachers and students, attempting to understand the meanings and interpretations given to events (Hennink et al., 2011). Those meanings can be specially used to identify and better understand in detail injustice(s) within the educational context as well as propose possible changes.

4.3 Research design

Given the lack of studies on SBA in Chile and the trends of the few made to date in the country, this research chooses to focus on depth over breadth, giving the most complete understanding of the assessment process in one school reality. Following a qualitative approach, the intention of the study was via teachers and student voices to discern a detailed picture of the complex SBA reality inside one subsidised school taken as a representative example of the Chilean reality, given that the education provision in the country is mainly concentrated in this type of schools (Acción Educar, 2021; Contreras et al., 2011).

The study design attempted to build an interpretive and critical narrative based on the richness, complexity, and detail of participants' perspectives and experiences rather than merely on statistical logic (Baker et al., 2012). This implies for the researcher to be open-minded, flexible, empathetic, and able to listen to peoples' experiences within the naturalistic school context where they live in order to understand the subjective meanings that people attach to their experiences (Denzin & Lincoln, 2008; Hennink et al., 2011). This aim was decided acknowledging the lack of studies of SBA in Chile and the need for descriptive and narrative accounts about what is going on inside the schools. Additionally, this has been complemented with evidence found in school documents that can describe how SBA is conducted within the school, such as the internal school decree and schoolbooks registers where information about assessment purposes and methods (RQ1) can be found. In this way, triangulation in this study was based on school documents and interviews analysis as a combination of methodologies about the same phenomenon, namely, SBA (Denzin, 1970). This study on SBA in the Latin American context follows that of Pizarro and Gómez (2019), who triangulated information coming from school documentation on assessment guidelines, teachers' conceptions, and real assessment practice (see Chapter 3, p.74). This present study adds to this literature base by also considering the school assessment decree and assessment methods registered by teachers in documents and teachers' conceptions of SBA. This supports a more complete understanding of SBA by combining teachers' perspectives with their real practice, at the same time that they are guided and configured by specific school policies. However, this triangulation was further expanded in order to include an analysis of the internal created assessment documents in contraposition to the Mineduc guidelines and students' perspectives on assessment. Several studies on the topic have stressed the importance of including students' voices, which are usually underrepresented in research when compared to

other stakeholders. The literature tends to be teacher centred or one sided, relegating students to a secondary position that goes against the central and active role that they have in their own learning, motivation (or otherwise), and ultimately within the assessment process (Cheng et al., 2011; Mäkipää & Ouakrim-Soivio, 2019; Remesal, 2011; Sivenbring, 2019; Tong & Adamson, 2015). Therefore, this study considered teachers and students equally with a particular interest in voicing students, who could offer fresh insights about SBA. Moreover, the research design included collecting students' perspectives on their results and their clear links to future consequences, their experience of being labelled in the process as 'good' or 'bad' students, and the overall fairness of the system.

The study comprised two phases, these phases involved a preliminary study where the suitability of different school documents to enrich the information on RQ1 on assessment purposes and methods was explored. The preliminary study also allowed to review of the interview questions designed to answer RQ1 on assessment purposes and methods, RQ2 on assessment information in terms of grades and feedback, and RQ3 on assessment fairness. This phase facilitated to improve, modify, and adjust the research design for the data collection phase. Therefore, the second phase involved the data collection process, followed by data analysis made by a thematic analysis. This will be explored in detail in the following sections of this chapter. Both methods used, document analysis and interviews, will be described and so will the sampling strategy, documents, and participants' selection.

4.4 Methods: documents analysis and interviews

This section will describe the reasons why document analysis and interviews were selected as methods of data collection for the study. The present research followed an iterative process design that involved movements and links between non-consecutive stages to promote reflection and clarification (Atkins & Wallace, 2012). In practical terms, this involved returning to and amending the research design, making changes when necessary, which, for instance, happened especially after the preliminary study (December-January 2019), in response to COVID-19 situation and also during the collecting and analysing data processes (Atkins & Wallace, 2012). This process will be described through the following sections of this chapter.

4.4.1 Documents

School and teachers' created documents were taken as representing data that pre-exists in the school context that can offer an abundance of information with which to analyse assessment practices. Therefore, they were considered in the first phase of this study, as they were created without the researchers' intervention (Bowen, 2009). In this sense, they were taken, following Scott (1990) as 'primary documents', as the material included in this phase is produced by those experiencing the school setting first hand. Documents shed light on assessment purposes and methods (RQ1), which are guided by the school's internal assessment decree and registered by teachers in the schoolbooks in their real practice. The rationale for selecting document analysis as a method was also based on the information presented in Chapter 3, where it was described how teachers in Chile, following their criteria, were free to plan their classes following the national curriculum and to create assessments accordingly (Mujica-Johnson, 2020). This chapter also detailed that each school creates its own internal assessment decree, which follows the guidelines of the national assessment decree N° 67 (Mineduc, 2018a). School internal assessment decrees are usually public documents that can be found in each school's web page. This was an interesting document to research on assessment purposes and methods (RQ1).

In terms of the assessment methods researched via schoolbooks registers, diverse studies have shown that there is no consistent pattern suggesting that assessment in one subject is more or less reliable than in another (Harlen, 2004). Therefore, this study considered several curriculum subjects for both the document analysis and subject-teacher interviews. However, following Mujica Johnson (2020), this study acknowledges that SBA in Chile is framed by a high-stakes system where teachers design, deliver instruction, and assess through these different curriculum subjects. Thus, book registers and participant teachers were selected mainly to represent the core subjects (Language, Mathematics, Social Sciences, and Natural Sciences) as they have a greater impact on yearly progression, certification, and university entrance.

4.4.2 Interviews

The use of semi-structured interviews in the data collection process was framed by the qualitative premise of understanding the meaning of phenomena, in this case, from the views of people involved in SBA procedures inside one school (Creswell & Creswell, 2018; Kvale &

Brinkmann, 2009). Semi-structured interviews involved extracting certain desired information from all the respondents, with an order that can be modified (Meriam, 1988). The interview transcriptions were the main source of evidence for the analysis of explanations in order to understand the Chilean SBA system at the secondary level. Therefore, the guiding interview questions were organised considering different purposes and topics on SBA.

4.5 Preliminary study and research design

A preliminary study was done in Santiago, Chile, between November 25, 2019 and January 3, 2020. The study was intended as allowing to explore the documents suited to answer (RQ1) and to pilot the interview questions for RQ1, 2, and 3. This was an essential step as the focus of this study is mainly teachers' and students' perspectives on SBA.

Access to the school was granted by the gatekeeper on November for the preliminary and the main study. The procedure was reviewed with the gatekeeper and direct access to enter the classrooms to explain the research project to teachers and students and seek interview volunteers for the main study was duly granted. The access to documents was agreed upon with the academic coordinator, who was in charge of facilitating the revision process and its analysis was understood as serving a preparatory role for the interview preliminary phase. The academic coordinator informed teachers via e-mail about the preliminary and main study characteristics, including the information sheet and informed consent for teachers, students, and parents/legal guardians. The researcher also explained the study to teachers who showed interest in participating. After the successful completion of this process, the researcher entered classrooms on December 2019 to find participants willing to help with the study. The research intended to represent a broad demographic in terms of achievement, this was highlighted to the students, as they all were invited to participate by considering all their voices as relevant.

4.5.1 Preliminary study documents and decision-making

During the preliminary study, the researcher was able to work physically in the school and had personal access and contact with documents, teachers, and students. This allowed the collection of information about eight schoolbooks from the secondary level (all the classes), ten teachers' planning, and the internal school decree, which is a public document available online in

the school's web page, all of them from the academic year 2019. These documents were reviewed to assess their suitability to provide an answer to RQ1. The documents' general information was registered during the preliminary study with written notes that allowed the later creation of a table considering the document type, access and conditions, information given, and notes (Appendix A). Data analysis to address RQ1 was studied using school documentary materials that allowed a general view of school assessment methods used by teachers within the school as they certainly happened. The schoolbooks allowed to register information about each class at the secondary level and the methods used for summative assessment (grades) and their frequency.

Additionally, the preliminary data found in the documents was contrasted with the information given in the preliminary interviews, keeping in mind that they could contain a radical disjunction between teachers' conceptions and their real practices (Pizarro and Gómez, 2019). This was a major and useful insight obtained during the preliminary study that allowed changes for the main study. These alterations and refinements were the result of the interplay between the theoretical considerations based on the literature and the experiences and observations lived during the preliminary study experience (Schmidt, 2004).

During the preliminary study, I wrote short comments about the school documents, taking particular notes about the mismatch between the Mineduc guidelines and the real teachers' practice. On the one hand, the internal school decree seemed to follow the Mineduc policy in terms of formative assessment promotion. On the other hand, I noticed the high dependence on traditional tests and the impact of external testing (PAES) on the assessment procedures used by teachers in the school. The schoolbooks contained a spreadsheet titled 'formative assessment'. However, it was empty in the eight books. Therefore, the monitoring of formative assessment in real practice should be done through the interview questions. Decisions made on these aspects will be further explored in the following section.

4.5.2 Interviews in preliminary study and decisions made for the main study

During the preliminary study, it became clear that the new assessment decree that was developed by the school and teachers' registers of assessment did not seem to reflect the high importance that the Mineduc new policy (2018) has come to wield in terms of formative assessment. This had an impact on how interview questions were elaborated and designed for the main study, as they adopted a free style. The aim was to understand what meanings associated

with the assessment were found, as constructed by teachers and students as they engage with the school life (Crotty, 1998). In this sense, the document analysis method served to suggest how certain topics should be covered or questioned in the interviews (Bowen, 2009). An interesting example of this practice is Hirsh's study (2020) where she studied assessment experiences on Swedish teachers and students without directly asking about assessment but rather about what constitutes a really good teacher (Hirsh, 2020). Thus, following the evidence showed in Chapter 3 related to the scarce use of the concepts in Chile and Latin America, the questions did not ask directly using the terms 'formative' or 'summative' to explore and understand how teachers talked about assessment purposes and objectives. This was decided also considering the complex division and difference between formative and summative purposes reported in the literature in Chapter 2 (Black & Wiliam, 2009; Earle, 2014; Flitcroft & Woods, 2018; Irving et al., 2011; Johnston et al., 2019; Mäkipää & Ouakrim-Soivio, 2019; Remesal, 2011; Yates & Johnston, 2018). This exclusion made it possible to explore what types of assessment purposes are associated with the term 'assessment' itself (Brown et al., 2011). However, the presence of key elements of both types of assessments was monitored through different questions, such as learning objectives and self-assessment.

Participants were invited to take part personally during the preliminary study, explaining that all voices were relevant to the research, namely, students with all types of achievement to obtain a variety of opinions. Three students and two teachers participated in the preliminary study, which was conducted before the end of the academic year and summer holidays (December 2019-January 2020). They were different in terms of age, gender, internal codes, and specific language (Kvale & Brinkmann, 2009). Participants' characteristics are detailed as follows:

<i>Category</i>	<i>Subject or class</i>	<i>Gender</i>
<i>Teacher</i>	English	Female
<i>Teacher</i>	Language	Male
<i>Student</i>	Year 11 B	Female
<i>Student</i>	Year 12 A	Female
<i>Student</i>	Year 12 B	Male

Table 2. Sample for preliminary study interviews.

The interview during the preliminary study was focused on issues about: the interviewees' (mis)understanding of questions; adaptation of the concepts, sentences, and words to the school context; the rephrasing of questions; changes in the order following the interviewees' logic and

answers and also considering if the information collected in the pilot would help to answer the research questions effectively (Hennink et al., 2011). During the preliminary study, the activities practiced involved making clear the study, creating a good atmosphere, allowing participants to choose the place for being interviewed, and also giving them room to open up whilst maintaining an independent posture (Hermanns, 2004).

After the preliminary study interviews, I recorded several audios containing general impressions and reflections about the process. The process was conducted based on the idea that it is often difficult to predict how interviewees, in this case, teachers and students, could interpret the questions included in the guide created (Hennink et al., 2011). It also helped me to learn how to ask, considering the characteristics of two different groups (teachers and students). Two different sets of questions were created following the literature review and main topics presented in Chapters 2 and 3 (pp.95-96) and used during the preliminary study (for teachers and students separately). Both interview guides were reviewed before the application with a Chilean teacher who holds a MSc in Curriculum and Assessment. The guide for the semi-structured interviews for teachers and students for the preliminary study contained questions on six items that followed the key aspects of assessment identified at that time in the literature review/conceptual framework structure: 1) assessment purposes, 2) assessment methodologies, 3) assessment methods criteria, 4) assessment feedback, 5) grade differences and 6) personal perspectives on changes, fairness and ethics of the process.

The modifications (Appendix B and D) and a literature review update allowed the creation of the final guide for the interviews (Appendix C and E). Modifications were made using a colour code, eliminating decontextualised and repetitive questions (marked in red) and adding other questions/topics (marked in green) that were mentioned by the interviewees during the conversation and in the final part of the preliminary interview, where space to make additional comments was given. For instance, the question: what do you think of the classroom assessment strategies? was eliminated because the schedule included several questions on assessment methods and purposes already, not adding new information. The themes addressed in the preliminary study interviews were also modified after several processes of literature updating for the international land Latin American and Chilean literature.

Finally, questions for the main study were organised in three topics related to the study RQs and conceptual framework: assessment purposes and methods, information, in terms of grades and feedback, and fairness.

Firstly, opening questions were used to warm up the conversation before the focused questions:

- Initial greeting and introduction of process, information, and consent (Hello, how are you? The interview process will be in this way..., do you have any queries about that?).
- General data: name and ID code, subject (for teachers), grade (for students), years in the school.

Secondly, the focused questions were created considering the topics and subtopics presented in the table at the end of Chapter 3 (pp. 95-96). For the first topic addressing RQ1, *assessment purposes and methods*, questions were created based on assessment purposes, objectives, types, and activities that students perform. Learning objectives in the questions were related to the checking of students' understanding, knowledge, and/or skills, to stimulate students to learn, to inform instructional-decision making, to establish a harmonious classroom environment, and to promote students' confidence (Zhao et al., 2017). Some ideas from the inventory listed by Brown (2006, 2011) presented in Chapter 2 were used in the description of students' abilities as a way to determine how much students have learnt, measuring students' higher-order thinking skills, and the possibility of the assessment forcing teachers to teach in a certain way against their beliefs or their teaching. Some indicative examples of the questions on this topic are:

- What objectives can you identify and tell us about from the assessment that you design/perform?
- What type of assessment do you use/perform? Do they include peer or self-assessment?
- Can you give me examples of the items and activities used/performed?

For RQ2, *assessment information*, the questions were created considering that SBA includes grades and feedback. In the country, grades are immersed within a high-stakes system as they have a direct impact on promotion, certification, and university entrance. Therefore, the interview guide included questions about the characteristics of feedback and grades, achievement, and the impact of assessment on grades, promotion, and ultimately university entrance. Some of the ideas suggested by Brown's (2006, 2011) inventories on assessment were considered when designing interview questions, specifically about providing feedback to students about their performance, making students accountable, and placing them into categories (by assigning a grade or level to student work), determining if results/information modify ongoing teaching of students and

allowing different students to get different instruction and finally, if the results can be depended on. This topic also includes ideas about the feedback's characteristics from Zhao et al (2017): focus (task-related, process-related, or person-related), nature (positive, negative, and what needs to be improved), mode (verbal, written), timing (immediate or delayed) and the teaching adaptation prompted by feedback.

Some examples of the questions on this topic are:

- What type of feedback do you give/receive? What are its characteristics?
- Do you think that assessment strategies have an impact on grades?
- Do you think that your grades reflect what you have learnt?
- What is your opinion about repeating students and the impact of school grades on university entrance?

For RQ3, *assessment fairness* as seen in the table of topics on SBA, the concept is linked with several assessment aspects. For instance, Brown's inventories (2006, 2011) considered the categories: assessment is irrelevant, bad and/or unfair to students, and an imprecise process. As seen in Chapter 2, the enormous variability in which the concept 'fairness' is used in the literature, involved questions addressing RQ3 being asked initially in general terms and then later extrapolating with further questioning. The same logic operated then for the concepts of 'formative', 'summative', and 'justice', giving participants the possibility to make links with the topics explained in this section or with new ones (and given the exploratory nature of this study). This topic included participants being asked about their concerns and possible ideas for constructive changes to the system. Questions in this item were:

- Is there any aspect of classroom assessment that concerns you? (in terms of testing, grades, frequency, learning, future possibilities).
- Would you say that the assessment system contains elements that secure its fairness?
- What changes would you make to how assessment works in the schools?

Since educational settings hold entirely different meanings to different people involved in assessment practices; the interviews were used to understand those meanings further via teachers' and students' voices. One of the main foci of the interviews was to interpret the assessment purposes that guided teachers' methods, as well as students' learning judgements about their results and intended/unintended consequences based on these dynamics, always with an overall emphasis on the fairness of the process. Both interview guides can be seen in the appendix section of this thesis (Appendix C and E).

4.6 Sampling strategy

The sampling strategy, in line with a qualitative approach, implied reflexive work and decisions that were made during the study (Emmel, 2013). It considered the limited nature of the research questions of the study, the scope of this thesis, and the time involved in doctoral studies. Furthermore, the sampling strategy considered that the intention of the study was not to generalise but to describe, understand, explain, and interpret (Maxwell, 2013) SBA in one school. Consequently, sampling in this study was impacted by the objectives of richness of information. This section will present a rationale for the sampling strategy in terms of district, school, documents, and participants selected. As this section presents the final samples for the study, they show the impact of some reflections and adjustments made during the preliminary study and due to the COVID-19 pandemic.

4.6.1 School district: San Bernardo

It is worth reiterating that, as suggested by Baird et al. (2014), the technical quality and social value of assessment need to be reviewed by taking into account the locality, setting, dynamics, and partially unique contexts in which and where they occur. This study has considered that in culturally respectful assessment, the students' context and their characteristics should always be considered. Thus, the district where the study, learning, and assessment procedures are taking place will be suitably characterised as a result.

'San Beca' is the name that local people use to refer to the place officially named San Bernardo. The district is located in the Maipo province and has a population of 301,313 in 2017, with a projection of growth of 11.13% for 2020 (Biblioteca del Congreso Nacional, BCN, 2020).

Following the district report from the National Congress Library (2020), San Bernardo is an area with a significant level of income poverty (9.42%) relative to the rest of the Metropolitan Region (5.4%) and the country (8.6%). Allied to this there are higher indices of overcrowded households (19.6%) when compared to the region (16.4%) and national figures (15.3%). There are also higher levels of multidimensional poverty (26.7%) when compared to both, the Metropolitan region (20%) and the country (20.7%). Multidimensional poverty expands the vision of income poverty to deprivation over at least five different dimensions: education, health, work and

financial security; housing and environment, and networks and social cohesion (Ministry of Social Development, Chile, 2015).

Concerning safety and security indices, San Bernardo has higher crime rates (3,000.9) per 100,000 residents, creating a greater social impact than the region (2,940.5) and the country as a whole (2,431.2). The domestic violence involves 643.6 complaints per 100,000 people, which is higher than the regional (510.6) and national figures (524.7).

As per the national panorama, in San Bernardo, the concentration of educational access is found within the subsidised private sector with 86 schools, where 39,411 students are concentrated (64.7%), in contrast to 18,144 in 41 free state schools (29.8%) and 3,375 in 6 private schools (5.5%) (BCN, 2023). 12,185 students aged 14 to 18 in the district are at the secondary level in contrast to 34,510 at the primary level aged 6 to 13 (BCN, 2023). Regarding PAES, the voluntary test used for university entrance exhibits the results of a highly segregated system: due to the changes in the test, the numbers increased from 33.6% in 2019 to 43.1% of students from the free State schools reaching 450 points, the minimum to apply to university in 2021. This lies in stark contrast to 67.7% of students reaching the minimum score in the subsidised private sector and 86,3% of the private paid sector in the district (BCN, 2023). Contrasting with national figures, students from private paid schools obtained higher scores on all the tests, followed by subsidised schools and in the last place, public schools. For instance, in Language, private schools scored an average of 748.9 points, subsidised schools 641.8, and public schools 607.7; in Mathematics private schools scored 701.4, subsidised schools 553.7 and public schools 527.6 (DEMRE; 2023).

The PLADECOS report (2011) highlighted that 3% of the population of 15 years and above, do not reach the basic instrumental learning levels in the district. This is translated into a social disadvantage of opportunities, insecurity to obtain a job, to satisfy basic needs without depending on public funds, and also increased factors of social risk and exclusion.

These contextual factors must be seen within the national reality of inequality and segregation, as described in previous sections of this chapter. However, suffice to state here, as the school is state subsidised and free, it can be assumed that most of the students in the sample come from the most vulnerable parts of San Bernardo, which is always considering that the academic segregation in the country is especially high in secondary schools (Valenzuela et al., 2014; Villalobos and Valenzuela, 2012). However, this must be always contextualised thinking that the district has, at present, a mixture of people coming from different socioeconomic backgrounds and also considering that subsidised schools concentrate students from medium-low,

medium, and medium-high socio-economical levels in the country (García-Huidobro and Bellei, 2003). This is a reality that inevitably exerts an impact on their perspectives about the evaluation made of them within the school and in their teachers.

4.6.2 The School

The school where the study was conducted is a governmental subsidised school, that was created during 2009 in order to cover the increasing educational demand in San Bernardo (Santiago), with “the objective of guaranteeing and promoting the quality of educational improvement” as explained in the school web page. This was relevant in the research design, because the only study conducted by Mineduc before the implementation of the new national assessment decree during 2018, did not include any subsidised school within the medium or medium-low socioeconomic groups (Mineduc, 2016). However, subsidised schools have the higher percentage of students’ and secondary students’ enrolment (57%) in the country (Acción Educar, 2021; BCN, 2010; Contreras et al., 2011, Mineduc, 2018b) and the district (64.7%) (BCN, 2023). The selected school is worthy of in-depth analysis as they depict the national execution of assessment within the school context. Therefore, it can be labelled as a general case, as it represents a typical case of school situation (Merkens, 2004). This also goes in the qualitative approach direction of studying people in their natural setting, identifying the impact of the context on them (Denzin and Lincoln, 2008). Thus, the school selected for data collection represents an illustrative case, as a typical medium-class subsidised school which embodies the reality and inherent contradictions between the demography of the country and its educational policy (BCN, 2010; Cabalin, 2012; García-Huidobro and Bellei, 2003; Gysling, 2016).

As usually occurs in the country, the school is one of a networks that a private society can have, with an owner labelled as “mega-funders” (“mega-sostenedor”), as they receive the top 5 higher amount of subsidies from the government in the country (Centro de Investigación Periodística, CIPER, 2014). Additionally, the school selected is well known to me as I worked there between March 2017 and September 2018, which supports the comprehension of the social and cultural environment in which the school, students, and teachers are situated. These are regarded by the literature as impacting the assessment strategies and participants’ perspectives (Brown 2011; Remesal, 2011; Yates and Johnston, 2018). Consequently, the school also represents a convenience sample, chosen according to the ease of access (Somekh and Lewin, 2005).

The students from the school come from a variety of socioeconomic backgrounds, which represent the reality of the district. A 24% is labelled as 'priority students' by the government because they are considered as socioeconomically vulnerable and disadvantaged, which attracts more funding for their education (SEP funding, described in Chapter 3, p.74). SEP students also receive free school materials, which is not common practice, as in the country parents/carers buy the students' school materials.

The students enrolled are nearly 1,100, who are distributed in pre-school education (pre-kinder and kinder for reception in the UK system), primary school (from year 1° to 8°) and secondary school (from I° to IV°, Year 9 to 12 in the UK system). Students throughout the school are distributed into two classes per level (called class 'A' and 'B'), counting 28 classes in total.

Teachers are approximately 40. They are usually classified by preschool education, primary and secondary level. Some of the teachers have been working in the school since its beginnings, but there is also some annual mobility. This mobility has different reasons, such as teachers being dismissed but also those that quit to transfer to other schools. Teachers' income is regulated in the country by a so-called educator statute where the payment for one pedagogical hour (45 minutes) of teaching is fixed: \$17.577 (£17.7) for preschool, primary and SNED teachers and \$18.495 (£18.6) for secondary school teachers (CPEIP, 2023). However, the payment is made when the teaching hour has been completed weekly during the month. Therefore, the value of the teaching hour is divided into four weeks, reaching approximately £4.65 per teaching hour.

Teachers' income ranges for a full-time contract range from £1,200 to £1,500 approximately. However, they have higher income in comparison with other schools due their Union agreement, which involves payment for lunch, transport, medical insurance, and work as guide teachers (responsible for one class).

The general culture of the school is marked by an engaged and motivated staff with low mobility, a high demand for school places, and the trust of families that have all their children and relatives in the school. In the last years, the SIMCE results have improved positioning the school in the second place of the district (for Year 4). All these elements create a very familiar environment and the people within the school are grateful for being part of the community.

4.6.3 Documents sample

As described in Chapter 2, assessment methods variety is a topic highly covered in SBA studies (Bota & Tulbure, 2015; Flitcroft & Woods, 2018; Mäkipää & Ouakrim-Soivio, 2019; Pope et al., 2009). Therefore, schoolbooks were analysed using seven core subjects and their teachers' records (Language: Spanish and English, Mathematics, Social Sciences (History), Natural Sciences Biology, Physics, and Chemistry). The textbooks were studied during December 2019 and January 2020 by reading and reviewing thoroughly each schoolbook for the eight secondary level classes and taking notes by hand, since they are documents that cannot be scanned, photocopied, or photographed, to identify the assessment methods used by teachers (RQ1). Thus, information on each subject, assessment methods and types, quantity, and comments taken were later registered into a Microsoft Word table (see Appendix A). This register also considered the use of external testing formats on the assessment methods used by teachers, as some of the registers reflected PAES/SIMCE standardised test formats (Bonner et al., 2018; Harris & Brown, 2009; James & Pedder, 2006; Tan, 2017). The impact of external testing on the country has been detailed in Chapter 3, Sections 3.3.1 and 3.3.2.

Primary documents were produced by the school without the influence of COVID-19, namely in a normal situation by those experiencing the school setting first hand (Scott, 1990). However, due to the COVID pandemic occurring during 2020, the sample of documents was increased, as the school was facing difficulties adapting teaching and assessment to online platforms at short notice. Therefore, the Ministry of Education published documents to support the assessment processes that were included in the data set as secondary documents, as well as three documents that the school produced during data collection. In any case, the documents were designed/created without the researchers' intervention (Bowen, 2009). These were included in the data as providing means of tracking changes and developments (Bowen, 2009) during the data collection process. Considering this, two sub-sets of documents were created and analysed: primary and secondary documents, as they shaped the SBA occurring in the school. The final sample for documents was configured as follows:

<i>Data</i>	<i>Type</i>	<i>Characteristics</i>	<i>Information provided to answer RQ1</i>
<i>Primary documents</i>	1. Schoolbooks registers, 2019 (8 books).	School created (6 subjects' registers, 8 classes from Y9 to Y12). Categorical and text data.	Types of assessment methods used by teachers. Presence of external testing formats used on subjects' class contents and assessment methods (identical formats).
	2. School Assessment Decree, 2020.	School created document. Text data	Guidelines for SBA: promotion of summative and formative assessment; information assessment purposes.
	3. School Assessment Decree (adjustments due to COVID pandemic, September 2020).	School-created document. Text data.	Flexibilisation of the assessment process and the high stakes (repeating students).
	4. Powerpoint presentation, Teachers Meeting 21/09/2020.	School created Assessment modification due COVID 19. Text data.	Promotion of formative assessment use (increasing understanding of what is 'formative' in the school).
	5. Powerpoint presentation, Teachers Meeting 28/09/2020.	School assessment and promotion modification for 2020 due COVID 19. Text data.	
<i>Secondary documents</i>	6. School Assessment Adjustment during COVID pandemic, June 2020.	Mineduc. Text data.	Flexibilisation of the assessment process and the high stakes (repeating students); promotion of formative assessment use (increasing understanding of what is 'formative' in the school).
	7. Assessment, grading, and promotion criteria from Y1 to Y12, August 2020.	Mineduc. Text data.	

Table 3. Overview of sample for documents (RQ1).

Primary documents were collected during the preliminary study, between December 2019 and January 2020. Secondary documents were collected between June and September 2020. This is due to teachers commenting that they were implementing formative strategies, providing three documents as evidence. All this evidence was considered relevant to understanding teachers' and students' perspectives during that academic year in particular.

To sum up, the final documentary data set contained 14 documents (Mineduc and School created) and ran into 412 pages. This wide array of documents provided relevant evidence (Bowen, 2009) to understand SBA within the school, especially in terms of assessment purposes and methods (RQ1).

4.6.4 Sample for interviews' participants

After the information about the study was given during the preliminary study, twenty-seven students and eleven teachers volunteered for the interview process, receiving an information sheet and consent forms that were returned duly signed. Teachers were selected from the main school subjects in the secondary school level: Mathematic, Language (Spanish and English), and Sciences (Social and Natural). Students were selected from all the classes at secondary level according to their willingness to participate in the research and by complying with the requirements of parent/legal guardian consent. For the main study, participants were contacted to confirm their participation and to agree on a specific date and time for an interview. The students who did not participate in the preliminary study (24) were invited to participate in the main study by email, which was registered in their consent forms signed in December 2019. The sample was completed after receiving the 10 first email confirmations, which were duly informed to the students. This process was slow and with understanding of the difficult times students were facing because of the pandemic and the changes into online education provision.

The final sample was conformed as follows:

<i>Class</i>	<i>Participant s N°</i>	<i>Gender</i>	<i>Age</i>	<i>Attaining</i>
<i>I°B, Year 9</i>	1	Male	14	Low
<i>II°A, Year 10</i>	1	Female	15	Medium
<i>III° A, Year 11</i>	3	Male Female Female	17	High Medium Medium
<i>III° B, Year 11</i>	1	Male	17	Low
<i>IV° A, Year 12</i>	2	Female Male	17 18	High Medium
<i>IV° B, Year 12</i>	2	Female Female	19 17	Low High
<i>TOTAL</i>	10	6 Female 4 Male	14 – 19 years old	3 Low, 4 medium and 3 high

Table 4: Sample for students in the school secondary level participating in the study.

All the year groups in the secondary level were represented, as the sample of interviewed students included at least one student per year group (Years 9, 10, 11, and 12), the group during 2019 had 339 students. This was considered important, due to the heterogeneous group involved in the sample, aiming to capture some of the group variability (Baker et al., 2012).

As 2 teachers participated in the preliminary study, the remaining 9 were interviewed for the main study. However, the intended sample for interviews was 10 teachers. Due to this, the documents were re-sent by e-mail to the nine remaining volunteers and all teachers at the secondary level, to complete the ten volunteers for the interviewing process. There was a swift response from one teacher that completed the sample. In this way, 12 out of 17 teachers from the school's secondary level participated in the study (2 of them previously participated in the preliminary study). The final sample was as follows:

<i>Secondary-level Teachers and Subjects</i>	<i>Interviewed for the main study, 2020.</i>	<i>Age</i>
<i>Language (Spanish)</i>	3	39, 35, 38
<i>Mathematics</i>	2	34, 31
<i>English</i>	1	32
<i>Social Sciences (History)</i>	2	37, 27
<i>Natural Sciences (Biology, Chemistry and Physics)</i>	2	29, 58
<i>Teachers total</i>	10	Age ranges between 29 and 58 years old.

Table 5. Sample for teachers' interviews at the School secondary school level.

The teachers' sample is based on features of the Chilean system of students having approximately 12 to 14 subjects at the secondary level, which in the school in question means 17 teachers in those areas. Teachers oversaw different subjects' lessons, and students came from different classes at the secondary level. It should be noted that teachers in the country usually work in several classes at the secondary level and students share them. For example, the Social Sciences, Mathematic, and Biology teachers that were interviewed gave lessons in all the classes at the secondary level (8). Therefore, all the teachers and students have common experiences with the assessment methods used, they receive grades and feedback from the same teachers, and they have an opinion about the fairness of their shared SBA experience. Besides, teachers are usually in charge of one class as the main teacher (or 'guide teacher', following the Spanish 'profesor jefe/head teacher' that is used in the country). That implies that they are in charge of parent meetings, students' interviews when issues are detected, and that have a deeper knowledge about their students' lives and learning processes than other teachers. The final interviewees' group was considered representative of the school, as the focus was placed on the secondary school level.

4.7 Reflexive account of research design

This section will describe the reflexive process that allowed several decisions on methods used for this study. Initially, it had been considered to include in the study contextual information about the school results (RQ2) as per external national tests (SIMCE and PAES) and internal (grades) measurements. I reviewed students' grades in the schoolbooks and also the public results of the school on SIMCE and PAES (PSU at that time). However, given the negative impact and perspectives associated with those assessments found in the literature on the topic, this approach was dismissed after the preliminary study (Bonner et al., 2018; Cheng et. al., 2011; Gysling, 2015, 2017; Harris and Brown, 2009; James and Pedder, 2006; Sivenbring, 2019; Tan, 2017). In that sense, the preliminary study was key to reflect on where to find quality information that could help to answer the study's research questions.

In the same line, the teachers' planning documents were also dismissed after the preliminary study, as the planning did not provide further information but the one contained within the schoolbooks. Therefore, it was decided to place the focus on the schoolbook registers that cover the secondary level (Years 9, 10, 11, and 12, with 2 classes in each level) which included information on assessment methods (RQ1) used by teachers and its frequency.

The research design was impacted by the knowledge that several authors have stressed the need to consider students' voices on research on assessment, given that they are often absent (Cheng et al., 2011; Mäkipää & Ouakrim-Soivio, 2019; Remesal, 2011; Sivenbring, 2019; Tong & Adamson, 2015). The study was designed to advance knowledge on SBA by including their perspectives, in line with a qualitative approach. This decision implied the exclusion of quantitative methodologies, as the focus was put on people's perspectives, in this case, teachers and students. This was thought of as giving a more complete overview of the classroom reality regarding SBA.

The decision to use individual semi-structured interviews as a method also considered that the only study that Mineduc conducted and considered when planning the reform of the national assessment decree, was based upon focus groups (Mineduc, 2016). Consequently, this study intended to gather in-depth information, creating a private situation where the interviewee is free to give their opinion in front of the researcher, without other persons' presence impact. It is known that the interactions between different participants may impact the full representation of the individual perspectives of each teacher/student, and they may impact confidentiality (Hennink et al., 2011).

However, methods were thought of as complementing one another: documents complementing the information given by teachers and students to have a detailed account of SBA in the real school context. Documents are usual evidence of the schools, teachers', and students' work. Moreover, if teachers' work overload described in Chapter 3 is considered, they spend a large amount of time on paperwork. Using the information that pre-exists in the school allowed to understand how certain aspects of SBA work in reality. This was complemented with teachers' and students' perspectives on SBA.

Overall, using more than one method of data collection (documents and interviews) allowed the comparison and cross-checking of the findings by seeking the convergence or divergence of different evidence within the data set that led to the same or different conclusions (Yin, 2009). This increased the reliability as one data set potentially corroborates the other one or otherwise brings in a different perspective, allowing methodological triangulation (Atkins & Wallace, 2012). However, the triangulation was also expanded to the interview transcriptions (data triangulation) as different teachers' and students' responses were compared to decide upon themes (Atkins & Wallace, 2012). This will be further explored in the following sections on data collection and analysis.

4.8 Data collection for the main study

Data collection was impacted by the COVID-19 pandemic, which involved several decisions that will be explained separately for documents data collection and interviews, in the following sections.

4.8.1 Data collection for documents (RQ1)

The academic year started in the country in March 2020, and students only attended classes for two weeks before the national suspension for the Coronavirus outbreak. Therefore, the 2020 schoolbook registers were empty during data collection and this evidence could not be collected as originally planned². This had a severe impact on the availability of relevant school assessment documents in 2020, as originally planned and therefore it was decided to utilise instead the 2019

² The assessment process was highly modified because of the pandemic. There were no data available about the possible return of student to the classrooms in the country, as America became the continent with the most confirmed cases, positioning Chile in 12th place of cases worldwide during the data collection process (WHO, World Health Organization, 2020).

schoolbook registers collected during the preliminary study. They represented the normal assessment strategies used by teachers during a regular academic year and were used to identify and understand teachers' assessment methods.

The first overview of the eight schoolbooks was made allowing for a general vision about the assessment methods used by teachers within the school. This analysis was made first by hand and later organised on Microsoft Word files describing each class in the secondary level of the school (a detailed table will be found in Chapter 6). Data was also based on the comments/notes section on the registers, which helped to register useful information about the assessment methods used by teachers in the different subjects.

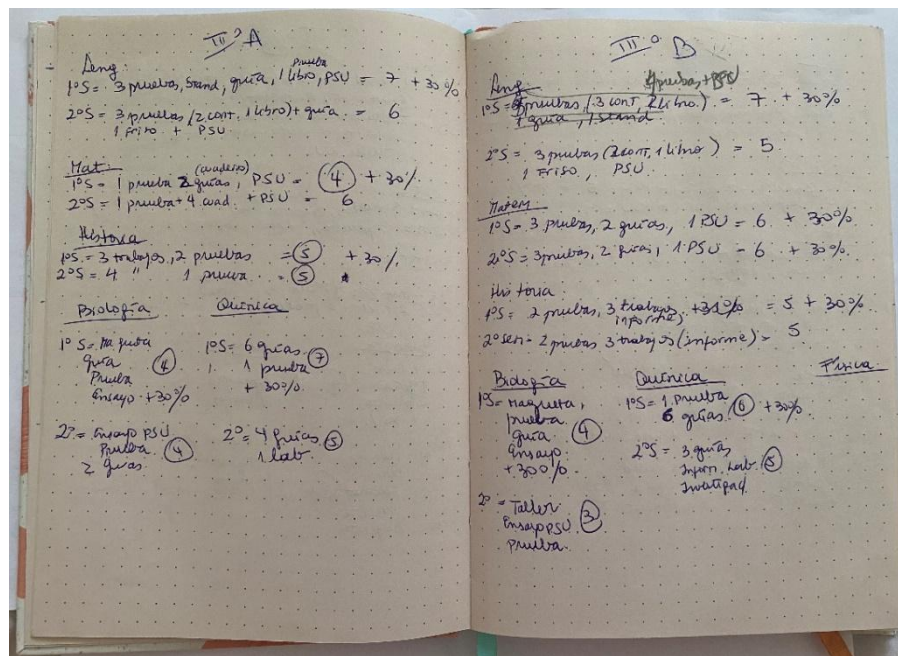


Figure 1: Image of preliminary study work notes analysing the schoolbooks.

4.8.2 Interviews for main study (RQ1, RQ2 and RQ3)

Due to the COVID-19 pandemic situation in 2020, the main interview data collection was conducted online to avoid physical contact risks (Lupton, 2020). Therefore, it has the characteristics of internet-based studies (Whale et al., 2017). This was done between July and September 2020. Interviews were conducted using Microsoft Teams as a video source of data, a process well established in qualitative research (Flick, 2009; Hammersley, 2012). The use of internet-based software provided several benefits, but also some challenges. Among the benefits, the platform helped to contact people who were geographically distant, allowing face to face

interaction (Whale et al., 2017). This was done by planning a day and time for the interviews that considered the time zone differences, as Chile was -5 hours during data collection (Lupton, 2020). During that period, teachers and students in Chile were communicating using internet-based platforms such as Zoom and Teams. Consequently, the internet-based data collection did not represent a new situation for the participants. After the initial conversation and explanation of the process, only voice recording was done, following the conditions established in the consent form previously signed by participants.

Among the challenges, the technical aspect of the data collection involved the check of webcams, audio, and stable internet connections, as they could have had an impact on the normal flow of information during the interviews (Lupton, 2020; Whale et al., 2017). The camera visual range allowed face to face interaction, listening to the participant's voices. The perceived power imbalance between the researcher and the participants was reduced by my own known status as a teacher, and because the project was explained personally to teachers and students during the preliminary study. This reduced the sense of authority possibly attributed to a PhD researcher studying in the UK and also prompted participants' trust and comfortability, as we knew each other already from the preliminary study when they showed their willingness and interest to participate.

In this way, twenty interviews were conducted with a date and time agreed with the participants via email and using Microsoft Teams. The sample for the interviews was equally balanced: ten teachers and ten students. These twenty interviewees were considered as suitable as after that number, little new information can be obtained (Green & Thorogood, 2004). That was also confirmed by informational redundancy criteria (Guba & Lincoln, 1985), as the data set showed consistent patterns on SBA among the participants.

Teachers were interviewed first, and that gave time for students to respond. It must be highlighted here that class guide-teachers support was key in order to motivate the volunteers registered to take part in the interviews during this study. Some of the interviews were shorter than others, but that was attributed to the personal opinions and characteristics of the interviewees. The length went from half hour to one hour approximately. Following a qualitative approach, the focus of the interview was the conversation content quality instead of its length. All the interviews were audio recorded, as stated in the consent form signed by the participants (and legal guardians) and stored on the University of Bristol personal OneDrive space, which is password protected. The consideration and description of the local research context was a

relevant aspect during the whole process, adding careful reflexive accounts of the research during this study (Mertens, 1998).

Interviews were transcribed immediately after being completed, allowing the identification of issues or topics that could be further explored in the upcoming interviews (Hennink et al., 2011). This was done in Spanish, and the transcriptions were member-checked by email. After the participants' transcription approval of the complete set of interviews, they were printed and labelled to be read, as can be seen in the following figure:

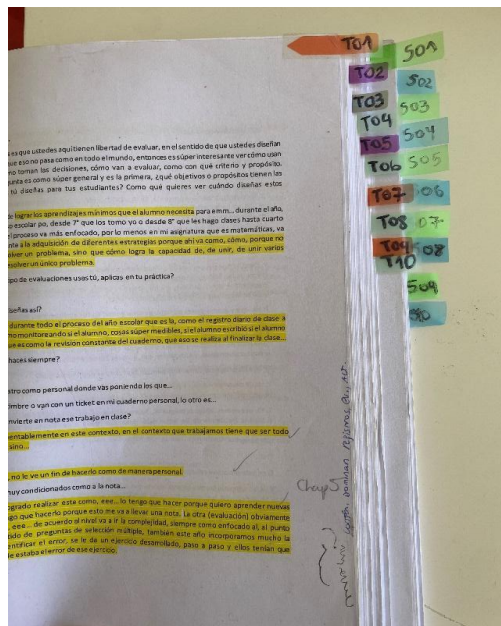


Figure 2: Printed interviews' transcripts folder.

4.9 Introduction to Data Analysis

The analysis of documents and interviews was carried out using a thematic analysis, which is a method for identifying, analysing, and reporting patterns of themes within data (Bowen, 2009; Braun & Clarke, 2006, Clarke & Braun, 2013). Despite this appearing to be essentially independent of theory and epistemology, it is compatible with constructionist paradigms since it provides a flexible and useful research tool (Braun & Clarke, 2006). This analysis method allows the organisation and description of data sets in detail and also the interpretation of various aspects of the research topic (Braun & Clarke, 2012; Clarke & Braun, 2013). It was selected under the premise that thematic analysis can provide rich and complex accounts of data (Braun & Clarke, 2006). The data set collected to address each research question was analysed thematically by means of a

flexible coding inductively to answer questions placed within three main topics: assessment purposes and methods, assessment information (grades and feedback), and assessment fairness. Even when the research questions creation process was deductive, as it was influenced by the literature review on SBA, topics contained in them were not used to classify the findings during the coding. This means that themes were identified linked to the data set itself, namely, data-driven (Braun & Clarke, 2006; Kiger & Varpio, 2020). Consequently, the data itself was used to derive the structure of the analysis. In addition, this inductive approach was suitable as it can be used where little is known about the study phenomenon (Burnard et al., 2008). That was the case for SBA in Chile, especially after the new assessment decree published in 2018.

The data analysis process involved transcription, coding, analysis, overall analysis, and written reporting for the findings chapters. Therefore, the analysis of the data set involved mainly an interpretive reading that involved me in constructing a version of what the data set represented and what could be inferred from it (Mason, 2002). The reading process highlighted that qualitative analysis is a diverse, complex, and nuanced process (Holloway & Todres, 2003). Thus, it implied an iterative process of exploring, reading, revision re-reading and constant revisiting of the results searching for understanding and the meaning of the data (Burnard et al., 2008), in their suitability to answer the RQs.

Initially, the corpus of notes and registers (from documents) and interview transcripts were separated into two different data sets: Word files, PDF files (RQ1) and interviews transcriptions (RQs 1, 2, and 3). The data set was analysed with an iterative process (Atkins & Wallace, 2012). Accordingly, the analysis included several reading processes with a first data analysis that was made by hand, by reading and re-reading the data set for both: documents and interview transcripts.

Set	Files	RQ's
<i>Set 1: documents</i>	<ul style="list-style-type: none"> - School assessment decree. - School Assessment decree adjustments. - Assessment adjustment (Mineduc). - Assessment grading and promotion (Mineduc). 	1
<i>Set 2: interviews' transcriptions</i>	- 20 files, one for each interview participant.	1, 2 and 3.

Table 6. Data set for analysis.

With an idea of relevant potential coding schemes in mind, the files containing the complete original set of documents in Spanish were imported to NVivo software, where they were codified. The reading and coding phase was characterised as a recursive process, where back and forth movements were needed. The information was later organised in one thematic map, which included the main themes for both, document analysis and interview transcripts analysis. Details on analysis for documents and interview transcriptions will be separately detailed in the following two sections.

4.9.1 Documents analysis (RQ1)

The analytic procedure entailed selecting, making sense of, and synthesizing data coming from the data set (Bowen, 2009). The documents set was divided following two purposes during the analysis: the school decree and school documents were used to understand the context (informational background), where the assessment occurred, and where participants operated (Bowen, 2009). These documents indicated the purposes under SBA was conducted (RQ1). The second intention was set following the schoolbooks' hand-written registers made by teachers, which indicated the evidence of the real practice methods, following the normative documentation. Document analysis was conducted and implied to skim, read, and interpret the documents as an iterative process (Bowen, 2009). The reading process allowed the identification of meaningful parts of texts contained within the documents set that were pertinent to answer RQ1a about assessment purposes and methods on documents. However, the iterative process also involved being able to situate this part of the study within the overall exploratory study of SBA and show connections among their different elements (see thematic map in Chapter 4). The thematic analysis allowed for a careful and focused reading of the documents to identify patterns. They appeared on the school documents and books and responded to SBA purposes and methods.

4.9.1.1 Analysis of schoolbooks files

The analysis of the schoolbooks files (Word tables) was made by the researcher based on hand-written notes taken during the preliminary study. Therefore, this analysis was not based on original documents, but rather on tables containing assessment methods (RQ1) organised by class. For this reason, the analysis was made more on words than full sentences or phrases. This is

because the assessment methods registered were usually made under short descriptions, grades-associated. These descriptions showed clear patterns of the most used assessment methods, which were classic paper and pencil tests. This information was contained under the theme ‘summative assessment methods’. However, the analysis also identified codes such as the impact of external testing (SIMCE and PAES), and the high frequency of assessment and grading in the school. Therefore, the analysis made from the schoolbooks files was later connected with information on assessment methods coming from other documents and also placed within one thematic map, integrating all the coding made during analysis.

The documents data set (RQ1) containing assessment methods as registered in Word tables created by the researcher, which were codified by hand. They were labelled as follows:

<i>Documents</i>	<i>Code name</i>
<i>8 tables (Word files) with a description of assessment methods used by class (from Year 9 to Year 12).</i>	I°A PD01
	I°B PD02
	II°A PD03
	II°B PD04
	III°A PD05
	III°B PD06
	IV°A PD07
	IV°B PD08

Table 7. Labelling of files based on the schoolbook registers.

The analysis of the assessment methods used by teachers allowed their classification in three subthemes during the analysis: high frequency of use of traditional tests, use of standardised test formats (PAES, SIMCE), and summative use of formative (formative graded work and graded rubrics). The details of these will be presented in Chapter 5.

4.9.1.2 Analysis of original documents

The original documents used for data analysis contained primary documents, created by the school, and secondary documents, created by Mineduc. Some of the documents were created by the school during the interview, transcription, and member-checking process and were mentioned by the participants as important documents to understand the assessment process. Therefore, they were included to understand SBA in the school during the academic year 2020. They were labelled as follows:

Document	Code
<i>School assessment decree</i>	PD09
<i>Decree adjustment document</i>	PD10
<i>File of school staff's meeting 21/9/2020</i>	PD11
<i>File of school staff's meeting 28/9/2020</i>	PD12
<i>Mineduc Assessment Decree adjustment due COVID 19 (Mineduc, 2020a)</i>	SD01
<i>MINEDUC grading and promotion guidelines due COVID 19 (Mineduc, 2020B9)</i>	SD02

Table 8. Original documents analysed.

Both hand analysis and NVivo software supporting the data management and handling made it possible to identify and code information about principles and purposes (RQ1) that guide SBA in the school by using PD09, PD10, PD11, PD12, SD01, and SD02. Therefore, the set was formed by the School Assessment Decree and its adjustment document due to COVID-19; two PowerPoint presentations used in the school's staff meetings, containing assessment adjustment due COVID-19, two documents published by Mineduc for assessment adjustment and grading and promotion criteria during 2020. The following figure is an example of NVivo coding:

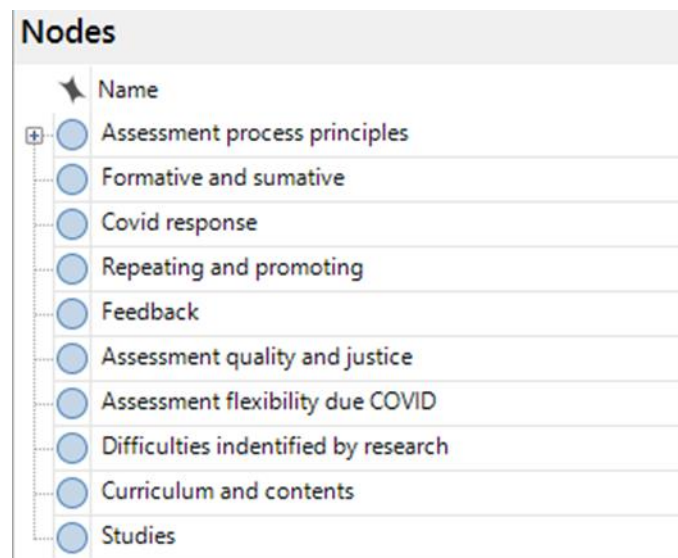


Figure 3: Example of NVivo coding for documents.

The documents set also included some examples of assessment documents used by teachers, provided during the member checking process of the interview transcripts. However,

they were not analysed but used as reference examples for assessment instruments, such as graded rubrics, self-assessments, and tests.

4.9.2 Interviews' transcripts analysis (RQ1, 2 and 3)

The transcription process involved the immersion of the researcher in data, to identify and interpret the unique participants' experiences and to understand their educational, and socio-cultural meanings (Hennink et al., 2011). Interviewees' perspectives in the form of written transcripts are textual data used frequently in qualitative research (Hennink et al., 2011). Participants were coded by number and the initial T for the teacher (T01 to T10), S for the student (S01 to S10), and R for the researcher. No further specific information has been provided about school subjects or classes, as that information can lead to the identification of the participants or their opinions.

After the interview process and transcription were finished, I reviewed and read several times the verbatim documents and identified several interesting patterns connecting with the documents (data triangulation). This was done by organising each interview question with the answers given by the ten teachers and students, separately. This allowed me to write notes in the margins of words or short phrases that helped to sum up what was being said in the text (Burnard et al., 2008).

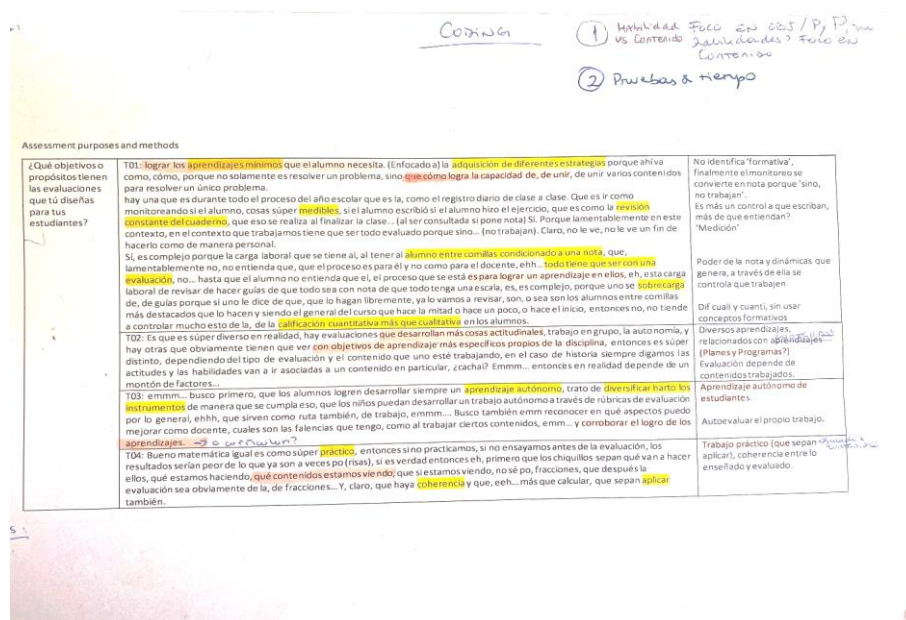


Figure 4: Example of first transcription analysis by hand identifying patterns (original version in Spanish).

This exercise, usually known as open coding, allowed to highlight words but mainly offered a summary of elements discussed in the transcript (Burnard et al., 2008). This helped to the recognition of repeated patterns and also topic changes and highlights. After this process, NVivo software was used.

The NVivo coding process was slow and separated into data from teachers and students. This was decided so that not mix their perspectives at the initial stage given that they would be brought into dialogue later. Thus, main codes were identified for the two groups separately to find commonalities and patterns between both of them. The codification made it possible to identify patterns among participants' perspectives and to organise the information into meaningful groups (Tuckett, 2005).

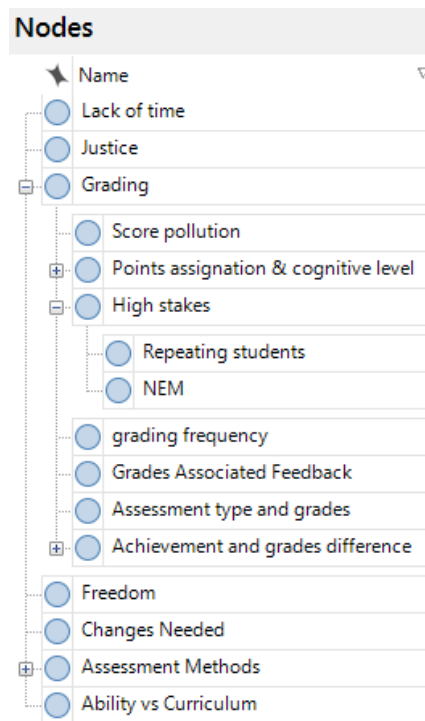


Figure 5: Example of NVivo coding for interview transcriptions. (Full list of final themes in Appendix L).

Although verbatim transcription (word for word replicas of the interview) and the analysis were done in Spanish, the relevant quotations coded by NVivo that were chosen later in the analysis to be used in the report were translated into English. Accuracy and appropriateness of the translation were checked by two colleagues-teachers and a proof-reader, both familiar with the educational language, Spanish, and also the idiomatic Chilean lexicon. The translation process was made following the idea that “a more trustworthy and accurate translation involves more than one translator” (Cormier, 2018; p. 336), and respecting meaning over form, namely, respecting

participant ideas rather than doing a strict sentence translation, which is labelled as conceptual equivalence (Cormier, 2018). The selection of particular data extracts was thought in terms of represent vivid examples of the points demonstrated in each theme (Braun & Clarke, 2006). They were also thought of as a main source of support for the narrative of the SBA within the school. This process finalised with a first scheme of themes for the entire interview data set. The codes were identified from the data and were raised from participants' perspectives (Hennink et al., 2011). These codes identified interesting features in the data in the most basic element of information, relevant to the phenomenon researched (Braun & Clarke, 2012). The validity of codes was checked by identifying if they were repeated across different participants during the study or if the code was highlighted by them as being important (Hennink et al., 2011).

4.9.3 Data Triangulation

After the first analysis of document notes and interview transcripts, they were triangulated by integrating and including information from the interviews' transcriptions into the documents containing information about the assessment methods and purposes (RQ1). This was useful to answer RQ1, as it involved assessment purposes and methods identified by both, documents and interviews. To organize the information for analysis and findings presentation, RQ1 was subdivided into two questions separating purposes and methods, as was presented in the introduction of this study (p.28):

RQ1a. What are the assessment purposes used by teachers, as indicated by the school's documents and teachers' and students' perspectives?

RQ1b. What are the assessment methods used by teachers, as indicated by school documents and teachers' and students' perspectives?

The triangulation was mainly made by the similarities of codes identified separately in the documents analysis and the interview transcriptions analysis. This allowed the confirmation of themes/codes related to assessment purposes and methods (RQ1). That means that documents and interviews were linked, generating an interconnected data set (methodological triangulation).

A table was created during data analysis, to add teachers' and students' opinions related to assessment purposes and methods. Since the information collected with this instrument was intended to address RQ1a and b, it considered the main use of summative assessment methods, the impact of external testing on methods, the classification of methods, their purposes, and also

the possible formative/summative distinction. Therefore, all the information related to these topics coming from the first section of the interview transcriptions was included in the teachers' and students' perspectives columns. The following table shows an example of triangulation:

Data from schoolbooks	Possible Theme	Backup from literature review	Participants' perspectives examples:
<i>High amount of tests during the academic year (293 tests).</i>	Summative assessment methods	Bota and Tulbure, 2015; Flitcroft and Woods, 2018; Mäkipää and Ouakrim-Soivio, 2019; Pope and al., 2009.	<i>"Because test, test, test, test... like you get stressed, like... I don't know, you are tense and all that, so if there were other methods to measure your capabilities, instead of a test, I think it would help also to increase your grades" (S10).</i>
	Lack of assessment variety, and use of standardised test formats.	Bonner et al., 2018; Harris and Brown, 2009; James and Pedder, 2006; Tan, 2017	<i>"Always you end up talking about standardised tests, so umm... accustom the guys to only one methodology of work I think that it traps them" (T10).</i>

Table 9. Example of data analysis and triangulation.

The use of NVivo software allowed the identification of nodes of information and groupings of quotations coming from the documents and transcriptions. This allowed methodological triangulation of the information for RQ1 a and b, and also to have a more developed understanding of the registers made by teachers regarding assessment methods in the schoolbooks and their purposes for the data analysis. This exercise was also relevant to analyse if data analysis and the resultant themes would be able to be contrasted with previous research on SBA for the international, Latin American and Chilean context.

4.10 Analysis of complete data set codes: themes and subthemes.

The final data set combining documents and interview transcriptions was analysed thematically (Braun & Clarke, 2006). Once the codes and sub-codes were developed, themes emerged in the analysis, thus creating a thematic map of the data and determining which aspects are captured for each theme (Braun & Clarke, 2012). Theme development involved sorting the different codes into potential themes, identifying, and collating all the relevant coded data extracts from the transcript (Braun & Clarke, 2006). In that process, codes were organised several times into themes and subthemes by integrating and combining data coming from the documents and the interviews, to organise the information for the presentation of the findings. In this stage,

themes and subthemes were linked back to the literature, which allowed the researcher to analyse their suitability to answer the RQs in the discussion section.

The themes used to create a map were constructed by analysing, comparing, and mapping by hand how they were related (Kiger & Varpio, 2020). The process involved searching for themes, such as ideas or common concepts identified during the data familiarisation phase (Braun & Clarke, 2006; Matthews & Ross, 2010). The revision of themes implied the reading of all the data extracts coded under each theme (Braun & Clarke, 2006). Several revisions of the complete data set were made to check if the themes worked or if there were any missing aspects, which led to reordering some themes.

A theme captures important elements of the data in relation to the research questions and represents a patterned response or meaning within the transcriptions and the data set (Braun & Clarke, 2012). Themes were decided in terms of their prevalence through by-hand analysis and NVivo coding across the entire data set. They were also considered for analysis when interviewees addressed them frequently. However, the themes were not decided entirely by means of quantifiable measures (Braun & Clarke, 2012). They captured important information in relation to the RQs and provided a rich description of the entire data set. The analysis followed an inductive, data-driven approach. Despite the interview questions did consider the conceptual framework of this study, the themes identified are linked to the data set itself. They were later contrasted with the conceptual framework in the discussion chapter of this study.

The themes were described, compared, and related by analysing the nodes that they implied (Bazeley, 2009). It was especially challenging to ensure that themes and sub-themes did not contain overlapping topics. The analysis process was recursive rather than linear, implying circling back in the analysis steps constantly (Kiger & Varpio, 2020). The revision was complex, as SBA is related to aspects of teaching, learning, assessment, and classroom conditions that are interconnected. The results are also impacted by external structural conditions related to the school setting such as social inequality and teachers' lack of time and increased workload. Following Braun and Clark (2006), thematic analysis seeks to theorise the sociocultural contexts and structural conditions that enable individuals' perspectives. Therefore, the analysis highlighted these structural conditions concerning the Chilean sociocultural context described in Chapter 3.

The refining of the themes was a long process where I asked myself what topics were associated between them, and which were not. As the process unfolded, the thematic map became increasingly complex (Bazeley, 2009) and included elements of the conceptual framework

of this study. For instance, when thinking about methods, I had clarity about the traditional formats used by teachers and the relevance of grading procedures. However, those elements were determined by teachers' time constraints. Therefore, several aspects of assessment methods (RQ1) were clustered under that sub-theme (teachers' lack of time and work overload). By analysing relationships among the initial NVivo nodes, I realised that teachers' lack of time was also a main motivator for the scarce feedback provided (RQ2) that was described by my twenty participants. Thus, the sub-theme was related to RQ1 and RQ2. Finally, a diagram of themes (Wellington, 2015) was developed to thoroughly answer RQ1, RQ2, and RQ3. The revision of the thematic analysis was concluded when all the relevant codes were included in the coding scheme and the themes and sub-themes structure did not require significant changes (Kiger & Varpio, 2020).

4.11 Themes and subthemes from data analysis

This section will present the main themes and subthemes resulting from the thematic analysis conducted for each set of data, in order to answer the RQ's of this study.

4.11.1 General Themes/Subthemes from Documents

Concerning RQ1, which addressed assessment purposes (1a) and methods (1b), the analysis of documents suggested that assessment purposes are highly impacted by the promotion of formative assessment use, which follows the spirit of the new assessment decree (Mineduc, 2018a). In the documents, assessment methods adopt a variety of formats that include graded formative and summative activities. However, schoolbook analysis based on a register of the methods used reveals that the assessment process is mainly focused on summative assessment through traditional tests and also tests that imitate the standardised tests used in the country, especially PAES. Additionally, the assessment methods revealed that several formative assessment methods are used, as suggested by the school assessment decree, but that these methods are always associated with grades. The methods are mainly graded formative worksheets and graded rubrics. This has been summarised in the following table:

<i>RQ</i>	<i>Topic of RQ</i>	<i>Main themes/Subthemes</i>
<i>1a</i>	Assessment purposes	Summative and formative assessment promotion, associated with grading.
<i>1b</i>	Assessment methods	Promotion of a variety of assessment methods on school decree. Summative assessment methods on schoolbook registers: <ul style="list-style-type: none"> - Low assessment variety (use of traditional tests and standardised test formats). - Summative use of formative: graded formative work and rubrics.

Table 10. Main themes and subthemes for document analysis.

4.11.2 General themes/Subthemes from Interviews' transcripts

Regarding RQ1, which explored assessment purposes (1a) and methods (1b), interview transcript analysis revealed that the assessment purposes are mainly guided by curriculum and content coverage, which seems to determine the dynamics and speed of SBA in the classroom. The assessment methods used by teachers are mainly tests, as they lack time to implement more varied assessment methods. Teachers and students also referred to the use of standardised test formats. Interestingly, assessment activities that differ from test formats are always assessed by grades.

These activities are defined by the school assessment decree as formative. However, as they are always graded, they represent a summative use of formative assessment methods. Additionally, participants show contradictory views on these assessment methods and reflections on the role of memory in learning.

The assessment information (RQ2) demonstrates complex interactions of grading processes used in the school. Participants show concerns about the high frequency of grading and assessment and also students' achievement and grade differences. Teachers and students showed concerns about the high stakes of the system (38 mentions), in terms of repeating the year and promotion. Additionally, they showed concerns about students' chances to enter university, as their grades have an impact on that possibility. Therefore, teachers attempt to help their students to increase their grades by assigning extra points, which is commonly labelled as "score pollution". As the results are mainly based on grades and contextualised in high stakes spaces, feedback is scarce. Its provision is highly impacted by the lack of teachers' time, number of students per class, and the length of the curriculum contents.

Finally, the reality of SBA in the school in their complex aspects and interactions ends up with several suggestions made by the participants about the need for changes in SBA, especially in terms of its fairness (RQ3). Teachers' and students' perspectives on fairness are highly influenced by the social inequality of the country, the grading process, and teachers' lack of time and work overload. This has been summarised in the following table:

<i>RQ</i>	<i>Topic of RQ</i>	<i>Main themes/Subthemes</i>
<i>1a</i>	Assessment purposes	Assessment purposes guided by curriculum contents coverage.
<i>1b</i>	Assessment methods	Summative assessment methods: <ul style="list-style-type: none"> - High frequency of use of traditional test (impacted by teachers' lack of time). - Use of standardised test formats (PAES, SIMCE), (impacted by social inequality). - Summative use of formative (formative graded work and graded rubrics). - Contradictory views on the use of a variety of assessment methods. - Irregular use of formative assessment: not for everyone, not always. - Reflections on the role of memory in learning.
<i>2</i>	Assessment information	Complexities of grading: <ul style="list-style-type: none"> - High frequency of grading. - High stakes: repeating/promotion, NEM, and PAES (Impacted by social inequality). - Achievement and grade differences. - Score pollution (impacted by social inequality). - Strategies for low performance. Teachers' lack of time and workload impacting on feedback provision. <ul style="list-style-type: none"> - Reflections on learning.
<i>3</i>	Assessment fairness	Need for changes towards a fairer system: <ul style="list-style-type: none"> - Social inequality and the impact of the stakes. - Unfairness of grading. - The need for a culturally responsive assessment. - Teachers' lack of time and workload.

Table 11. Main themes and subthemes for interviews' transcripts analysis.

Further description of themes, subthemes, and topics associated with the interviews' transcript analysis can be found in Chapters 5, 6, and 7 and Appendix L.

4.12 Overall themes map

Visual displays have the potential to provide useful information in qualitative research (Onwuegbuzie & Dickinson, 2008). The network created within the thematic map is intended to depict relevant relationships between themes and subthemes (Verdinelli & Scagnoli, 2013). As such, the thematic map integrated the data gathered (Bowen, 2009) by means of documents and

interview transcripts, allowing a better overview of SBA in the school. It is important to mention that this map is relevant to demonstrate cross-connections between themes, sub-themes, and concepts (Kiger & Varpio, 2020). The thematic map displays a network that is created as a collection of nodes connected by lines, showing many variables addressing results (Verdinelli & Scagnoli, 2013). Within the map, themes and sub-themes are independent but also fit within a bigger picture, working together as a whole story (Clarke & Braun, 2006). Their names were reviewed and decided to ensure that they represented the findings briefly and were adequately descriptive (Kiger & Varpio, 2020).

The thematic map was created as follows:

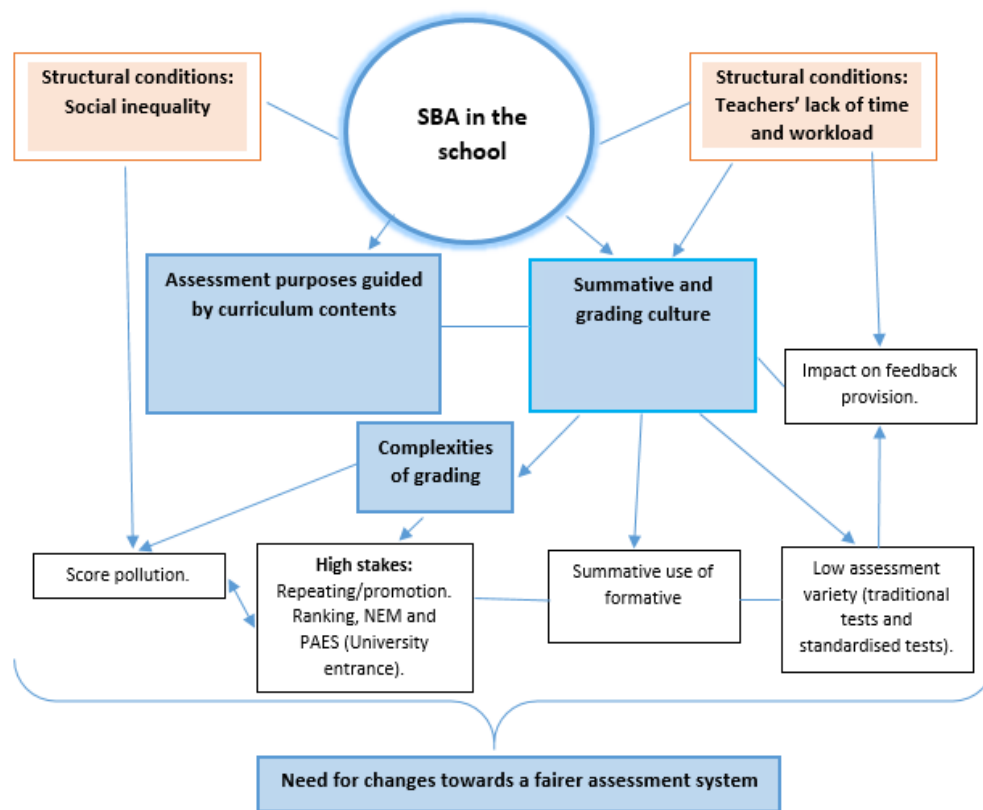


Figure 6. Overall themes and subthemes map obtained from the complete data set.

4.13 Themes and sub-themes overview

Themes were not used merely as tools to organise data but as elements that allowed me to reframe, reinterpret and connect different elements in the data set (Kiger & Varpio, 2020). As

themes were data-driven, they did not necessarily follow the interview questions structure or the conceptual framework topics (Kiger & Varpio, 2020). This allowed me to identify relevant aspects of SBA within the school, especially the way in which those aspects were interconnected. Through that process, SBA themes were clustered under six main themes and five subthemes, with several concepts associated among them. Two of the themes were labelled differently. First, structural conditions (in red), which transversally impacted the study results, and responded to a national reality described in Chapter 3. Then, non-specific school conditions included social inequality, teachers' lack of time, and work overload.

The "Summative and grading culture" theme implied that the school decree followed the same line as the Mineduc policy decree on SBA, promoting the use of formative – but graded – assessment. The document also set common rules for the use of summative assessment. As details on the use of formative are typically scarce and lack formal register systems, they were monitored by the interview questions. In any case, the focus seemed to be placed on summative strategies that are later symbolised with numerical grades and a few formative strategies, such as learning objective communication and a reduced use of co- and self-assessment.

The main themes for SBA in the school are related to RQ1b on assessment methods, with a higher frequency in the data set (629 coded mentions in the school documents and 271 in the interview transcripts). In relation to these themes, the use of standardised test formats appeared in both analyses, exhibiting low assessment variety, the use of traditional tests, and also the summative use of formative assessment. These features were later linked to the "complexities of grading" theme (RQ2), which are also connected to the high stakes of the assessment system.

Complexities of grading (181 mentions in the interviews) are presented in their relation to summative assessment impacting the assessment information (RQ2). The presence of grades were connected to the high stakes of the system, in terms of promotion and repeating of the academic year but also due to their impact on university entrance. Hence, they are interconnected to the use of standardised test formats as assessment method (RQ1b) that imitates the ones used in the country to gain access to universities through the PAES. The complexities of grades also reach the inclusion of secondary school grades to apply to university (NEM). Therefore, teachers and students reported several score pollution strategies that are used to "do justice to the students" and help them to navigate the system to increase their chances of being accepted. Thus, RQ2 and the high stakes of the system are connected to the social inequality found in the country, as a structural condition (84 mentions). Understanding that grades are not the only expected result of

the assessment procedure and that it should also contain feedback (RQ2), its provision was highly impacted by the lack of teachers' time and work overload (65 mentions). In this way, information received by students other than grades is almost inexistent.

Following in frequency, assessment purposes (RQ1a) were guided by the curriculum's content coverage as indicated by Mineduc's plans and programs (46 mentions in the interviews' transcripts). RQ1 (a and b) themes are highly impacted by the lack of teachers' time and work overload. These structural conditions also have consequences for RQ2 (assessment information). The higher frequency of these themes must be interpreted following their presence on two different data sources, i.e., documents and interviews.

Finally, the overall features of SBA that were understood in their complex connections and interactions in the thematic map are linked to the profound need for changes towards a fairer system (118 mentions). Following the tendency of the literature on the fairness of assessment presented in Chapter 2, this theme includes a variety of topics. However, participants' perspectives about the fairness of the system are linked to their main features of SBA in the country and, once again, framed within structural conditions that are imposed by the system, particularly in relation to social inequality and teachers' working conditions. The holistic view of the data set (documents and interview transcriptions) and resultant themes and subthemes obtained through an inductive analysis, was explored regarding their suitability to promote discussion, following previous literature on SBA. Indeed, the connections and interaction established between the themes as part of the thematic map provided fertile ground for discussion.

4.14 Ethical issues

This study included a broad ethical structure that incorporated UK/University of Bristol considerations (Bassey, 1999; Stuchbury & Fox, 2009). The ethics form for the pilot and the main study was approved on November, 2019 (see appendix K). This was decided in order to consider different aspects coming from different cultural realities. Some of them were respect for democracy, in terms of the freedom to express, give, and receive ideas and information, considering the responsibility with participants in terms of truth (Bassey, 1999). Truth crossed the data collection, analysis, and reporting of findings and also was related to respect for persons, recognising them as human beings entitled to dignity and privacy (Bassey, 1999). Consequently,

the ethical procedures (Appendix F and G) were guided by universal codes and principles with a special focus on trustworthiness (Bassey, 1999). They considered the time and observation dedicated to the research, the check and triangulation of the data, and the detailed accounts of the findings and conclusions. All the writing and report processes respected strictly participants' and the school's anonymity at all stages of the study. The data collection procedures and interviews were carefully conducted to avoid prejudice in labelling students (BERA, 2018). Since the data collection process involved the participation of underage students for interviews, fully informed consent from legal guardians and students was sought in all cases (Appendix J). All the procedures were done respecting teachers' and students' autonomy and right to express their views freely (Brooks et al., 2014). Information sheets (Appendix I) were given to participants and legal guardians sought to inform them regarding the process in which they were consenting to contribute (BERA, 2018). This included information about the disclosure of excerpts and the possibility for underage participants to be accompanied during the interviews (BERA, 2018).

Regarding validity, the study considered two aspects of validation. On the one hand, interview transcripts were member checked by all the participants. That was especially relevant to understand what teachers understood by formative assessment in the country. On the other hand, peer validity was sought by reviewing the data analysis and its results with teacher colleagues. Respondent validation and peer validation are highlighted as key elements for a qualitative study validation, as they can help to guard against potential bias, providing additional insights into theme development (Burnard et al., 2008).

This study followed an internet-based data collection process for the main study, due to the pandemic situation occurring in 2020. This type of data collection removed some dangers considered in every study ethics procedure (Whale et al., 2017). However, challenges were related to technical issues, the absence of physical interaction, and the power dynamics being increased for the internet situation (Whale et al., 2017). It was important to ask volunteers beforehand how they would feel about being interviewed via online platforms (Whale et al., 2017). Equally important was to let participants know about the ideal situation of them being in a quiet room for the interview, allowing the presence of a family member only for underage students that required it. Despite these efforts, some of the interviews turned out to be quite noisy and also interrupted by people talking to the participants. In terms of having an internet-safe environment, this brought some issues regarding the researcher's commitment to privacy and confidentiality, as some of the

aspects were no longer in the researcher's hands by the use of online platforms (Chiumento et al., 2018). However, all decisions were negotiated and participants were duly informed.

As the study involved participants whose first language is not English, translation excerpts were proofread, to obtain a faithful transcription of their voice (Baker et al., 2012). In the initial phase of transcription and excerpt selection, I had several conversations for data analysis with two Chilean teachers/colleagues at the University of Bristol. This stage was crucial to faithfully translate the specifically Chilean vocabulary. After the writing of the findings chapters, they were proofread by a British PhD from the University of Bristol who speaks English, Spanish, and Portuguese.

The transcripts files were labelled using the number of the interview next to the letter T for teachers and S for students (T01, T02, S01), in order to secure their anonymity.

Finally, despite the funded character of this research (ANID), there were no constraints regarding its orientation, dissemination, conduction, and results. Therefore, this study was conducted with full independence with no conflicts of interests declared (Brooks et al., 2014).

4.15 Reflexivity and positionality

The interpretive approach used in this study acknowledges that the researcher's background, positions, or emotions are an integral part of the process of producing data (Hennink et al., 2011). In this case, it is related to the school for data collection, where I worked as a teacher during March 2017 and September 2018. I decided to go into the field, not as a stranger who would interrupt the school's atmosphere. Thus, reflexivity on the emotional value of a school placed in the district where I was born and a school where I worked are present as background features of importance. Reflexivity is a concept that allows the researcher to explore and deal with the relationship between the object of research and the researcher (Brannick & Coghlan, 2007). It also allows an understanding that the aim, research questions, and results of a study are connected with a unique configuration of the researcher's personal qualities, which joins the data and the whole research process (Merriam & Tisdell, 2016). Building on my previous professional and insider role in the school familiarity facilitated the research process. For instance, the interviews followed a conversation pattern, where having previous acquaintances acted favourably. This is because even when the interviews were conducted using online platforms due to COVID19, I had the opportunity during the preliminary study process to explain the study to

teachers and students and to have a list of volunteers that were engaged with their participation, face to face, before the online data collection commenced. Luckily, this did not change, despite the stressful and unusual situation that they were living through when the interviews were conducted.

Positionality is related to the researcher as an insider and/or an outsider. The language of insiders and outsiders at the research place (in this case, the school), is usually addressed in the social sciences. Thomson and Gunter (2010) define “insider” as an individual who is familiar with the research setting in a way that an outsider could not be. An insider shares language, themes, and experiences with the participants in a study (Kim, 2012). However, insiders can lack distance and perspective on the events studied. On the other hand, an “outsider” refers to a person that can observe events with “fresh eyes” However, outsiders can misinterpret local meanings and practices. As can be seen, the researcher’s position involves both benefits and drawbacks to consider. In this study, I positioned myself between those two absolute poles and shifted from one to another. I am a teacher and I worked in the school, but at that time I was visiting the school and conducting online interviews as a researcher coming from a European university. In this sense, it is difficult to be further removed from the position of the bus driver’s daughter who was raised in a poor village in the same district. Thus, there was tension between my nativeness (Chile) and the knowledge acquired from a non-native system (England) (Smith, 2021). For that reason, my subjective experience was rather fluid, continuously shifting, and “messily blurred in particular places and times” (Thomson & Gunter, 2010, p. 26). Therefore, I adopted a perspective where my positionality was dynamic, which involved a complex interplay between my experiences and my adaptation to various situations (Barnes, 2021; Couture et al., 2012; Thomson & Gunter, 2010; Milligan, 2016; Rahman, 2023). Before the preliminary study I reflected on how to balance my multiple and shifting identities that represented who I was and who I am today. I felt a sense of loyalty to my colleagues; they are my colleagues and I had an opportunity to make claims about their/our reality. In this sense, my overriding sense was that I was not coming back as a privileged person but rather I was returning and asking for their participation as a peer and contemporary who emerged from a complex socio-economical background to be where I am today. Therefore, I was an insider and an outsider at the same time (Hult, 2014). As Couture et al. (2012) suggest rather than dichotomising my identity as a researcher by labelling it as insider or outsider I embraced my numerous and complex inherent identities. Thus, given the characteristics of this study, my reflection led me to ponder on this in terms of reducing the sense of intrusion, since the

particular reality of the district is well known to me. Moreover, I could evoke a sense of belonging (McNess et al., 2015), especially during the interviews. Furthermore, there were benefits in terms of having access to documentation, data participants, and the local context (Brannick & Coghlan, 2007; Hult, 2014). This was confirmed during my time in the school during November 2019 and January 2020 and during the main study procedures as well. My experiences with the school helped me conduct protocols and procedures in a manner that was considered socio-culturally appropriate (BERA, 2018).

However, reconciling these often-contradictory background factors that wield a significant influence on the research was a major challenge during data collection procedures. I was aware of the difficulties of a research process where shifting positionalities were present by being critically reflexive. During the interviews with both teachers and students I did not assume that I immediately understood everything but rather sought ways to elicit clarifications and ask for examples. Similarly, I did not take for granted the participants' knowledge about the assessment system in the school but rather kept asking to elaborate and clarify meanings (Cormier, 2018; Kanuha, 2000).

This study is in agreement with the idea that language impacts positionality (Cormier, 2018). Thus, my insider position posited some benefits, such as speaking the same language as my participants during data collection, which reduced my power over them as the researcher's language was not imposed (Cormier, 2018). By being a "linguistic insider" (Cormier, 2018), I was able to speak and understand the "everyday jargon" used in the Chilean context and specifically in the district where the study was conducted. This also allowed me to make follow up replies and create richer data in the interview process (Brannick & Coghlan, 2007). However, my in-between positionality regarding language was also a deep point for reflection afterward in the data analysis process, where issues on translation and interpretation emerged. This is because my original data set created in Spanish was later presented in English. Concerning translation, I faced cultural-linguistic difficulties because of certain English words having no equivalences in Spanish (i.e., School Based Assessment, assessment, fairness), an issue that has been highlighted by authors such as Cormier (2018). I acknowledged these difficulties from the beginning in the conceptual framework of this study, and also by creating two literature review chapters that could explain conceptual differences in relation to the assessment process for the English language and for the Latin American/Chilean language. They were markedly different, posing a challenge in conveying the diverse meanings of a process (assessment) that was culturally embedded in two different

languages. Thus, I adopted theoretical ideas such as “socio-culturally responsive assessment” and “cultural scripts” in order to consider and to include the impact of the culture on SBA (Brown et al, 2022; Elwood & Murphy, 2015).

Additionally, I had the responsibility to translate from Chilean jargon to English, being aware that “no one language directly matches any other” (Dalby, 2003, p. 271). Consequently, translation had an impact on data validity. Thus, I decided not to treat the translation and interpretation process as invisible; rather, I made my decisions explicit in Chapter 4 to demonstrate transparency. These decisions focused on maintaining the data set in its original language for as long as possible, to apply proofreading, and to have discussions with three bilingual academics (one British and two Chileans). The translation of the excerpts that are included in this thesis was made only after data collection and data analysis to present the findings, which is called a “later translation”. It sought to keep the participants’ voices in their original language for as long as possible, recognising the ontological importance of individuals’ first language in research (Cormier, 2014; Temple & Young, 2004).

In line with Milligan (2016), I considered how the interviewees’ perceptions of my role shaped my positionality. I think that I was seen as “one of them”, as the social structure for this zone of Santiago makes us maintain a closed posture, distrust external individuals, and observe privilege with suspicion, features which I inevitably share. I think that a great deal of the conversation process was based on their acknowledgment of my persona as someone who knows how it is like to be inside a Chilean classroom and who knows first-hand the schools’ reality. I was respected for this, and they trusted me with their honest perspectives in the process, which added trustworthiness to the study.

4.16 Findings chapter’s structure

It must be noted that the theme presentation is data-driven. Therefore, they have been organised into the findings chapters in relation to their suitability to answer the RQs in a slightly different way in which they are presented in the overall themes map. They also attend to the complexities of the interrelationships among the themes, subthemes, and topics presented. When organising the themes for the findings chapters, they were related to the study questions regardless of the frequency of quantity or data clustered under each one (Kuger & Varpio, 2020). This was especially true when, for instance, analysing assessment methods. This clearly showed the

relationship between teachers' lack of time as a structural condition that was not asked directly in the interviews. The same situation occurred for grading complexities, that is impacted by a different structural condition, that has an over-presence on this study's findings: the social inequality in the country. Both conditions have a different level of impact on each theme and subtheme, configuring how SBA is lived inside the school by their main actors: teachers and students.

Therefore, findings will be presented in three chapters, following the three research questions of this study and the main themes associated with each one. Chapter 5 will present results on assessment purposes and methods (RQ1a and b), highlighting the assessment purposes guided by curriculum contents and summative assessment methods and testing culture impacting on SBA. Chapter 6 will present information on grades and feedback (RQ2), highlighting the impact of the high stakes on the complexities of grading. Besides, it will present how teachers' lack of time, as a national structural condition, impacts the scarce feedback provision. Finally, Chapter 7 about assessment fairness (RQ3) will present the need for changes toward a fairer system as per teachers' and students' perspectives, in their relation to several aspects of the Chilean educational system and its inherent structural conditions.

Chapter 5: Assessment purposes and methods.

5.1 Introduction

This chapter, the first of 3, introduces the findings obtained from the analysis of the school documents and interview transcriptions. Specifically, this Chapter will present findings on RQ1: What are the school assessment purposes and methods used in one secondary school in Chile and what are teachers' and students' perspectives about them? This will be structured in two parts, presenting first the findings about assessment purposes (RQ1a) and assessment methods (RQ1b) using an analysis of school documents and interview transcripts. Two main themes will be drawn from this chapter: assessment purposes guided by curriculum contents and also the summative and grading culture.

5.2 Assessment purposes on the school assessment decree

The internal school assessment decree document, coded as PD09, was reviewed and updated by the school in 2019 to incorporate the new requirements made for the Ministry of Education for all the schools in the country during 2018. This new school assessment decree was published online on 2019 (n.d) and communicated to the school students and families on March, 2020. It contains information referring to the assessment process, grading, and promotion of students. The document details guidelines for the assessment, following the school's educative project, general dispositions in the form of definitions, and the assessment process, where it is described as SBA as contains summative and formative purposes, SNED assessment, grades, semester and year averages, special situations of assessment and promotion, personality reports requirements, formative and summative grades and yearly promotion (PD09). The main change between the wording of the new school assessment decree and the former one, following the Mineduc requirements, is based on changes in the rules for repeating students, inviting schools to avoid the practice, and also a reduction in the grades needed per semester, lowering them. However, decisions on both topics are school taken, and in that vein, they can be contrary to the decree spirit. Assessment remains suggested as a process containing formative and summative aspects. In the school assessment decree document, 'assessment' is understood as

“the set of activities led by education professionals which the aim to empower students to obtain and to interpret information about their learning in order to encourage decisions that promote the learning process and also give feedback on the teaching process” (PD09, p.3).

Overall, the school assessment decree seems to be a general regulative framework that allows several unspecified aspects of the assessment implementation. As such, even when the decree promotes the use of summative and formative purposes of assessment, it does not include details of their interaction, potential issues, or how to protect one from the purposes of the other (see Chapter 2, p.38).

The assessment, following the assessment school assessment decree (PD09) is understood explicitly in both, the summative and formative sense, as an intrinsic part of the teaching process. Therefore, the school assessment decree seems to follow the Ministry of Education guidelines promoting the use of both forms of assessment (Mineduc, 2018b). The formative purpose is understood as integrating teaching to monitor and support students’ learning. Hence, the assessment is defined as formative when students’ performance is obtained, interpreted, and used by professional educators in order to make decisions about the next steps within the teaching-learning process (PD09). On the other hand, summative assessment aims to certify, generally by the use of qualifications (grades), and by extension the students’ learning achievement (PD09).

The school assessment decree makes explicit the definitions and distinctions of formative and summative assessment respectively within the school:

“Formative or qualitative assessments are understood as directed towards measuring learning quality for advancement (monitoring). It will have a permanent character and will be planned based on the aims and objectives accomplished by each subject and student. The teacher will give this assessment a formative character, that will lead the students’ learning process. The accomplished and not accomplished objectives will motivate feedback” (PD09, p.4).

However, the formative aspect has been linked in the school assessment decree with the concept of ‘measurement’ and also with some activities that finally could result in a grade. In that sense, the assessment purposes detailed by the school assessment decree are each categorised under the themes ‘summative and grading culture’ respectively. Formative actions tend to be structured and formal and their main focus should be the way in which they effectively guide and

support the students' learning process, especially by the use of feedback. However, the document does not state how this formative aspect can be registered or evidenced by teachers nor does it give more detail about the feedback characteristics, procedures, or minimum requirements.

On the other hand, summative assessment refers in the document to the *“numerical representation of the objectives of learning abilities”* (PD09, p.8). The document does not include further information about how this representation can be made or the points assignment process used to generate a grade, despite being supposed to be related to the student's accomplishment of learning abilities, which are usually outlined in the Ministry of Education objectives within the national plans and study programs. Overall, the document contains basic definitions of concepts, but does not contain details about how the interaction between summative and formative purposes can be made or neither entails how one can feed into the other one. This has been stressed as relevant by international literature on SBA.

The school assessment decree (PD09) also contains relevant information about the grading process inside the school. It defines *‘grade’ or ‘qualification’ as “the learning achievement representation through an assessment process which allows a shared meaning to such learning by a number, symbol or concept to be conveyed”* (PD09, p.3). The decree establishes that students have the right to be informed about the assessment criteria and to be assessed and graded according to an objective and also transparent process (PD09). Grading can be linked, for instance, to specified scores within tests and the summative rubrics used by teachers when they use assessment formats that vary from tests. In any case, the representation is always numerical.

5.3 Assessment methods on the school assessment decree

According to the school assessment decree (PD09), the school promotes a socio-cognitive approach to education. There are no details about what this approach means in the document. However, it is suggested a participative methodology for teaching is also considered for the development of activities that students perform in. The decree states that the school promotes an active pedagogy as it *“believes in inspiring the development of creativity and the innovation of methodologies applied, by using a variety of resources that imply the incorporation of modern information technology”* (PD09, p.2). Consequently, the socio-cognitive approach seems to be implicitly related to participative, creative, innovative methodologies of teaching and activities and the use of a variety of resources.

The school assessment decree offers various examples of methods used for summative assessment:

“Possible strategies for this type of assessment, considering the diversification of learning are: projects, role play, focus groups and/or forums, monographies, bibliographical files, expositions, open questions test, structured and semi-structured tests (multiple choice, paired terms, completing information), acting, practical work, research work, exercise worksheets, learning on field, dissertation, lab work (experimental), oral interrogation, essays and others that the teacher considers pertinent for the assessment” (PD09, pp. 8-9).

It also suggests formative assessment methods:

“Worksheets with exercises, questionnaires, homework and notebook revision, essays, oral questioning, among others” (PD09, p.8).

The ‘other’ type of formative or ‘qualitative’ assessment methods are not further detailed in the document.

As can be seen, the school suggests a wide range of assessment activities both formative and summative for teachers to use with their students with specific methods associated, placing the main decision in teachers’ hands when they plan their classes and activities (as was stated in Chapter 3). The document does not establish an explicit role for students in these decisions. Even considering teachers’ training on formative and summative assessment methods, the document does not explain how the same activities suggested for formative and summative purposes, such as essays and worksheets, would be used for one or both purposes or what characteristics each process could have. This is especially relevant within a school culture where everything seems to be guided by grading. In fact, within the school assessment decree, assessment methods, formative or summative, are under the “Grading” section (PD09, pp. 8-9).

5.4 Assessment methods on schoolbooks registers (PD01 to PD08)

Following the analysis of the schoolbooks registers, it was highlighted that assessment methods were highly based on traditional tests, exhibiting a low assessment variety within a summative and grading culture. Assessment methods included graded tests based on SIMCE and PAES formats found in their data bases or online. It is also a common practice that Pre-university enterprises, that exist in the country as private paid preparation for the university test (PAES) that

students take once they finish secondary school, provide materials to the schools to enrol students in their pre-university organisations³.

Following the analysis of the schoolbooks (PD01 to PD08), Language and Mathematics are the subjects that use the bigger number of tests, consistent with the higher hours that students have in these subjects. Grades include one grade per semester based on workshop assessments. These workshops are taught by the same teachers and are based on preparation for standardised SIMCE and PAES formats tests. The score obtained in the tests and exercises is translated into a numerical grade on a scale of 1.0 to 7.0. The final grade obtained by semester counts as one grade in the main subject, Language (Spanish), and Mathematics. Apart from this compulsory workshop, two teachers also used them, one in Biology in four classes (Year 10 A and B and Year 11 A and B) and one in History and Social Sciences only in Year 10 A.

Finally, the schoolbooks evidence the summative uses of formative assessment. This is done by grading formative activities, such as in-class worksheets, and also by using points-based rubrics to grade. The themes found by the schoolbook registers analysis are detailed as follows:

Theme/Subtheme	Description	Associated codes by theme
<i>Theme: Summative and grading culture. Subtheme 1: Low assessment variety (traditional tests) and use of standardised test formats (SIMCE, PSU). (Word count 293)</i>	Test format, registered in the grades section and the class-activities registers. The test is understood here as “a form of written assignment consisting of questions and/or problems to be answered/solved by pupils individually during a limited period of time” (Hartell & Strimmel, 2017, pp. 783-784). Tests are conducted within the school space, without use of textbooks or notes, and are related to one specific content area (Hartell & Strimmel, 2017).	Contents test. Unit test. PSU test. SIMCE test. 30% Semester test.
<i>Theme: Summative and grading culture. Subtheme 2: Summative use of formative (graded formative activities). (Word count 222)</i>	Worksheets and work are understood as activities performed mostly during the class and associated with classic paper and pencil work. They differ from ‘tests’ as students can have teacher/material support (books, texts, notebooks) and can be individually or group-performed. They also can be teacher-supported during the class, providing oral feedback. This is the main difference with the ‘test’ category. They are labelled by the school assessment decree as formative, even when they are always graded. These activities can be also finished at home or given as a dated assignment. In some cases, worksheets have the same test formats (a combination of multiple choice questions and open questions and/or exercises).	Worksheet. Written report. Work. Guide. Assignment. Notebook revision. Questionnaires. Class to class exercises. Portfolio.

³ Due to the COVID-19 outbreak, SIMCE was not taken in the country during 2020. Former PSU (PTU in 2020 and PAES since 2022) had been impacted by several changes after the social uprising started in Chile in October 2019. Changes attempted to acknowledge the social segregation and inequality of the country.

<p><i>Theme: Summative and grading culture.</i></p> <p><i>Subtheme 2: Summative use of formative (graded rubrics).</i></p> <p><i>(word count 110)</i></p>	<p>As was stated in Chapter 2, rubrics set coherent evaluative criteria which help students to work identifying levels of performance quality and setting a score strategy (Brookhart, 2013; Popham, 2000). As a tool for students, rubrics help them to identify what is expected from them during the assessment, which is a formative feature. However, in the school, each level of performance quality is associated with a score that ends up in a grade that is registered in the schoolbooks, becoming high stakes (see figure 11, p.164).</p>	<p>Lab work. Video. Stand. Frieze/Model. Oral presentation. Acting. Debate. Self and peer assessment.</p>
---	--	---

Table 12. Analysis of schoolbooks registers: themes, subthemes, description, and coded words for assessment methods.

A main characteristic identified by the schoolbooks register analysis was that all activities that are associated with different assessment methods are graded. Therefore, the formative strategies linked to summative assessment methods were labelled ‘summative use of formative’, as the result is not strictly associated with a process of learning improvement. Additionally, even when the use of rubrics that specify levels of performance are considered as formative, they have points associated which end up in a grade. The following graphic allows us to identify the test, graded formative activities and graded rubrics in the different subjects across the secondary level in the school:

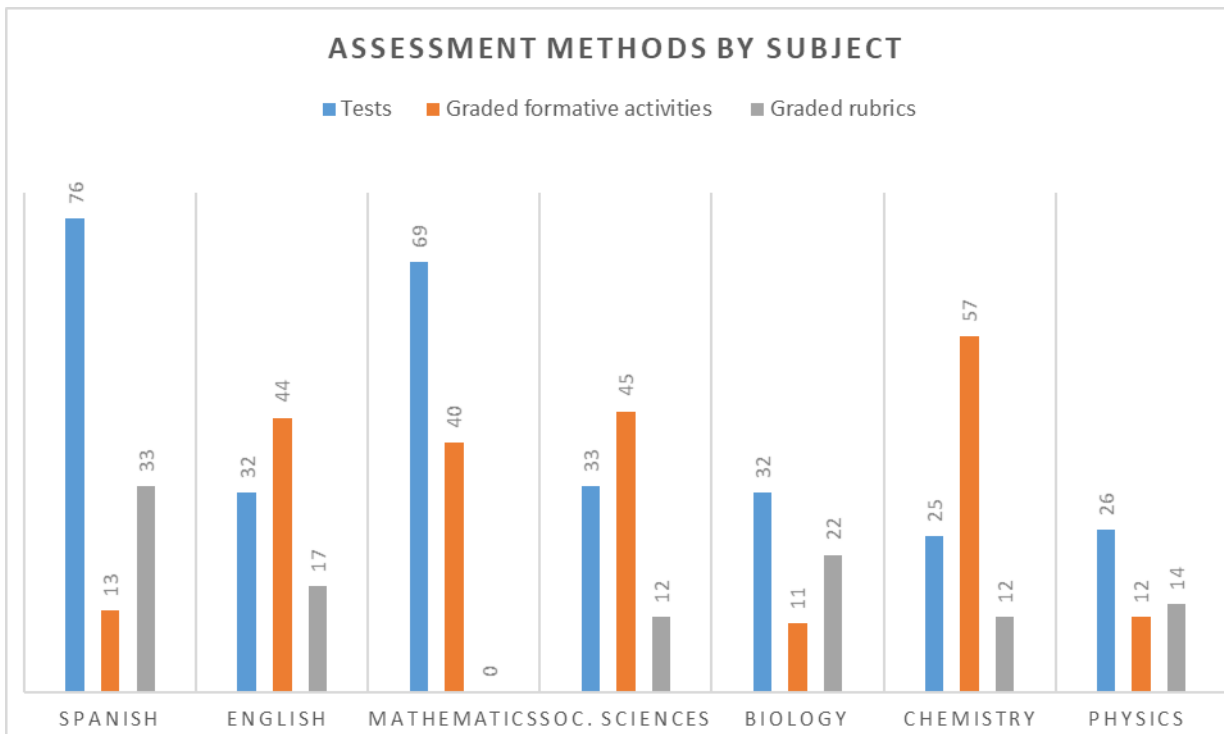


Figure 7: Detail of assessment methods used by subject and their type (following word patterns).

As can be seen in the graphic by subjects, following the three main themes, the most used assessment method registered in the schoolbooks is 'test', constituting the main significant pattern within the primary documents data set of schoolbooks registers with 293 mentions (PD01 to PD08). Summative tests are followed as pattern through a variety of graded formative activities, which get 222 mentions in the schoolbooks. This item includes a variety of activities which have a common feature being based on a range of paper and writing activities done usually in the class. They follow a similar format to tests (multiple choice and open questions) with the difference that can be made with teacher support and/or classmates and material support, such as textbooks or notebooks and more time. Therefore, the main difference between these activities and tests is that the latter are done individually and in silence, and the worksheet/work activities can be done collectively and with support and feedback. However, all these activities end up in a grade.

The third type of assessment, performance tasks was graded by using points-based rubrics that were given to the students in advance (110 mentions). This assessment method, labelled as summative use of formative, can be also done in the form of written reports or assignments. These can be related to specific subject content and include analysis or research on specific topics that are graded by assigning points to each requested section/topic to generate a grade or following graded rubrics. These activities are done usually after class time and with specific deadlines. Examples include performance in lab work, creative tasks such as video creation or models, and stand presentations in school fairs. These activities are not linked with classic paper and pencil activities and are always assessed by point-based rubrics and a grade. This type of rubric also allows a self and peer assessment in some subjects, which is a not widely used practice, as will be further explored and complemented by the interviews data. As all these activities result on a grade and usually does not imply detailed feedback, they are considered as summative.

To understand the summative and grading culture and the summative use of formative assessment methods used by teachers, which will appear also in the findings from the interviews' transcripts analysis, it is worth mentioning that all the work done is graded in two different ways. Firstly, activities can be directly graded, such as tests, worksheets, and rubrics. Secondly, activities can be graded by written activities that are registered in worksheets, in the students' notebooks, and in teachers' registers, which later become one 'accumulative grade' that can be also labelled as 'accumulative work', 'process assessment' or 'notebook revision'. In that way, this assessment method was identified in

the schoolbooks registers and can imply several different activities that would become a grade or tenths used to increase grades. This will be further explored in Chapter 6 on assessment information in terms of grades and feedback.

The assessment methods used by teachers in the school show a variety that depends on each teacher, given the freedom that they have to decide on assessment methods. It is important to say that all assessments identified and categorised as tests were teacher selected or designed/created by the teachers themselves, and in that sense, they follow multiple formats and styles. Teachers have their test databases and also are free to select materials found online. Examples in this section were provided by teachers who voluntarily offered and sent assessment instruments created by them during the interviews.

PREGUNTA 6:

El gráfico adjunto muestra las temperaturas máximas registradas durante un mes en una cierta localidad. Todos los intervalos son de la forma $[a, b[$, excepto el último que es de la forma $[c, d]$. ¿En qué intervalos se encuentran el segundo decil y el tercer cuartil, respectivamente?

- A) $[24, 28[$ y $[28, 32[$
- B) $[20, 24[$ y $[28, 32[$
- C) $[20, 24[$ y $[32, 36]$
- D) $[24, 28[$ y $[32, 36]$
- E) $[20, 24[$ y $[24, 28[$

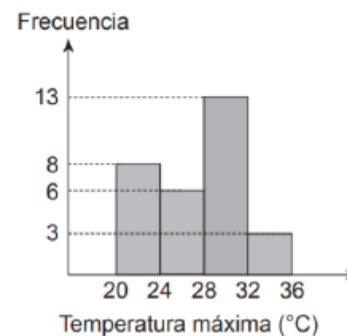


Figure 8: Example of multiple-choice test question used in Mathematics (original version in Spanish).



Figure 9: Example of multiple-choice online test questions used in Biology (original version in Spanish).

- 1) ¿Cuál restaurant muestra el mayor rango de espera?
- 2) ¿Cuál restaurant minimiza el tiempo de espera?
- 3) Si se establece un tiempo de espera de 8 minutos. ¿Cuál restaurant ofrece mejor servicio? ¿Por qué?

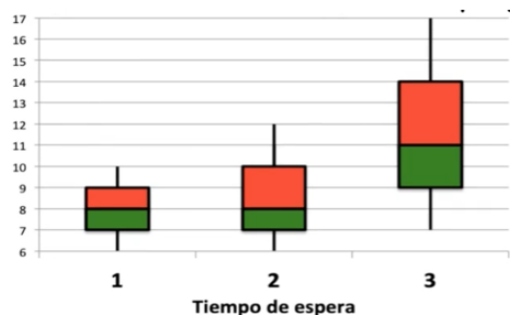


Figure 10: An example of an in-class activity in Mathematics, where students need to find values on box-plots exercises and answer three questions in their notebooks for an accumulative grade (original version in Spanish).

Finally, it is worth mentioning that the schoolbooks (PD01 to PD08) do not contain information about the formative aspects of the assessment process even when they contain a spreadsheet for the formative assessment register (that is empty in all of them). They only show evidence of the high frequency of grades and testing (summative) and assessment methods labelled by the school assessment decree (PD09) as formative, but with activities that are always graded.

The following sections will present findings on assessment purposes (RQ1a) and assessment methods (RQ1b) following the interview transcript analyses.

5.5 Assessment purposes following teachers' and students' interviews

This section will present findings on assessment purposes which, following Chilean literature on the topic and teachers' and students' voices, seemed to be guided by practices highly influenced by the coverage of wide national curriculum contents. This means that their main purpose is to advance covering the contents and topics that are suggested by the Mineduc national plans and programs for each subject. In any case, it is useful to remember that literature on the topic reports complex and various interactions about learning when thinking about assessment purposes. Curriculum contents had 46 references coded by the analysis with NVivo software in the interviews with teachers and students and also appeared consistently when they reflected on different aspects of the assessment and learning processes. When asked specifically about the assessment purpose, six teachers (out of ten) and one student (out of ten) identified directly that checking the curriculum's content coverage was the main purpose.

In the case of Language teachers, they tended to identify the main assessment purpose that students acquire specific communicational abilities, making links with the curriculum objectives or standardised tests likewise.

Assessment is done, with the main aim of checking the curriculum's content coverage mainly by testing (as could be seen in previous sections of this chapter), namely, by summative strategies to advance on the learning units suggested by Mineduc, where promoting learning is subordinated to that first mandate:

"I mean, I think that most of the teachers don't do an assessment thinking about the student, because with that we should have to create a lot of different assessment types, then we think only about objectives accomplishment with fixed learning standards [following the national curriculum]" (T09).

Teachers do mention the assessment of students' abilities as being subordinated to the curriculum's content coverage, which seems to be the main purpose of the assessment process, even when this teacher identifies that assessment purposes are diverse, finally, they depend on the content as primary guideline:

“It is like, super diverse in reality, there are assessments that develop more attitudinal things, group work, autonomy, and there are others that obviously are related with more specific learning objectives, suited to the discipline [of the subject]. So, it’s super different, depending on the assessment and the content that one is working on. In my subject, the attitudes and abilities are always related to a particular [curriculum] content” (T02).

Several teachers stated (not explicitly) that the assessment purpose is formative, in the sense of checking specific objectives for the subject that are linked to particular abilities, that are once again, subordinated to the curriculum mandate. About the attitudinal aspects, group work and autonomy were mentioned, and they are assessed by activities with graded rubrics, which means that they finally receive a grade in a summative way, contrary to formative practices. However, they help to promote learning and to check learning progression, being formative for those characteristics.

For teachers, it is difficult to separate their purposes when assessing their students from the huge pressure to cover wide curriculum contents imposed by Mineduc: *“I have like, contents, topics, contents, contents, contents... learning objectives, it’s all about contents” (T08).* Therefore, students’ abilities are subordinated and considered only for the points assignment processes that are part of the summative grading strategies used by teachers, rather than as a main objective of the educational process itself.

Both, teachers’ and students’ perspectives refer to the size of the curriculum contents as the assessment focus and purpose, questioning the meaning of the curriculum as it attempts to cover several contents, but with no success: *“I still feel that it is like, I want the student to learn everything, but at the same time [it is so much], [than they learn] nothing” (T01).* Therefore, the purpose of the assessment is marked by the over-presence of curriculum contents that are always in teachers’ minds when they think about the purposes of the assessment. It seems to be an enormous concern and also, a burden: *“We never get the curricular coverage, I mean, from four units, sometimes, if we are lucky, we get to cover two” (T01).* Or:

“For me, it’s unreachable, is unreachable in all the classes, you can’t accomplish the specificity that the curriculum requires, being so quick, like assuming that only because you cover it, they [the students] learn it... so I’m always delayed a lot, always” (T03).

The over-presence of the curriculum contents, in their extension, seems to have an impact on how teachers think about assessment purposes, as they are constantly worried about them when teaching.

For students, the assessment purpose is related to *“control that all the students reach the same level inside the school, [...] they can notice that with the tests when you can do an exercise or you can understand the content”* (S02). As can be seen, the assessment purpose is related to contents, tests, and performance on them: once again, the summative and grading culture theme. Even when one student can recognise that the assessment purpose is formative (without using precise terminology), to know if the students have comprehended, they still feel that the content and test performance are the main focus:

“The assessment objective was to know if we learnt something, I feel like it was like that, but in some cases, it wasn’t... it was to cover contents and make a test only to get everybody having good performance... it wasn’t like to be worried [teachers] about us really learning the contents” (S03).

Therefore, the double purpose is present: to know if students are learning, which is formative, but in another sense, to cover contents and to test, which is summative, and that was the main focus for this student, instead of learning checking.

With a strong focus on summative aspects, it is not surprising that four students stated that the assessment purpose was to measure students to rank them, and three of them specifically mentioned ‘intelligence measurement’ instead of learning as an assessment purpose. Students can identify that they are not assessed in their abilities:

“At the end, they don’t assess the ability, it’s what the student learnt... I don’t know if I explain myself... umm, like, maybe the student knows what the content is but maybe he/she doesn’t know how to develop that very well, because it’s not his/her ability” (S08).

Curriculum contents first, instead of students’ learning or abilities, that go in a secondary place, seems to dictate the dynamic and set the times inside the classroom, more than learning itself and in that way, more than formative purposes: there is no time to come back, we need to advance to cover the curriculum.

“But also, I feel that I don’t agree with all that we are asked to cover, with all the programs [curriculum contents programs], with all the contents because it is too

much, is too much content, and scarce time to do the activities that one would like to do” (T04).

Even when teachers identify learning as an assessment purpose, that is usually subordinated to the contents, impacting the formative aspect that learning could have in the school. Time constraints are also highlighted by this teacher. The programs containing the curriculum contents created by Mineduc become, in this way, a main idea, beyond students’ learning itself, even when it is based on them. Thus, students’ learning or abilities take a secondary place, especially considering that testing is a common and predominant practice within the classroom, as was explained in previous sections.

Time constraints have been identified in this research as a main theme by teachers and students, however, they are imposed by historical national conditions in which teachers work, namely, external. This is also highlighted by students, once again giving priority to contents coverage: *“Ummm... no [there is no time to reflect on learning], because truth is that the class is too short, so the contents are covered, then an activity and that’s all” (S03).*

It is also relevant to highlight that teachers barely use the concept “formative”, as could be expected from them when asking about the assessment purposes and especially when considering the tone of the new national assessment decree. “Summative” or “formative” concepts seem to be not appropriated by teachers and students, who simply talk about “assessment” and what the process entails, especially in terms of curriculum contents checked by paper and pencil tests. This is in line with the style of some Latin American and Chilean literature presented in Chapter 3.

It was interesting to listen to two teachers talking about the assessment purpose as to check with their students if *“they accomplished the minimum” (T01)* or the *“be able to see if the students learnt the basic contents that they need to know in that specific [learning] unit” (T02)*, opposite to one student identifying as assessment purpose *“to get the maximum of the students” (S02)*, revealing very different expectations. In any case, curriculum contents are highly relevant.

Finally, prompted by a school culture where assessment seems to be guided by grading aims based on the curriculum coverage, teachers and students were invited to reflect on what learning was for them. If it is true that learning was greatly associated with a summative culture, students reflected mainly in terms of the persistence in time that learning should have. In that sense, they considered school instances where they felt they had learnt, but with a learning that faded away in time:

“To learn is to try to... to know and learn like something new, but like, it’s like that doesn’t always endure... because with the time one forget, for instance, some assignments like [subject name], the content is like I did learn it, but likewise it was forgotten with the new contents that we covered and because we didn’t come back to that content” (S06).

As can be seen, students learning is directly related to the endurance that it should have in time and students identify that some learning fades away for different reasons, such as the sensation they have to keep advancing to cover new topics involved in each subject’s curriculum. The impact and difficulties surrounding the national curriculum in Chile have been detailed in Chapter 3.

5.5.1 The role of memory, writing, and notebooks to accomplish the assessment purposes through students’ voices

“In the school, they teach us to memorise, but they don’t teach us about love”, says a famous song from the Argentinian group “The Fabulous Cadillacs”. It is not a surprise that, in a school context where assessment purposes are guided by contents, the memory has an important role to manage contents that are assessed mainly by tests. The previous section has highlighted some elements that are key for students when they are assessed: intelligence measurement. In this light, for some students, the assessment objective is reduced to checking what they have memorised instead of their real learning. Seven students make references to the memory role in their learning and assessment. For one of these students, after the test, all is forgotten: “At least for me, I was really focused until the test, after that, all was forgotten” (S02), “it’s worthless because later everything is forgotten and you are left as a blank page” (S03) and “the system assesses only that which you memorise, truly, only what you can retain” (S08).

Students also reflect on the emotional aspects impacting the recovery of what has been memorised:

“You learn, maybe a text by memory and then the test comes. I forgot it because I got nervous. A week later, maybe I can remember something, and maybe one month after that I don’t remember... and the PSU [current PAES] comes and the same text will appear but I can’t remember and I would need to study to forget afterwards, and in the University is not going to be useful” (S08).

The recovery of information is also linked to the usefulness of the contents. Several of the contents that students need to memorise and study are written by themselves in their notebooks, along with exercises, questionnaires, and class activities. However, notebooks are not only important to study before a test, but they are also the core of several graded formative assessment activities, such as accumulative grades. Therefore, notebooks are at the core of the assessment process. The revision and grading act ensuring that students comply with specific subject-related activities, conducted especially during the class. This includes grade students only because they register the class contents in their notebooks. However, students critically comment on this form of being graded:

“I think that he [the teacher] needed to produce a grade because there was a moment where we were way behind with the subject contents [curriculum] and he needed to register grades in the schoolbook.... But it wasn't the form... [...] I think that he checked the titles [registered in the notebook] and some exercises, if an exercise was missing, he discounted one point... a lot of classmates failed the grade... that assessment does not reflect anything! I mean, it can't be possible that I am assessed by something that I wrote or not... what happens if I know something and that is the reason why I didn't take notes of that?” (S01).

This practice raises questions about how points are assigned to have a grade, and if rubrics indicating assessment criteria are provided. Students indicated that some teachers used rubrics and others did not, they simply checked the notebook sheets and assigned a grade. As can be seen, it prompted students' concerns about what that grade could reflect about learning. Additionally, it prompts questions about the assessment methods used as being traditional but also promoting shallow learning: learning to later, simply forget.

5.6 Assessment methods in interviews

Tests are usually featured as paper and pencil multiple-choice tests that are teacher-created to assess the Mineduc curriculum contents/units covered during the academic year and the abilities expected from the learning process. All these instruments must be sent to the academic coordinator, who oversees reviewing all the assessment instruments used and ordering photocopies for every student in each class. The assessment instruments creation process was

regarded as time-consuming by teachers, therefore teachers chose tests as they are easy and quick formats:

“I also use multiple choice tests and I make a development item [open questions], but it is not the most common thing because... because... being super honest, to facilitate myself the revision and marking... emmm... I try to not put so many open questions [where students must write], but when I do that it is to ask very specific things... Because the truth is that in time, you learn as well... when I started being a teacher, I used a lot of open questions, and then, on reflection, that is counterproductive, you got stuck with the assessments, reviewing them a lot of time, so it’s really hard” (T03).

As can be seen, the use of fixed response test formats (e.g. multiple choice) as an assessment method goes beyond pedagogical reasons. Teachers seem to decide based on their time constraints and workload. Students also recognise this as impacting the extended use of tests: *“It’s quicker for the teacher and it’s easier, it’s easier to have one test with all the answers and only mark if it is right or wrong” (S01).* However, test preparation is also a time-consuming task.

Following the specific needs of each subject area, the tests created and designed by teachers can consist of multiple-choice questions, such as in Mathematics and Chemistry, or a combination of closed and open questions where students develop their answers. This pattern is especially used in humanist subjects such as Language and History and Social Sciences

All the subjects include a semester final test (school decided and designed by teachers) that results in a grade that accounts for 30% of the final students’ average on the subject (due to the inclusion of self-assessment during 2020, this was lowered to 25%, the assessment is still used in 2023 as stated in the school’s assessment decree). These tests are based on multiple choice paper and pencil tests that assess the subject contents covered in one semester. As these tests are taken at the end of the semester, they require that the previous assessment made during the semester is graded and represented as an average grade that will be counted as 70% of the final grade in 2019 (that was modified in 2020 due the COVID19 pandemic). Thus, this implies several assessments and work to close the final grade before the semester test: *“is like... in the last time they bombarded us with tests and works, it was a lot, yeah... so we felt under a huge pressure to attempt the semester tests” (S02).* This semester's test was also regarded as placing increased

pressure on the students for the impact of the resultant grade (30%) on the semester average and also their frequency:

“Where I put a lot of effort in, was on the semester tests, that were the final ones, because if I had a bad performance, my whole average went down... [because of their cancellation] ... I was so happy, really happy, because they are scheduled one after another, so you end up studying for one, you have the test, and then you have to study for the other one, and you take another test, so it’s horrible...” (S03).

The frequency of tests impacted the students and the relevance of having an opportunity to increase the numerical averages obtained, highlighting the summative culture, appeared once again:

“Because test, test, test, test... like you get stressed, like... I don’t know, you are tense and all that, so if there were other methods to measure your capabilities, instead of a test, I think it would help also to increase your grades” (S10).

In this way, students seem to be acutely aware of the need for using other assessment methods related to their capabilities, although within this summative and grading culture, always linked to measurement. Perspectives on the use of a variety of assessment methods will be explored further in the following section.

5.6.1 Graded formative work as an assessment method

Student participants exhibited different opinions about the use of assessment methods other than tests. On the one hand, the low prevalence of these assessments compared to tests used make students feel that there is no space for showing wider abilities: *“I think that [the assessment] it’s very general, like everyone must adapt to the same [tests], if you fail, it’s your problem... there is scarce variety” (S03).* On the other hand, they conform to these well-known formats because this is what ‘all schools do’:

“Ummm... in the end, it is what is considered as appropriate, because in all schools they do that, every school is like that, multiple-choice tests. I think that it is not the most appropriate because you must inquire into all the abilities that one student has. So, they should make assessments that help to develop all those abilities, but if you take the matter in general, it’s the correct way, because in the end, all schools do the same” (S03).

This adherence to norms is reinforced by another student who says that the test system is fine only because *“it is what has been done forever”* (S05). As students are used to test formats, they consider it as a comfortable method, but they also are capable of valuing assessment opportunities that differ from tests or worksheets. However, being used to the application of the tests, they reflect on the abilities required for different assessment methods, impacting some students preferring the test formats:

“I do like that they make us write reports because you can give your opinion and you can express yourself, as you truly are... I think that for some of the students, it is harder though, I like to write, but that isn't the same for everyone, most of them can't develop a written idea, so they prefer the multiple-choice format, which is the most accessible to everyone” (S03).

Differing from the difficulty of these assessment methods highlighted by student 03, some students value that these assessments help them to get better grades: *“[the assignment] It's better than the test because they help you to increase your grade, they help in your average”* (S09). However, what is interesting within these different views about the variety of assessment methods is how students' abilities, which are not usually considered by teachers and students when thinking about assessment purposes (the focus is the curriculum content), appear as a relevant factor when they talk about formats other than tests and seem to be required when working on non-traditional paper and pencil tests or worksheets formats. In any case, grades are always present, highlighting the summative culture and the grading theme once again.

Worksheets with multiple-choice were used in subjects such as Mathematics, Chemistry, and Physics. In any case, students value this type of graded formative assessment: *“I like them because it helps me a lot, and make me understand better because in the worksheet I can ask if I don't understand something”* (S08). This excerpt highlights that what students value is that they receive formative feedback and help to learn.

Teachers also reflect on the importance of variety to prepare students for different types of assessments:

“One always ends up talking about standardised tests, so umm... accustom the guys to only one work methodology I think that it traps them... now, umm, the variety, I insist, always is going to have them prepared to assume any assessment” (T10).

However, for some teachers, there is no space for these activities as they must focus on the cognitive aspects of learning:

“My focus is on the other type of assessment [written], but I try to... to incorporate only to increase the variety and not to lower like... the achievement level, moreover it is like I don’t want to focus that much on the psychomotricity stuff [when doing activities other than tests] but on students managing the cognitive aspects” (T07).

Once again, this talks about the freedom that teachers must decide how to assess following their criteria about what is more important and the level that their students must have. Students also recognise that whatever form the assessment takes, it is teacher dependent and involves the trust of the teacher with each class or student: *“[the assessment] depends on each teacher, depending also on the trust level that the teacher has with the class or with each student” (S07).* Therefore, each teacher decides on assessment methods following personal perspectives. If it is true that they know that variety is important, they also identify that cognitive aspects are more relevant than activities than differ from tests as they help to keep certain achievement levels inside the classroom.

5.6.2 Use of standardised test formats as an assessment method in interviews

As was explained in some of the earlier sections of this chapter on assessment methods following the schoolbooks register analysis, teachers use SIMCE and PSU standardised test formats to grade their students in the school. Even when the format of the SIMCE and PSU workshops and their orientation towards standardised tests are decided by the school, teachers feel that they must be aligned with these formats to help their students to succeed in the national external standardised tests. This is especially true for the tests that impact on their university entrance possibilities:

“At the end of each Unit, there is a more standardised test, with multiple choice and an open questions section, that has as its aim, to identify if the students learn the basic contents that they need to know in respect of that Unit, and emmm... obviously relating to that in Chile, for instance, the university entrance is made up based on standardised tests, or that the SIMCE is going to measure the quality of education with standardised tests... [...] So you must content yourself with that, in fact, my tests always contain PSU questions, since immemorial times” (T02).

Therefore, teachers use these formats considering that their students will be exposed to them inevitably and their preparedness will impact their chances to enter higher education. It

must be reminded (as was explained in Chapter 3) that the Chilean educational system considers 'NEM'⁴ which is obtained with the numerical grade average during secondary school as part of the score that allows a student to go to the university. The score also considered a 'Ranking' item, which is a score obtained from the ranked position of the student among the classmates of the yearly cohort, following their grades. As the impact of standardised tests on teachers and students is in the data analysis linked to the complexity of these scores based on the school grades within the high stakes system, this will be further explored in Chapter 6 to avoid repetition as this chapter will be based on the findings of assessment information in terms of grades and feedback (p.185).

5.7 Irregular use of formative assessment: not for everyone, not always.

Considering that the school documents stated clearly that formative assessment methods needed to be used and evidenced but, contrary to this a strongly summative assessment style where formative work is treated summatively and graded, some of the specific characteristics of this formative assessment were investigated through the interview questions, as suggested by the extant literature. Apart from these questions, participants did not exhibit an extended use of the words 'summative' or 'formative'. They simply talked about assessment. The assessment purpose is, as the school documents and interviews analysed show, summative, and the few explicit mentions of 'formative' made by teachers (only 2) are finally also related to grading purposes as they associated with the assessment process made class by class. An example of this is what they do with the accumulative grades which will be explained in a separate section (6.7.2).

Document analysis did not provide further information about formative assessment in the practice, as the formative was always linked to grades and therefore, summative purposes. Thus, the interview questions asked about the use of peer and self-assessment, learning objectives sharing, in-class conversation about performance, achievement differences, and strategies used for low performance (feedback use will be analysed in RQ2, Chapter 6).

During 2019, as can be noticed in the schoolbooks (PD01 to PD08), some teachers used graded peer and self-assessment in some subjects such as Language (Spanish) and Biology. Even when this assessment practice was used only by two teachers in the subjects analysed in this

⁴ Despite the conditions of online learning during COVID-19 in 2020, 'NEM' was still considered that year as part of the University application scores and has not been changed. It was also considered in the 2023 process.

study, not all their classes were assessed in that way. In Language it was done only in Year 9 B and Year 10 (A and B) and in Biology, it was done only in Year 9 (A and B) and Year 12 (A and B). They can raise questions about the specific characteristics of that classes as allowing the strategy.

Following the interviews, once again time constraints seem to impact the use of self and peer assessment: *“I can’t use them all the time for, you know, time... you need to stop after the assessment, but when I can, yes, it’s done”* (T10). As was stated before, these opportunities are also linked with the summative use of rubrics for performance activities. Considering the high stakes of the system, they seem to be an invitation for good grades and an opportunity to increase numerical averages. One teacher expressed her concerns about *“the lack of appropriation of that, that assessment exercise”* (T05) as the students seem to have difficulties in objectively assessing their work:

“The truth is that I feel two different things: the first one is that [self-assessment] depends a lot on the students’ personalities. I mean, if there is a guy with high self-esteem and lack of reflective skills, at any time he is going to positively assess what he has done... there are groups where several people with huge egos are reunited and it’s difficult for them to recognise their mistakes [...] In other cases, some students are really self-destructive with themselves, they feel that they did it all wrong and that they don’t match with the grade assigned” (T05).

Teachers’ perspectives evidenced that self and peer assessment are more negative in the beginning. However, in time, as students get used to the strategy they can be more reflective and honest, making the space more meaningful for everyone. Students also value that space: *“It’s good because you can recognise your mistakes and not necessarily the teacher has to be the one saying that, you can do it by yourself”* (S07). Self and peer assessment are not perceived as common practices and the first quick answer to the question about its use is ‘no’. The process always ends up in a grade and, as was stated earlier, they were implemented only by two teachers in 2019. Eight teachers did not use this strategy at all and even when six students that were interviewed were from the classes that had self and peer assessment, only two of them could remember the activity, and the other four stated that they did not have the activity at all: *“Not that I remember”* (S02) or *“No, I believe we didn’t have”* (S06).

Some teachers used the dialogue as a formative strategy during the academic year more in the form of informal conversation about students’ learning (T01, T02, T07, T08, T10): *“We do a lot of non-formal conversation, about their learning”* (T01), *“A lot of conversation, mostly with the*

higher classes" (T02). They do not label the strategy as formative even when it is, is a practice that they simply do. This represents not only a relevant aspect of formative assessment practices but also evidence of feedback provision strategies.

Learning objectives are shared, once again, about the curriculum contents of the subject, mostly at the beginning of the academic year or in a learning unit (T01, T02, T03, T04, T05, T10): *"I communicate them as we are changing objectives and Units"* (T03). Only one teacher stated that the practice occurs in each class. The impact of the curriculum contents is also reflected in the learning objectives shared with the students verbally:

"Besides that on my subject is like ability, content, methodology, it's already a huge thing, so the learning indicators are easier, and yes, at the beginning of the Unit I always say: 'guys, look, this content is a bit more complex, this content is easier going... we are going to make links here, we are going to do this'... Always trying to create a context, so the students know where we are going, what I expect from them" (T05).

It must be noted here that this practice is mainly described as something that teachers 'try to do'. Teachers also write the learning objective on the blackboard in each class.

Related to the analysis of the classes and learning, this is impacted again by the learning objectives to demand on time and capacity. Therefore, teachers just keep going with teaching to cover the curriculum contents, without deep reflection about the process due to the lack of time: *"There are a lot of times when the teacher rushes to give the lesson, a reason why they can't go back to help students that were left without the learning"* (S07). The same idea was given in another interview: *"It depends on the subject, because there are contents that like, well, contents are covered automatically to all the class, so it's like there is no time to reflect if you learnt or not, umm... if someone needs help"* (S10).

However, despite teachers' perspectives on conversations about learning, students feel that the class space is barely used for reflection on learning, also impacted by the curriculum contents and time constraints: *"Truth be told, the class is like too short, so they [teachers] cover the contents, then an activity and that's all"* (S03) and *"There is no space to say 'you know what, I didn't understand, could you explain me in another way?'... no, there is no space for that"* (S04).

Only two teachers stated that they talked after an assessment had taken place or reviewed the questions together with their students. One teacher stated that *"it is really complex to see all the requirements of 45 students in the classroom... it's complex... in what moment?"* (T01). Only

one teacher used a change in the strategy: *“I change the strategy and come back to cover them [contents] in a different way”* (T03). Time pressures, class size, and focus on curriculum contents seem, once again, to impact formative strategies in the school.

The aspects suggested by the literature as being formative detailed in this section, are usually linked by teachers and students with curriculum contents and grades, and notoriously impacted by the lack of time, which has been identified as a theme prompted by external conditions, imposed as national characteristic to teachers and students. The following sections will describe graded formative strategies described by teachers and students during the interviews. They are relevant to understanding what is formative to them and how it is implemented in the school.

5.7.1 Summative use of formative: rubrics for grading activities

The study found that teachers made summative use of rubrics to assess performance and to grade activities other than tests and worksheets, in different school subjects (a definition of graded rubric can be found in section 6.4, p.145). It must be noted that any of the participants identified the graded rubric explicitly as a summative use of formative. Once again, for them, it is simply an assessment. Rubrics include the aspects to be assessed and points obtained in each criterion following performance criteria. It must be reminded that this was the less frequent type of classroom assessment used by teachers in the secondary level (110 references), as most of the assessment is done with tests (293 references) and graded formative activities (222 references) during the academic year. The assessed activities using graded rubrics included a variety of tasks such as lab work, video creation, oral exposition in the classroom and also in school stands/fairs, friezes, models (made of different materials), acting, debate, and self and peer assessment. The main feature of this type of assessment is that it implies performance in activities that usually differ from writing and can be done individually or in group. An example of these activities are students' presentations in the “Book Day” stand, and grades registered in Language (Spanish) for all the classes in the schoolbooks (PD01 to PD08). When I asked during the preliminary study, the academic coordinator explained to me that in a group, students choose a book to explain to the public, dress themselves as characters, and decorate the stand accordingly. This activity was done usually on a Saturday and all the school staff and students' families participated. Several pictures of the activity can be found in the school webpage. This type of activity was also found in Biology,

where students presented a seminar on sexual responsibility, as explained by one teacher: “We did other activities, they [students] did hormonal regulation models, the chemical structure of DNA models, they work in a sexuality fair as well”. (T03)

Oral presentations, debates, and acting are also part of this type of performance activity based on spoken language skills. Other examples included in this type of assessment involved the creation of products, such as fossil creation in clay, frieze (foldable stories), maps, and lab work, such as the dissection of a cow’s brain. It must be noted that these activities were performed during 2019, before COVID-19 changed education to online.

18/18

70

RÚBRICA AFICHE O INFOGRAFÍA DE PROPAGANDA

INDICADORES	LOGRADO (3 pts)	MEDIANAMENTE LOGRADO (2 pts)	NO LOGRADO (1 pt)
PRESENTACIÓN	El trabajo es limpio, claro y entendible. Tiene sentido completo.	El trabajo está desordenado. Sin embargo, es entendible y tiene sentido.	El trabajo está desordenado, sucio y no presenta sentido.
TIPO DE TEXTO	El trabajo es, fácilmente, reconocible como un afiche o una infografía de propaganda.	El trabajo es reconocible como un afiche o infografía, pero, requiere análisis.	No se puede determinar si el trabajo corresponde a un afiche o a una infografía.
CONTENIDO E INFORMACIÓN	El contenido y la información son entendibles y apoyan la construcción del texto completo.	El contenido y la información no son muy claros, pero, apoyan la construcción del texto completo.	El contenido y la información no se entiende o no corresponde a lo que se pretende comunicar.
PROPAGANDA	Se evidencia un objetivo de persuasión dentro del texto, lo que le otorga el carácter de propaganda.	Si bien el texto quiere persuadir, los elementos utilizados son ambiguos.	El texto no trata de persuadir o convencer de una idea.
CREATIVIDAD	El afiche o infografía es muy creativo. Presenta imágenes y decoraciones didácticas y precisas.	El texto es creativo. Las imágenes utilizadas explican el contenido.	El texto no es creativo o no presenta este carácter.
USO DE IMÁGENES	El texto presenta imágenes acordes a la intención e información presente en el afiche o infografía.	Si bien, se presentan imágenes, estas no coinciden del todo con la intención del texto.	No se presentan imágenes o ninguna de las presentadas es acorde al contenido del texto.

Figure 11: Example of points-based graded rubric used in the school in Language (Spanish).

As a common feature, all the assessment activities involved in the performance theme were graded by using points-based rubrics that were given to the students in advance. They make

explicit the characteristics that the work must have and assign points following the accomplishment of those indicators. The image used as an example set 6 indicators about the work that must be done in Language (Year 9), advertising a product: presentation, type of text, content and information, advertising, creativity, and use of images. For each indicator, students can obtain 3 points if the indicator is accomplished, 2 points if the indicator is half accomplished, and 1 point if the indicator is not accomplished. The numerical addition of the points became a grade using a scale from 1.0 (minimum grade) to 7.0 (maximum grade). Therefore, students know the criteria under which they will be assessed and the specific score for each one:

"I always give the rubric before the work is started and it's like a key for me, because it helps me to collect information, for me, and for them, I take the rubric as a 'route of work'. So, there is always the rubric from the beginning, I try to be really punctual with that, and obviously, the scores are already established following accomplishment levels" (T03).

In general, these planned activities are considered as makings students to increase their numerical grades average as they are opportunities for good performance, and as formats that are regarded as focusing beyond the checking of curriculum contents: *"Because... these assignments are like a way to increase your numerical average, it's like work, but it's easy, although the test is different because they are only about contents" (S06).* However, the planning and use of these activities depend on each teacher, therefore, some subjects do not use the strategy.

Impacted by the high presence of testing strategies, teachers recognise that these types of assessment are labelled as harder by some students, exhibiting resistance:

"Sometimes I feel that there is more rejection to the flexible assessments than to the tests because they are really used to the tests, really used. For instance, it happened to me that they [students] said 'I had excellent performance' [in the subject], and I asked why, and they answered 'Because I took the test and did well on that'... do you see? So, umm... it happens a lot that there is resistance to work in other formats because that implies more effort, they have to create something, and they react like 'that's boring', for them the test is easier" (T05).

This perspective was also found in several students who finally, expressed a contradictory preference for the test format, as was stated earlier. However, performance activities are also identified as an opportunity for students to show their multiple abilities, beyond written content management:

“I try to incorporate other things, that maybe are not related with my subject, but can be more plastic, more concrete... that they [students] create videos, that they make representations, things that maybe are not necessarily related with the proper abilities of the subject, but that allow students that don’t like it very much, that doesn’t have much ability to analyse a historical source, they feel that they also have a space within the class to show their abilities” (T02).

As can be seen, the impact of the testing culture implies contradictory reactions towards assessment where students need to perform following other abilities than the involved in a classic test format. On the one hand, students like these spaces and they use them to increase their grade averages but on the other hand, they can find it difficult as they are more used to working on classic assessments, such as tests and worksheets, which are the most common assessment experience that they have. Interestingly, these activities allow teachers to talk about abilities, which is not the common perspective, as they usually are aimed towards curriculum contents. In any case, each aspect of these activities is scored by points that end up in a grade, as detailed in each rubric that teachers create to assess them. Once again, they decide freely on using this strategy or otherwise, generating contradictory opinions among students and reinforcing the summative and grading culture that is found at the school.

5.7.2 Accumulative grades: formative made in Chile

Within the assessment methods used in the classroom, a particular strategy that teachers use to grade can be found, revealed by the interviews. Instead of grading only one worksheet or specific work/homework, teachers accumulate a series of activities done to finally, after a certain period of time, assign a grade. Teachers have a reason: within the Chilean school culture, if the work assigned to students is not graded, they do not work: *“Children are conditioned to work for grades and not for the learning itself or enjoying the learning process” (T03)*. Another teacher commented: *“I had a really complex class group... where I had a lot of grades because everything must be graded, so they work, otherwise, they don’t do anything” (T05)*. Grades frequency here could be related to students’ behaviour involved in a ‘complex class’. This is again reinforced by another teacher: *“If there is no grade, the guys don’t work, they are conditioned to that” (T06)*. This also means that if you do not comply or you do not work in the activities, you will get the minimum grade (1.0).

It must be noted that the 'formative grade' was made a compulsory aspect of the semester averages at the end of the first semester in 2020, when I was interviewing teachers and students. Therefore, the main feature of this 'formative grade' is that teachers use a series of smaller assessments and combine them into one final grade. Teachers voluntarily shared some PowerPoint presentations (PD11 and PD12) where they discussed this compulsory 'formative grade' in staff meetings with me during that time, and especially when they were checking their interview transcripts afterwards. This was useful information to explore how teachers understood formative assessment and their perspectives shown in the interviews. As the 'formative' concept appeared again in the interviews, I made some informal questions about what was involved to two teachers who mentioned that they were using the strategy of 'formative grade' to assess their students as compulsory grades. It was difficult for me to understand what they were doing as 'formative' until a teacher confirmed that this accumulative work was "*formative... Made in Chile*" (T02), which was included in the title of this thesis. Of course, they had been describing the strategy during the interviews, although the 'formative' aspect involves class-to-class work that is monitored by the teachers, it finally became a grade, as all the formative assessment methods used by teachers in the school.

Therefore, this accumulative work done in class to obtain a grade, is monitored and reviewed by the teacher for instance through Mathematics (see example on figure 10, p.149) or Chemistry exercises, to answer questionnaires in History and Social Sciences, or simply writing the class contents from the blackboard in their notebooks. Questionnaires usually are referred to as open questions and are used commonly for curriculum's content preparation and checking before the test. The main difference, as was stated before, is that these activities count on teachers' and peers' support. Like many activities that can be done during the complete semester, they obtain a grade as a result of their work. As such, this assessment is usually considered as part of the accumulative grade, which counts as an assessment of the process, which was later identified as what formative means for these teachers:

"There is one [assessment] that is done during all the process in the academic year, that is the, like the daily class to class register. This is like monitoring if the students... super measurable things, if the students can write, and if the students can solve the exercises, such as the constant notebook revision, which is done at the end of the class. [The register] is made with a stamp or revision ticket that I register on my personal notebook" (T01).

Following the 'measurement' intention of formative assessment on the school assessment decree (PD09), this teacher seems to define it as they do and understand the concept in real practice and also as was evidenced by the addition of a 'formative grade' to the semester average during 2020. As has been evidenced by different means, formative is, therefore, to grade as the student works and advances with different activities during the semester. It can be understood that these activities also help to monitor students' learning and receive feedback, as they are teacher-reviewed. However, this accumulative work and activities can also represent points (or tenths), instead of a direct grade, that can be used to improve the grade on the upcoming test, a practice that could be labelled as 'score pollution'. These specific features of the Chilean system, as are related specifically to grades, will be further explored in Chapter 6. The obvious question here is how this student's work can be considered formative when it is graded and becomes a formal grade, becoming, by any definition, summative. This also questions how the school actors understand and apply formative assessment, challenging the understanding of the concept presented in Chapter 2 but agreeing with the different focus of the assessment study in Latin America and Chile presented in Chapter 3.

Chapter 6: Assessment information: grades and feedback

*“The people who are struggling the most, are those hit hardest by the system”⁵.
Carlos, 17 years old.*

6.1 Introduction

The previous chapter covered RQ1 on assessment purposes and methods following the school documents analysis and teachers’ and students’ voices. As the Chilean educational system is characterised as high stakes, this chapter will focus on findings on RQ2: What are teachers’ and students’ perspectives about the information provided by the assessment, in terms of grades and feedback, in one Chilean secondary school? Assessment information is understood as the data that students receive from the assessment process. Following the Chilean assessment policy, it will be presented then, in two main sections, following the features of SBA: assessment results in terms of grades (summative aspect) and assessment results in terms of feedback provided to students (formative aspect) and how this information is received by them. This chapter incorporates the complexities of grading, the high stakes of the system, score pollution, and feedback provision themes, always influenced by acute social inequalities in the country and the teachers’ lack of time and work overload.

6.2 Grades on the school documents as background information

This section will provide background information about grades by the schoolbooks register analysis and was identified as relevant to understanding teachers’ and students’ perspectives on the topic. Related to the frequency of assessment, during the 2019 academic year, it was found that students at the secondary level were subjected to 625 assessment instances in the core units analysed in this study, each student had between 88 and 65 graded assessments that are registered in the schoolbooks. All these grades are considered within the high stakes of the assessment system (promotion, certification, and university entrance) and have a direct relationship with the summative and testing culture theme. Following the logic of assessing all the

⁵ This excerpt was part of an assignment submitted for the Unit ‘Introduction to Qualitative Research Methods on the Social Sciences’, where interviews with two students on SBA were conducted in January, 2019.

work that students do, the following table contains the details of the number of graded assessments made in each class and subject, and in all the classes at the secondary level of the school during the academic year:

<i>Class/ Subject</i>	<i>Language (Spanish)</i>	<i>Language (English)</i>	<i>Math</i>	<i>History and Social Sciences</i>	<i>Biology</i>	<i>Chemistry</i>	<i>Physics</i>	<i>Total</i>
<i>I° A (PD01)</i>	17	11	16	12	9	16	7	88
<i>I° B (PD02)</i>	19	12	18	11	7	14	7	88
<i>II° A (PD03)</i>	16	12	13	11	10	12	7	81
<i>II° B (PD04)</i>	16	13	17	12	9	12	7	86
<i>III° A (PD05)</i>	14	12	11	11	9	12	6	75
<i>III° B (PD06)</i>	13	12	13	11	7	13	6	75
<i>IV° A (PD07)</i>	14	11	11	11	7	7	6	67
<i>IV° B (PD08)</i>	13	10	10	11	7	8	6	65
<i>Total</i>	122	93	109	90	65	94	52	625

Table 13. Grades registered in schoolbooks per class and subject in the 2019 academic year.

This impressive amount of assessment procedures implied for a teacher the revision of thousands of assessment instruments during the school year, considering that most of these classes had 45 students and also that there were between 122 and 52 assessment instances for each teacher during the academic year. When these numbers are considered, it is necessary to recall that one type of grade called ‘accumulative’ (or ‘formative grade’) implied several different assessments and that work that is subsumed under one final grade after a certain period of time. Whatever form the assessment takes, it always results in one grade. Even when this study had a qualitative focus, this data reinforces the high frequency of the assessment and also the summative and grading culture present in the country. The schoolbooks analysis (PD01 to PD08) contained mostly grades registered, evidencing the primacy that this procedure has inside the day-to-day life in the classroom. The information is usually presented in a numerical language, with fewer details about different aspects of the assessment, other than grades. It will be noticed that this will impact feedback provision as well: students only receive a numerical grade for their work and that is the main language communicating their results to them and their parents/carers.

6.3 Summative and grading culture through teachers' and students' voices

As described in Chapter 5, it is relevant to remember that all of the assessments are graded, as was explained by teachers in the interviews, evidencing a highly summative assessment style:

“Sadly, in this context the context in which we work, everything needs to be graded. Otherwise... they [the students] don’t work. They just don’t see the point of working in a personal way... we haven’t accomplished moving towards...mmm... ‘I have to do it because I want to learn new things’ from ‘I have to do it because this involves a grade’” (T01).

This teacher’s opinion highlights that the dynamic inside the classroom is marked by graded assessments (summative), with the belief that all must be graded, every activity or homework, otherwise the students do not work. This also reinforces that when talking about assessment, teachers always link the activity with grading, evidencing the summative culture. Despite the Mineduc policy intended to promote formative assessment, the school assessment dynamic remains mainly characterised by grades production: *“It is just that, when the reforms are on paper, they are very pretty, but then when you try to apply them to reality, they are undoable”* (T03). Teachers critically reflect on the testing strategies producing grades and on the accuracy of these grades to reflect students’ learning:

“Sadly the, the model of, grading that we have is not always adapted to the knowledge that the students acquire, umm: the written test. If it’s true that it is an instrument that is favourable for the teacher, several times it doesn’t reflect what students have learnt. It happens a lot to me that the students have horrible grades in my subject, but if you ask [theoretical] questions to them, they do like the subject and they do have the knowledge, there is... a disparity generated there, I think for the type of instrument that we use, it is a very rigid instrument” (T05).

Even when teachers choose test formats for practical reasons, they recognise that they do not always seem to be the best choice for reflecting students learning for their rigidity, especially for the use of multiple choice formats, which also invites us to question its validity and reliability. This is especially worrying when the high amount of tests that students have is considered (293 references in the schoolbooks as a main pattern, as was detailed in Chapter 5), as they could reduce opportunities for students to show their learning in other formats. They are also a practice that keeps occurring. Another impact of this reality is, as has been mentioned in the previous

chapter, that the main aim is to get a good grade and not to learn: *“They study, or whatever [activity], they do the work only for the grade”* (T08). A contradictory element here is that even when the school assessment decree requires different types of assessment, methods are impacted by teachers’ lack of time and their freedom to decide. Therefore, the numbers production provokes an essentially empty learning or superficial style, mainly based on memory skills, that does not reflect real students’ abilities: *“I don’t know if the grade is the best way, because it does not necessarily reflect the abilities or capabilities that a given student might have”* (T03). This reflection has been also presented in the previous chapter, as the focus, rather than students’ knowledge and abilities, seems to be the curriculum and grading its coverage. Beyond how reflexive a teacher can be, they keep grading mainly by paper and pencil tests.

Teachers’ reflections about learning often include the possibility of making changes to the grading system. However, grades are a language that parents understand and even demand:

“The huge transformation that must be made should be in the line of, at some point, eliminating the grades, that grades were not a number that defines how much you learnt or not. Moreover, as happens in other places, the student could be assessed holistically as a complete human being... the change is deeper and more radical, and also harder to accomplish because it has to be with a mental change. I mean, if I told parents ‘You know, your son is very sociable and has lots of leadership’ but you give them a 4.5 average, that parent is not interested in that, and will ground the student likewise... Because we are programmed to function in relation to grades, so I think that the change must go in that direction. We need to understand that education is a tool that will help them in their lives, more than simply accumulation of grades that, in the end, I left school and the grades are not useful at all” (T02).

This teacher places a question about if it is realistic to think about eliminating grades within a highly summative culture, whereby the system functions only by grades. In this sense, it must be noted that students’ reports contain only grades and attendance percentages and do not include any written information about the students’ learning or abilities. This is the only information that parents officially receive in parent/carers meetings of each class. Individual parent/carers meetings are usually focused on low-performance students. For instance, if parents/carers bought materials for one activity, they demanded a grade as a result:

“I would love that there were no grades, it would be cool... but in fact, sometimes parents come to complain when there isn’t any grade for something, they say ‘but I bought materials and where is the grade?’” (T03).

Grades appear at the centre of the learning process, including practices and a language that seems to be profoundly rooted, impacting beyond the classroom towards parents, who have learnt to demand grades as a result of the students’ learning process. This shows how learning is a systematic process affected by elements that go beyond the school environment. This can be perfectly exemplified by the structural conditions impacting the findings of this study, such as social inequality.

However, as can be seen, teachers and students are also capable of questioning the meaning and adequateness of those grades, reflecting on what they represent. This is because the system keeps the grades production within a rhythm that is based on the school’s demands and not students learning rhythms, leaving several unintended consequences in the process. It was seen that the grading produces a large number of grades and assessed instances. Some of the consequences are directly related to the high stakes of the system. However, other consequences are related to teachers’ and students’ perspectives of the grading and also some psychological consequences that will be further explored in the following section.

6.3.1 Negative consequences of grading

Teachers and students highlighted some psychological consequences of grading that are relevant for this study as the assessment results (RQ2) could be prompting a negative impact:

“Yeah, the grades are like a threat, a tremendous vice... yes, I feel like, it is related to..., everything is related... in the end, it is a strategy of pressure as well and of control... So yeah, it causes me... I don’t know, it is uncomfortable that pressure exists for the grade and I see the students super stressed about the grades thus it is sad for me, because I try to give them the discourse ‘Guys, you are not a grade, you are worth more than that’ but on the other hand, all the time I am grading them... you see?” (T03).

Teachers grade because they must comply with assessment calendars in a dynamic where it can be used as a strategy to threat, where students feel that the resultant numbers do not have real meaning, being labelled and stereotyped in the process:

“For me, it’s like... [grades] don’t measure your capability, don’t measure what you know, don’t measure what you know about life, what you know about yourself... Because in a certain way, for teachers, it is like, the grade is to know how you are, you got a 7 [maximum grade], you are smart, you are responsible... but if you got a 2 [minimum grade], you are dumb... no! you are lazy, you don’t work” (S04).

Students also commented on the stress caused by the grading processes that also have an impact on their self-esteem: *“There were days when I felt super vulnerable in that sense, I felt dumb, I felt like... I wasn’t motivated like to improve on my studies” (S04) or “I would love to have better grades... yeah, to feel good about myself” (S10).* The assessment results impact students’ self-esteem and motivation, which must be understood within a classroom when they are constantly assessed, as was stated previously, mainly by using traditional assessment formats, such as tests. Additionally, these grades will have a say in their future possibilities to access higher education. The impact of years of grading for these students at the secondary level also brings other unintended consequences and elements to consider, as can be seen in this teacher’s answer:

“It is extremely complex that they acquire like the, the love for the subject, because they come already frustrated from past years. So, how you can take off from their heads, from the students that come with that mind-set of ‘is that I suck at this’, ‘it is that I don’t understand’ [...] ‘I never have understood [subject name] and I never will’... there is an emotional aspect of them, the frustration towards the subject. And besides, parents are also a part of that, sometimes in a negative way, because they pass on to the students their own ideas, like ‘I was bad at this, so my son is also bad at this, I understand and I let him be’... so they understand, don’t provide support, they just let that be as it is” (T01).

Therefore, self-labelling seems to be linked to the intergenerational transmission and appears to be impacting students’ expectations about specific subjects, prompted by the grades that they obtain for their performance. This could be also related to the social inequalities of Chilean society and the cycles of power of privilege described in Chapter 3. Some students even declare that they are allowed by their parents to simply fail the subject because they cannot learn: *“Here they have told me like ‘this is not going to be useful at all’ or ‘I don’t care, in this subject, I’m allowed to fail” (T06).* In that sense, teachers considered parents’ participation in their

children's education a key element that must be in line with the school and also the student efforts.

Students also referred to different situations prompted by grading, such as some teachers using students' names as examples of good performance all the time, or making them feel like their opinions do not count and even giving students' low grades out loud, embarrassing them. They also feel that students are treated differently solely because of their grades:

"Yeah, I felt that [female classmate name with the best grades average] she is treated super differently, and it's like [male classmate name] he is constantly scolded, '[male classmate surname] sit down, shut up', things like that, and when she [female classmate name] talks, they don't say anything, they let her be, things like that" (S06).

The grading system impacts a number of different elements, such as self-esteem, self-perception, in-class students' treatment, and labelling/stereotyping.

It is interesting that, when reflecting on students' learning, teachers mentioned the importance of self-reflection beyond blaming only the student for their learning, which could be a key element to understanding the negative aspects of the assessment process:

"But I think that the teacher there has a lot of responsibility, in the constant learning, you see? Because the teacher tends to blame the student, 'Ah, you did it badly, I did a brilliant instrument', I mean, I have zero analysis capability or critical capability, you see? I think that teachers lack that more generally, they lack umm... they need to get out from their egocentrism of telling all the time 'I know everything'" (T09).

This teachers' reflection about their responsibility for the students' learning seems also to expand to external factors to them, such as the reliability and validity of teacher-produced assessment instruments and their impact on students' results.

Additionally, teachers reflect on other elements, such as the students' emotional well-being, and also consider their family situation as impacting their possibilities to learn:

"Look, I'm convinced that more than results [...] my pedagogical style is centred on the emotional, for me that one student, I don't know, left the classroom happy is far more potent than a result. And also considering that at least our students here, they share such family backgrounds or contexts that... sometimes, you can even imagine the problems that they bring, so for my part, the focus is not on results but rather on my empathy towards them and I think that that is far more important" (T10).

In that sense, for this teacher participant students' learning seems to be mainly focused on creating a supportive emotional environment where students can study and learn that goes beyond academic elements and it is more interested in students' wellbeing than their numerical results. Repeatedly, the analysis of learning is made beyond the huge amount of numerical data that the assessment procedure creates. Therefore, the following section will focus on different aspects revealed by the interviews, that showed elements that are worth considering when talking about the assessment results, especially in terms of grades and their consequences within this high-stakes system.

6.3.2 Grades differences and achievement: motivation, expectations, and cultural capital

The grades that are a result of the assessment process are related to students' achievement, an analysis that is an important aspect of the educational process. Teachers exhibited high reflexivity levels about their responsibilities regarding students' achievement:

"I think that we have to stop blaming the student, first of all... to work with the student, the directive team, you know? This comes from above because the teacher could have brilliant ideas, and a lot of creativity to apply in the classroom, but they cut you down in a certain way. Umm, here there is a system that you must comply with, rules that influence our work... So I believe that we must stop blaming the student and put more focus on the teachers, what teachers do I have, and the tools that the teacher needs to give a good education" (T09).

These reflections show how teachers engage with their practice in different ways, that are always affected by external school structures that impose conditions that can limit their work in the classroom. Most of the participants talked about their responsibility for their students' results as a first approach to low performance.

"I don't know if the problem is a student's problem or mine... but umm... no, the problem is focused on the teacher. Sometimes we take for granted things, I understand that the student manages things on some level [...] and sometimes we have to recognise, as teachers, that I take for granted things that can affect the students, but that is also because of the fast-moving nature of the curriculum, all the stuff that you have to cover, so that can be affecting" (T10)

Once again, teachers' work seems to be impacted upon by curriculum coverage which could cause them not only to reflect properly on how they are doing in their classes and their students' needs. However, teachers also cited several different reasons impacting results, such as students' aspirations and motivations and also their relationships with each teacher:

"I think that the thing is like the motivation of each of them. Yeah, I think that the main theme is motivational and aspirational. Often it happens that students with low performance, it is because they don't have any clarity towards, about where they are going. It's like their path is extremely uncertain, so they don't have any clarity in life" (T01).

This was especially important for them as the difficult context where they work demands an emotional relationship with the students and an awareness of the difficulties that they could face in their lives. Teachers also mentioned as an important factor the need to make different activities where students with different interests or abilities can be engaged by considering the diversity when planning assessment methods. However, this was in contrary direction of the high amount of tests taken by the students in the school. In that line, teachers mentioned that the system seems to be highly structured and rigid, making it difficult for some students to fit in and also highlighting their cultural capital as a reason for failure:

"There are guys that are better suited to the school system as it stands than others. There are children that for family issues, or personality, they fit better in the school system and that allows them to have a better performance... versus, versus other students who sadly, they don't have the capabilities, they have strong personalities that sometimes umm... clashes with the system that is too structured, more authoritarian. And that generates low grades... Now, I feel that the cultural capital topic, at least on my subject, is fundamental. I mean, it can be really noticed when children bring a greater cultural capital... that is interested in topics that the others are not, and that generates a huge gap... I have worked in other schools where if you, as a teacher, don't generate that cultural capital, that student is not going to have it even if he/she is very smart or adaptable to the school system, they won't, because they don't bring that from home, their contexts don't allow them to generate cultural capital, and that's umm... generate an important distance" (T05).

It is interesting how teachers consider the influence and role of the family and parents on students' cultural capital impacting upon their results and even creating gaps within the

classroom that are a challenge to consider when working on their subject's topics. This also relates to the segregation and inequality of Chilean society described in Chapter 3 (structural condition theme) and the subsidised schools where students from different backgrounds and abilities can be found:

"The lack of knowledge about the diversity of students that we have, the different systems, and also the worry of, of this segregation, that I feel that, every time the students are going to fell [the segregation] in some sectors, be more incompetent in some abilities. I mean for example, the opinion [as ability], as I was saying before, the critic, to take these elements, these argumentative spaces where the guys, I don't know, can be open to dialogue... umm, that is also a factor which worries me" (T10).

This teacher shows worried about some of the students' abilities, such as giving their opinion, argumentation, to criticise something. It is usual that students from low socioeconomic sectors speak differently and use Chilean slang, which I commented on in the methodology chapter I can talk about. It is worth mentioning here as well that students' motivation and expectations are impacted upon by the well-known results of a few of them being able to go to University. For teachers, it is clear that sometimes students feel that they are a grade and also exhibit evidence of years of feeling not capable, which impacts on their academic performance and motivation:

"If they have a bad grade, they assume that they are that, that he is that grade, and therefore is bad for that subject. It is hard to free the children from that grades influenced self-perception of themselves that they have, but I think that with a systemic change, that can be done" (T03).

The bad performance stereotype threat and its impact on grades also make students feel completely closed to some subjects, by simply saying that they do not know anything, even after years of being taught on the subject. Changes can be made, following this perspective, with a "systemic change", that could be related to the overall educational system rather than particular settings of the school. This reinforces the relevance of identifying external historical structures impacting SBA in the school that this study has highlighted as transversally impacting its findings: social inequality and teachers' lack of time and work overload. These themes also are part of the systemic change that is needed, as highlighted by one teacher here.

For teachers, students' learning is monitored by observing certain behaviours such as motivation and participation rather than the grades analysis per se. Therefore, even when the

assessment produces an impressive amount of numerical data, they reflect on learning by analysing other variables:

“I think that students’ participation, when students do participate in the class, I think that is an indicator for me that it is working, that the guys are processing, acquiring knowledge... umm... reflecting, etc. Umm... it happens to me that there are classes that are difficult that they participate in, and there you have to find other indicators, other variables... participation or motivation, and being super honest and self-critical, I don’t like the grades, like the numerical theme is my indicator, truth to be told, no” (T05).

One interesting feature of reflection achievement is that for students, it is impacted by their capabilities and interest in their studies. However, they also mentioned teachers’ methodologies and the relevance of mutual interest:

“I think that [achievement] is impacted upon by the methodology that the teacher has, but also it involves the student, because one, it is like you excuse yourself by saying ‘it’s that I don’t understand this teacher, he/she doesn’t know how to teach’, but maybe it takes effort from us as well... and we don’t seek for help in other spaces... Although I know, that shouldn’t happen because the teacher is in the classroom to teach you... I think it cuts both ways” (S03).

Interestingly for students, central external factors from the classroom, such as motivation or expectations are not mentioned in relation to their achievement. However, they are mentioned when the high stakes of the system are considered in terms of repeating a year and/or university entrance. Therefore, the achievement is considered more in terms of the grades’ consequences rather than the achievement of learning itself. This will be further explored in the following section.

6.3.3 High stakes: repeating/promotion and university entrance

The school assessment decree (PD09) establishes the school as being responsible for certifying students’ promotion from one level to another, however, the leaving certificate at the end of secondary school is finally issued by the Ministry of Education. This certificate or licence of secondary school allows students to progress to University, if and when they accomplish the requirements established by the law and by the individual Chilean universities. As has been stated previously, these requirements are based on the score obtained in a special university entrance test

(PSU during data collection, PAES since 2022), also from the grades obtained during secondary school (NEM) and the students' position by numerical average based on grades, compared to the cohort students in the school that year (Ranking).

Following the national assessment decree explained in Chapter 3, the students will be promoted to the next level or class when they pass all the subjects in their study plans, or when they fail only one subject but have a yearly grade average of 4.5 (including the failed subject) and or when they fail two subjects but compensate with a yearly grades average of 5.0 (including the failed subjects), (PD09).

When the high stakes of the system are taken into consideration, two of its consequences appeared immediately as being related to the assessment and grading process: repeating and university entrance. It is worthwhile recalling that the assessment methods used by teachers are designed following external test formats, such as the one used in the PSU (PAES since 2022). Repeating and university entrance are framed by teachers and students within a scenario where students "*really deserve or need*" (T04) to repeat or go to university. In this study students who had the experience of repeating an academic year, leavers who were later able to enter the University, and others who did not all participated. The next two sections will explore repeating and university entrance separately.

6.3.3.1 Repeating the year

Due to the COVID-19 pandemic, discussions and requirements about flexibility were held and evidenced in the school documents (PD10, PD11, and PD12), calling teachers to "avoid repeating students" in response to the sanitary context. This was considered relevant, as the pandemic forced the school to make more flexible the learning process beyond its typical summative culture. Even when teachers could grade students qualifying for repeating, the school principal and an 'ad-hoc' commission (PD12) had the faculty to promote them when exceptional circumstances could be corroborated. Typically, at least one or two students per class in all year groups repeat the year, which also happened during the COVID-19 pandemic⁶.

Both teachers and students presented similar perspectives where some students deserve to repeat if they have not been fully engaged in their learning. However, that is a complex opinion

⁶ The country had intense discussions about repeating students during the COVID-19 emergency who did not have computers or internet connection at home, as well as students who took care of their siblings while their parents were working or had dead family members as a result of the virus.

when they balance the positive and negative impacts that the decision can have on students' lives to finally, showing themselves to be insecure about the repeating strategy. Teachers reflect on the lack of time and monitoring of the students' progress during the academic year to prevent them from having to repeat. Additionally, they consider the repeating students as being stereotyped and disengaging from the educational system:

"I feel that repeating is an inherent vice of the system, because no, it is not necessarily going to generate an attitudinal change in the student. Several times it generates the opposite, to reject school altogether...umm... to force him/her to be taken away from its context, sometimes their friends and all... and he is super disadvantaged. It generates rejection from classmates in the class where he/she arrives, all of them know that he/she is a repeating student, so... And additionally, 'he must be lazy', so nobody wants to work with him, 'he must be dumb', and thus a series of prejudices are generated. And it becomes difficult, let's say, some students never get to fit in other classes... umm, I feel that this is a systemic failure because that means that there wasn't a proper monitoring process, not only from the school or teachers, I'm not blaming the tutor teacher for not realising, because you usually know where the matter is going to, but also there is a family theme meddling that we didn't know how to read with clarity" (T05).

Besides stereotyping and disengagement from school, teachers recognise that once again, a lack of time prevents them from identifying their students at risk on time. They also reflect on other factors such as family issues. For a repeating student, the impact of the experience is also related to family pressures and other feelings such as shame:

"Because I didn't study for myself, I studied for my parents, that's what I got, I studied and had good grades simply for my parents, so they wouldn't be ashamed and stuff, because when I repeated, to see their faces was horrible... And [repeating] doesn't help us all, not everybody takes that impact in a good way, because there are ones that feel like dying and don't want to keep studying anymore. For me, it was useful, but maybe for the person next to me, it is not at all" (S04).

Teachers also recognised that if there were more than tests as assessment methods, they could realise early on when a student is struggling, differentiating between them more clearly:

"Besides, the fact of not diversifying the assessments, umm... and not diversifying how I assess my students differently, umm, depending of the context. Sometimes it

prevents us from identifying things that, that we could have identified, maybe there is a student that hasn't accomplished the required level, I don't know..." (T05).

Students also recognise the impact of the assessment methods used on the results and also on repeating students:

"I think that several teachers or several persons consider that they [repeating students] are people who don't... they make them repeat because they can't advance because they didn't learn something. But the truth is that they are people that can advance and did learn, but in other ways, not in the way in which they assessed us" (S01).

As has been previously set out, despite the negative impact of the assessment methods used by teachers as preventing students from showing their learning, they keep being based mostly on testing and, also, testing in SIMCE/PAES formats. That practice is identified by one student as a recurring contributor factor: *"That people that are in repeating danger, I'm sure that if they were to be assessed by a presentation, an oral exposition, I'm sure that they would perform so much better than they do in a test" (S01).*

When reflecting on students' responsibility for their results, some teachers tended to mention students' immaturity to face their studies. Therefore, repeating could help them to be more mature: *"Now I feel that several times, the repeating is generated by lack of maturity" (T05) or "more than anything else for the immaturity that these guys have [...] immaturity and laziness" (T06).* Even when the repeating responsibility can be placed only on the student, most teachers interviewed in this study recognised the negative impact of repeating and showed themselves to be reflexive about their responsibilities in that regard:

"We can't just make the students responsible, or their parents. I'm convinced that we, as teachers, should be constantly making those revisions, what have we done, I mean, if I as a teacher, openly I'm going to grade with, I don't know, 3 or 4 reds [failing grades] and then, by the end of the year or by the end of the semester I'm going to be telling the student 'you know what? You failed the subject'... I don't know how meaningful and how responsible, as a teacher, is that" (T10).

This reflection evidences the way that teachers are aware that the repeating culture should involve different actors and deep reflection about the teaching strategy. In any case, teachers seemed to consider that students' efforts are a key indicator to decide who 'deserves' to repeat and who can be saved from that. This can also be connected to grades and score pollution

strategies and mainly with education as a wider system with elements and actors beyond the school setting:

“It is going to depend a lot, I insist, on this triangle [student-teachers-home] that I’m telling you. Thus, if the student has family issues, if the student maybe umm, is a good student and for some family issues factor the student has to repeat, but you considered that, I don’t know, he/she was a good student, I think it is a punitive measure. Now, if the student obviously didn’t put effort in, there wasn’t work from his/her part, and the family is not interested in his/her education, the educational process as such, and we as teachers didn’t worry about the process of the student getting motivated, then I think no... I think that in that case, well there is a questioning of all the sectors, actors involved... we tend to, teachers tend to blame only the student for repeating. I think that repeating in, in some cases we can say that it is a punishment strategy, in others it is as a way to make students understand that I must put effort in, that I need to overcome all... all the difficulties that can appear, and that I need to put all of it on me to be able to work and make it” (T07).

Teachers balanced the impact of different elements when reflecting on repeating students, making the topic complex in ways in which all these elements are finally evaluated to decide at the end of each academic year, that is December in the country (usually students know if they have to repeat before Christmas). Teachers also recognise the impact of grading and their responsibility in the learning accomplishment, and even societal pressures in having to repeat:

“How is it possible that the children didn’t accomplish nothing, a minimum at least... as we work here in the school, that happens though... that sometimes the kid doesn’t get the grades, that is a 4 as a minimum, or in truth that reflects the fact that the children didn’t learn nothing... because it depends a lot on the assessment instances that you present to the children so I don’t know... I think that repeating is not good, I think that children shouldn’t repeat, and I think that it is a challenge, like balancing the assessment because in depth, the grade is like... not necessarily going to reflect what the children have or haven’t accomplished. Yeah... and that also tells you about a deeper social issue, not only here in the school, because for parents it is also a power instrument, to be able to control children with, through the school grades” (T05).

This reflection tells us about the balance that the assessment should have to prevent students from repeating and, above that, to make them learn. Interestingly, this teacher also reflected on grades as a power instrument for parents, which adds to the idea of grades as a problem of the system that can be even used to control students' behaviour. Within a high-stakes system, it is not a surprise that parents/carers appear in the grading complexities theme as an important element when considering grades, as they care about their consequences and are impacted by them. In any case, a consequence for repeating students was clearly shown as they were found to be disengaging from school and even deserting the system:

"I'm against repeating, that creates a system where the children just get disengaged and demotivated... lot of students that repeat are brilliant guys, very capable, with analysis capacity, very critical, but they simply disengage from deep inside, and as a rejection sign of this structure, umm... they don't do, they don't work and end up repeating..." (T09).

Repeating is, therefore, seen as a major sign of rejection of an educational system where students cannot show themselves for different reasons. This was also reinforced by another teacher who had experienced working in adults' education, which are programs where people who had deserted the traditional school system can finish their secondary education during the afternoons:

"There are several desertion cases where the students repeated, or a teacher openly said to them or their parents that they should look for another school because they weren't capable... I don't know... and they leave the school. And they are children, young people, you find guys that are 13, 14, 15 years old approximately, and you say but why aren't they at school?" (T10)

The link between repeating and students' desertion was also clear for teachers. However, the strategy was shown to be complex, in the sense of being necessary in some cases but also containing contradictory evidence of its positive impact on students' learning, which reinforces the view of mixed feelings and perspectives about repeating. As such, the topic seems to be surrounded by reflections on student effort, the impact of the family, and also about the assessment methods used, making students disengage from their learning process. Although repeating students is one side of the high stakes of the assessment system in terms of promotion, there is another side which is the impact of grades on University entrance.

6.3.3.2 Use of grades for university entrance and structural conditions: inequality and segregation⁷

Teachers immediately recognise the impact of school grades on university entrance in relation to the differences and a generalised social segregation of the educational system (structural condition theme), which previously appeared when the assessment results were considered by the interviewees:

“Look, with the system that we have, with this super segregated system that we have, umm... I don’t agree with that; you know? Because we have to be honest, I mean, maybe this 7 [maximum grade] that I got here is not going to be the same 7 that I could get in another school, or without looking at it as less I don’t know, in a comparative way, state, private, subsidised... I insist, comparing a little bit the patterns, if maybe the educational system was more balanced, the NEM would be more meaningful, but as I see this segregation... umm, no, it’s not fair, it’s not fair [...] When you have a segregated system, grades shouldn’t be a factor” (T10).

Interestingly, teachers seemed to identify that grades lose their meaning as they are produced within a highly segregated system that has been described in Chapter 3 and also identified as a main structural condition theme. In that sense, the education itself would be unable to help to solve these structural issues from society and the educational process would be placed in a dynamic where nothing could be done instead of being the place where they can disappear. As has been previously stated, the ‘NEM’, which is the students’ grade average during the four years of secondary school, represents a part of the final score that will allow them to enter the university, or not. Students showed themselves to be aware of the impact of the grades on their future but they still think, as has been stated previously, that they cannot show fully their capabilities:

“You must have a good NEM, have a good PSU [PAES] score, and that until today continues to be difficult due to all the things that I commented on before, that the grades system is not as accurate as it should be, because there are people that have different learning capabilities” (S07).

The extended use of tests is once again cited as constraining students’ possibilities to show their full capabilities. As a result, secondary school grades or NEM are seen also as meaningless as

⁷ Two students from Year 12 interviewed in the study accessed the university, both in traditional public universities with high demand and scores requirements, and obtained also scholarships to study for free.

they only reproduce the societal segregation where grades impact negatively students coming from lower-income sectors of society:

“It is complex because you end up in the NEM, which is super unrepresentative, because there are students from schools umm, with academic excellence, graduating with a 6.0 average so it’s not the same that for a school, that a student leaving this school with a 6.0 average. Obviously, gaps are generated, here our student with a 6.0 and a student from the National Institute⁸ with a 6.0, that one is going to get 800 points whereas our student is going to get 420⁹” (T01).

According to teachers, despite having the same numerical averages, students from different schools’ backgrounds will show significant differences in their university test results, impacting upon their possibilities to access higher education. This has been sufficiently dealt with in Chapter 3 when highlighting the high social segregation of the country and also when explaining that the Chilean educational system does not have compulsory examinations, therefore promotion and higher education access are directly impacted by students’ grades following the SBA system. This speaks to a low correlation between school grades and university test scores that shows differences following the students’ socioeconomic background. Therefore, they question the impact of grades on the students’ university entrance:

“I mean, I think that is good to consider the performance during these years for the university entrance, but I don’t know if the grades are the better way... because of the same factors we have been talking about, grades do not necessarily reflect the abilities or capabilities that the student might have... sometimes they reflect more only the motivation [...] NEM is like a sentence because the truth is that in Year 9, I don’t think you are mature enough to know the importance that [grades] are going to have in your future so I think that is fine to consider the students’ performance but I don’t believe that NEM is the best way to do that, not in these conditions” (T03).

As can be seen, the impact of NEM on university entrance is seen as being a complex factor because of the conditions in which the grades are generated, where they are not a trustable

⁸ The National Institute is a public secondary school in the centre of Santiago (the capital city), that selected its students, before a new law forbade the practice. The school had historically good results in all the measurement tests used for university entrance (PSU/PAES) and school accountability (SIMCE).

⁹ During the preliminary study, I analysed the correlation between students’ NEM and Ranking with their PSU/PTU scores in 2020. The amount of students obtaining high grades in the school is dramatically reduced when PTU scores are considered, as they are far lower and only a few students obtain scores within the 600-650 category, on a scale up to 800.

product of the students' abilities and learning and therefore, they are an unfair factor for teachers and students. The latter recognises that these combinations of factors into copy and cheating strategies used in the tests:

"I think that the copy thing in the tests is something done out of despair because you are seeking not to repeat, or looking... they always tell you 'you must have good grades' and now 'you must have a good NEM, a good ranking', and in that despair, you want to copy or make a cheat sheet, or do something to ensure a good performance" (S01).

Even though only one student out of ten recognised this practice, it was also inextricably linked to the tests used and the pressures generated by the high stakes of the assessment. Limitations of the assessment methods reduced them to tests and to standardised test formats, focus on numerical results, limitation of possibilities of showing different abilities and capabilities in the assessment, the perpetuation of the based inequality where students have been born and also the characteristics of the system as preventing students with low performance to be helped on time. Some of these consequences are based on structural conditions that are imposed on teachers and students by the educational system. However, others are based on actions designed by teachers within an educational space where lack of time seems to be guiding the process more than the policy or assessment guidelines themselves.

6.3.4 Grades manipulation through score pollution

When it is considered that grades are the main resultant data of the assessment process, several strategies that teachers used to manipulate them appeared during interviews. The strategies for low performance appeared as opportunities to increase the numerical grade that would later be registered in the schoolbooks. In any case, these strategies were linked with learning increase or abilities improvement: the only and main objective is to increase the grades. Consequently, these practices are linked to summative strategies rather than formative ones. One of the first activities for low performance was already mentioned in the previous chapter: teachers offer tenths (increments) for activities that increase students' grades in the upcoming assessment, usually a test. The tenths can be used as 'SOS' in the case of grades under the passing level of 4.0, such as 3.8 or 3.9. In other cases, the tenths were associated with activities to reinforce the acquisition of the contents, such as oral presentations or written questionnaires:

“For instance, I use the system of preparing the tests. Then I make them questionnaires for study, or I make them create a scheme or a conceptual map... that synthesizes the Unit’s contents, and for that I give them tenths, that in general are scarce, I put three tenths, which doesn’t generate an impact in the sense that is not going to modify, to inflate the grade, but to motivate them and secondly, it helps them to study, it is a study way, for your test” (T02).

Therefore, the activities are made to motivate students to be prepared for the upcoming test, and in this case, the teacher considered that the practice did not modify greatly the grades obtained by students. Generally, these strategies are used to avoid low grades in the assessment and even when they are proposed as voluntary, almost all the students tend to comply with them. On other occasions when students obtain a failing grade, they can do assignments with grades that will be averaged with the failed grade, allowing them ultimately to pass. Students mentioned that, in some cases, they can retake the test and also mentioned that the numerical scale could be modified and adapted to increase the students’ results. That is done by considering the highest score based on points that one student had and creating a new scale for them from there. The accumulative grades were also mentioned as an achievement improvement strategy, as if they comply with them they will secure a good grade that also helps to increase their numerical average in case they get lower grades in some subjects. A final reflection about this topic is that when I analysed the schoolbooks and identified the high amount of grades registered (PD01 to PD08), it was impossible to know how many of those grades have been increased through the ‘tenths’ system, questioning the validity and reliability of the grades and their overall meaning.

6.4 Feedback on documents and through teachers’ and students’ voices

Previous sections have presented findings on information provided by the assessment results on summative terms, namely, grades, describing several consequences, some of them, at the psychological and parental level. These sections will focus on another relevant aspect of the information provided by SBA, which is the formative aspect of feedback. In this regard, the school assessment decree (PD09) contained scarce information about feedback. It establishes that feedback for students will be always considered, which seems to be a formative requirement highly promoted by Mineduc (2020b) (it was also highly promoted during the pandemic, SD02). The document does not define feedback characteristics nor a strategy to check its presence. Even

then the school assessment decree explains that homework must have feedback and follow-up activities about its quality and pertinence, this is usually associated only with points to increase grades, accumulative grades, or simply grades, with no feedback at all. During the worldwide pandemic in 2020, Mineduc (2020a) published a document about “Assessment adjustment due COVID 19” on March 27. The document (SD01) suggested feedback as a formative assessment that “could be transformed into summative assessment with a grade” (p.9). Once again, the document did not include any guidelines about the feedback characteristics or formal requirements.

This section will discuss the feedback associated with the assessment process through students’ and teachers’ voices, understanding that feedback is a vital part of students’ learning and improvement. The first impression analysing the data set is that feedback is scarce, due to lack of time, as students receive a numerical grade and that is all the information resultant from SBA:

“On the other hand, with the written assessments, I have to recognise that for me it’s harder to give feedback. I feel that [a test] it’s a more helpful instrument, useful when you are in a rush, so the truth is that there [feedback provision] I fail... for the written test is difficult and I must assume that with total humility... I mean generally speaking, the incorrect ones are marked, the correct ones, the grade and that’s all... [...] It’s hard to give individual feedback when you have to give that for 45 students” (T05).

The grade represents the main information that students have to keep learning and trying to improve their performance, making it harder to provide feedback beyond that. For teachers, feedback seems to be impacted upon by time constraints (as a structural condition), by being focused on mistakes identification and also more linked to tests, especially standardised ones. For instance, a teacher working in all the classes on the secondary level marks approximately 350 students only for one assessment. Time constraints motivate teachers to make informal, verbal collaborative feedback rather than formal, associated with the results of standardised tests:

“Feedback is always done collaboratively. So they work in groups [analysing their individual results], first, they analyse what they answered, I’m talking about standardised tests, they analyse their answers and they talk as a group, like ‘you marked A, I marked B, let’s see what the mistake was’. I do that, they do the feedback and I serve as a mediator in that sense. I work a lot as a mediator, and I always do feedback after the assessments, always” (T07).

This feedback experience is in a group context rather than individual, by checking individual answers on a standardised test. The focus on mistakes identification seems to be also linked with

short comments that are mentioned only for higher grades: *“Look, they know, from a 6 to 6.5 I put a ‘Very good’, and from 6.5 to 7.0 I put ‘Congratulations’”* (T04). That type of practice seems to be reduced only for those obtaining higher grades. However, feedback is linked and reduced to checking test answers. The informal characteristics of the verbal feedback highlight the reduced spaces to do that activity but also the flexibility with the grades assigning process:

“When one, I don’t know, you say to the guys ‘you know like here, in this paragraph, I don’t know, it lacks coherence with the second or the third one, read it, we can comment, we can check’. I always said to my students that if the grade needs to be changed, we change it, but we need to reach an agreement... But as I said... the sad part is that there is not always space for that” (T10).

The situation radically changes when rubrics are used: students obtain scores that are related to specific performance assessment criteria that they know beforehand, in a process that is sometimes collaborative and also labelled as meaningful to everyone:

“We created the rubric all together, and it worked really well because in depth is like, all of us agreed: these are the criteria that you gave to me, therefore, you are the ones that complied or did not comply with what you decided. They knew because they said to themselves ‘These are going to be our criteria; we are engaged with accomplishing these criteria’. There were groups that, I remember there were groups like really in line with the criteria that they created. But some people didn’t have that order, so they did other criteria... But it was funny at the end because we had feedback from that activity that allowed them to say ‘let’s see, these are the criteria that we decided ourselves, but we didn’t accomplish that, it was difficult to get that one’” (T03).

Once again, if it is true that rubrics guide students’ performance and help them to identify their achievement in each criterion, and in that sense, that could be considered as relevant feedback, even when this is always points and grades related. As the use of rubrics that assess performance (and usually group) activities helps students to think about their learning, generally speaking, the assessment process also helps teachers to reflect on their practice:

“And then they [the students] shed light about what you are doing not so well maybe and... although it is important to reach good enough trust levels in order to them telling you ‘teacher, you know what, this content wasn’t worked well enough’, or

'teacher, you know, that question was super confusing, I understood this'. Thus, they will be helping you to improve your assessment instruments for the future" (T02).

However, teachers' self-reflection is, once again, heavily impacted upon by worries about time constraints:

"What am I doing wrong? Where do I need to put more emphasis? I don't know, how can I accomplish... what can I do in my class to get all the students to learn... but also... the time versus this, this, this idea of wanting all the students to learn is complex to harmonise" (T01).

Again, the curriculum coverage intention as a time-consuming prime task appears to be impacting upon the feedback provision, deeming it to be a "waste of time": *"I think that teachers were really closed, even accepting the idea that feedback was a waste of time, that you didn't cover the curriculum, because your time wasn't enough for the feedback" (T07).*

The impact of curriculum contents as a theme can be also seen here as a reflection of feedback as also more easily related to them than abilities or the learning itself. However, some teachers reflected on the relevance of feedback beyond grades provision, as it motivates students' learning:

"Umm... students do care about us umm... like that we notice their individuality. The students are not interested only in getting one grade, but to know that you really read their work, that you give the time, to write little things, to comment. Then, when you give it to them: 'hey, your work was good!', I mean, the student knows that his/her individuality is being considered. Thus, for the next, for the next assessment, they will do more, because they know that they are not one more in the group... that you are really looking at what each one of them is doing in particular" (T02).

This reflection is related to the common practice where students only receive a number, or grade, as a result of their work. However, it also shows evidence of a system where ideas beyond grading practices can also be found. In that sense, the demands of the grading processes involved in assessment seem to be acting against feedback provision, reducing the space and time for its development. For students, when they obtain feedback it is usually in a private form, with comments that sometimes show them how to improve but depending on each teacher, as they tend to highlight only errors:

"I think that they mark more what you have done wrong... for instance, if you get a 7 [maximum grade] they write 'very good', 'congratulations'. When you, I don't know,

get a worse grade, they write a full paragraph 'there was a lack of cohesion, a lack of coherence' but they don't say 'you could do this'" (S01).

The assessment method seems to influence the feedback given, as it is more common for written assignments than for test formats (that are the most common assessment method used). In any case, some students feel that the feedback is negative and especially targeted towards low-performance students and errors:

"I think that those with the worst performance, it is like they got more comments, even it's like 'just put in more effort, you can do this'... but for the ones with the best performance, it is not like I feel superior, but they didn't say almost anything to us, it is like they assume that we will have a good performance" (S02).

Therefore, it seems like a contradictory situation, where good students feel that only low-performance students get feedback, and on the other hand, low-performance students feel that only good students get what is viewed as good comments on their performance. This seems to be prompted by the feedback characteristics: focus on error without improvement guidelines. Assessment results have been also labelled by students as delayed, despite the school decree requirements about timely communication of results. As these are verbal or written comments on the paper, parents could read them or not: the official results are always only grade reports. Here again, the effects of the summative culture of the assessment practice are noticed, creating spaces where words and dialogue are not a common element but rather numbers.

Despite being identified as an uncommon practice; students know that the information contained in the feedback is important for their learning: *"it needs to be known where one is mistaken, in order to not make the same mistake again" (S06), or "it helps me to improve my work, to have good work and a goal to accomplish" (S05).* This represents a clear formative aspect of feedback identified by students.

The lack of time and class size act again as a constraining factor for feedback provision. Overall, students stated that what little feedback they received was at the discretion of each teacher and could be verbal or written, which is less common as they usually receive only a few short comments about their work. As such, comments are usually related to mistake identification, in feedback that is given mainly privately but often only means a numerical grade, proper of the summative culture of the country. Thus, feedback is not an extended or formal practice and seems to be insufficient to support students' learning improvement. The following chapter will analyse

these features in light of the fairness of the SBA in the school, once the main features of the system, in terms of purposes, methods, and information provided by the assessment have been revealed.

Chapter 7: Assessment and Fairness

“Assessment procedures are aimed at an improvement not only for the results but also for the rationality and fairness of the educational practices”.
Santos Guerra, 2003.

7.1 Introduction

Chapter 6 presented findings on assessment information in terms of both grades and feedback provided to students. It highlighted the scarce information and feedback that students receive from the assessment procedures to improve their learning and again, found that the prevailing system is mainly focused on grade production placing feedback on a secondary plane. This chapter will focus on findings on RQ3: What are teachers’ and students’ perspectives about the fairness of the assessment system in one Chilean secondary school? This research question opened up a space for a critical approach where reality can be criticised to bring more justice (Merriam and Tisdell, 2016), as was stated in the research design chapter. It will also allow consideration of that reality which needs not only to be described but also transformed (Cannella and Lincoln, 2015). Thus, this chapter focuses on the theme “The Need for changes towards a fairer assessment system”. However, findings in this chapter will be inevitably influenced by two external themes identified by the analysis: social inequality and teachers’ lack of time and work overload in the country. In this way, the overall fairness of the system as per participants’ perspectives will be presented in relation to the social inequality of the country and the high stakes of the system. Here, the impact of assessment results on students’ future possibilities within a society where inequalities are marked are explored. Following the tone of the international literature on assessment fairness, this section juxtaposes other interconnected topics, such as the possibilities for formative assessment and the diversity of the classroom. Secondly, following the need for changes theme, the profound need for a culturally responsive assessment system is presented, highlighting the relevance of the cultural context, which has been a main topic of this study. Finally, the fairness of the structures under which teachers work and the impact of their time on implementing assessment procedures are presented.

7.2 Social inequality and the impact of the high stakes of the educational system

The major and overwhelming theme for this chapter is that all of the twenty participants in this study, teachers and students, identified different elements that allow them to state that the assessment system is not fair. Following the tendency of the literature shown in Chapter 2, they made links between fairness and different elements of school life. However, some clear patterns were found within the data set which are linked with elements that are already presented in previous chapters. Fairness is mainly associated with the high stakes of grading and its impact on university entrance by the use of NEM, which will be further analysed in the following paragraph. This is framed by segregation and inequality within the Chilean society and educational system, impacting SBA. The social context is key to understanding how SBA is lived inside the school:

“No, it’s not a fair system, is not a fair system because it wants to assess all the students thinking that they are all the same, and nobody is like the other one. I think that this has relation to the gaps, with, with all the awfulness that the system has, not only educationally but the system itself that rules our country, and our daily lives” (T09).

The deep inequalities within society are also exemplified by a teacher who makes explicit differentiation between districts in the country and their respective possibilities:

“It depends on the reality of each school, I mean, a school in La Pintana who has a vulnerability rate of, I don’t know, 90% versus a school in Las Condes which is private, there if the assessment of these two students is contrasted, it’s not fair for any of them” (T01).

Teachers showed clear consciousness about the knowledge and abilities gap generated by a segregated society and students’ assessment as being a difficult reality to overcome.

“It’s not fair, how can it be fair for children that have disabilities, fewer competencies, fewer abilities..., the culture, of course, here we have children that don’t care, and obviously I can’t demand much from them, I mean... you can demand only within certain ranges” (T08).

Students’ groups evidence differences that are difficult to cover when assessing them. Making things worse, the system assesses 45 students within a classroom in the same way without respecting differences is also highlighted as unfair by teachers: *“It’s not fair, no, no... because it is centred towards certain abilities developed within certain types of assessments, it is not diverse, it*

doesn't attend to diversity and it's not integral" (T02). This reflection acknowledges students' differences but does not include the fact that in the country, teachers also decide the instruments that will be used to assess their students and therefore, are suitable for change. In the same sense, they exhibited the difficulties of creating fair assessment instruments as not one size fits all:

"I think it's unfair, precisely because of that, because it's hard to accomplish an instrument wide and diverse enough to assess all of the students' abilities. In the end, I think that one takes that option that it is, we create an assessment in this way, that one in another way, or this way, to diversify... but an instrument is never going to be fair enough as an instrument" (T03).

As has been stated in this study, the curriculum is assigned by Mineduc's policies as something that must be complied with, but the assessment strategies are left for every school and every teacher to decide upon, following their school assessment policies. During the COVID-19 pandemic, Mineduc (2020b) published a document about grading and promotion that required assessments "to be fair and of quality" (SD02, p.5). The same document asked the school community to "only use summative [assessment] when it guarantees quality, fairness and learning opportunities to all students" (SD02, p. 7). Even when this Mineduc document mentioned the concept of fairness, assessment decisions were once again, in teachers' hands. The internal documents produced because of the pandemic (PD10, 11, and 12) did not include any concept of fairness on them. For students, to continue assessing them in the same way even when they have low performance, with no changes, is not fair:

"No, it's not fair because they know that there are students that don't perform as well as the good ones, they keep doing that, like in a certain way forcing them to learn in that way, to be like the others, like...umm, I don't know, he got a 7 and keep doing the same test and keep going and going, and you got a 2, a 3 and nothing changes for you" (S04).

Despite is expected that some students will perform better than others, this speaks to an inflexible system without space for changes or strategies to support low-performing students, even when those strategies and methods (as described in Chapter 5) are actually in teachers' hands and were especially requested during the pandemic with fairness objectives (SD02). Linking back to Chapter 5 on assessment purposes and methods, students feel not being assessed on their real capabilities and they highlighted the abuse of the memory, especially for the extended use of tests. Accordingly, teachers face groups of 45 students without enough time or space to tackle any

learning gaps identified. The grade production within that space also appears to be a constraint factor. In that sense, the school seems to be in a place where inequality is perpetuated rather than diminished or dissolved and this is accepted with conformity by teachers: *“The grades topic is really heavy... it is like negative in reality, for the guys, but sadly it is our system, and we have to... it’s the game’s rules and we have to simply follow”* (T06).

Again, teachers seem not to acknowledge the freedom and responsibility that they have to make decisions about the assessment methods used as a way to improve learning and results or make them more meaningful and representative of students learning. They see the school structures as part of a system that is not possible to change but simply to follow.

7.3 A socio-culturally responsive assessment: the fairness of asking teachers for formative and summative practices

My first impression after analysing the data set, as has been stated, is that the language ‘formative’ and ‘summative’ is not common or extended either for teachers or students, following the tendency of other national research described in Chapter 3. Besides, formative assessment methods are generally linked to grades, responding to a strong summative and grading culture. This invites the question about the fairness of asking teachers, here and elsewhere, to implement policies based on international ideas that keep promoting both concepts and methods, within school spaces impacted by social inequality and teachers’ lack of time and work overload. Teachers simply talk, in a holistic way, about ‘assessment’ identifying the practice usually with grading processes, even when they include formative strategies. Formative assessment methods that are constantly and systematically found in the school are those which are graded. Formative strategies, such as learning objectives sharing, conversations on learning progression, or self and peer assessment are less common as they are not necessarily associated with grades nor needed to grade. When the formative and summative features of SBA are considered, teachers feel a contradiction between the two:

“It’s a huge contradiction, I feel all the time within that contradiction, and for the same reason, I try to modify it, to reflect on my own assessment and sometimes I realise like ‘oh, this assessment guideline is awful’, or ‘these criteria are really wrong, they don’t measure what they are intended to measure or ‘it’s unfair’. Some students make accomplishments in other ways and you keep reflecting and modifying as well.

Some people never do that and they keep repeating the same instrument year after year without any modification, without reflecting about the process” (T03).

For these teachers, the formative aspect focuses spaces for reflection about the assessment process quality and fairness and about being ‘really’ formative but always related to grades: *“I feel that that is the way, to make a processual and really formative work, and that all was reflected in the grade, which is, for me in my case, is the most difficult in assessment terms” (T05).*

In any case, the concept is not widely used by them, and the ‘process’ seems to relate to assessment that is done during the time in an accumulative way more than in any particular instance, such as in tests. Thus, formative means accumulative grades where teachers can monitor the students’ progress and also the use of points-based rubrics that always end up generating another grade that is registered in the schoolbooks. In this way and as has been stated, formative is summative. This could be corroborated when in September 2020 the school decided to include a ‘formative grade’ (as registered in the school documents PD11 and 12) in the students’ final results. However, the focus of the strategy is the points assignation rather than learning improvement, as there is no time to do so: if a student is having difficulties with learning, they just keep moving on with the curriculum contents.

Teachers were asked about the possibilities presented by Decree N° 67, the new initiative designed to keep promoting the use of formative assessment, to have fairness in the assessment. However, they kept the reflection about grading needs and how difficult it could be to make the changes proposed by the decree: *“I feel that this modification is like a modification in the form, but doesn’t change the focus of education, so I feel that can be even worse [...] I don’t know if our system is prepared for such change” (T03).* This is because the summative requirements remain the same, within a summative and grading culture where numbers matter, as was explained in Chapter 5 but where the repeating students' practice should be avoided, changing the rules of promotion, which can be resisted by teachers as some of them think that students ‘deserve’ to repeat, as explained previously. Therefore, there exists a ‘repeating culture’ that focuses on grades, and within the high stakes, has enormous relevance for students.

Questions made in the interview about the new assessment decree did not make teachers use the concept of ‘formative’ extensively. It is important to say that the decree did not modify grading or certification requirements for the schools so that processes seem to keep teachers doing what they have been doing always. The difficulties of making changes to have a fairer

system are identified as a cultural process that needs to be done rather than a change that has already happened:

“I think that we are still in that process, in this process like, cultural. First, we need to get certain, certain paradigms out of our heads and then open our minds, once we open our minds as teachers, we will start making deep and important changes. But I think that we are not there yet, we are not culturally prepared” (T07).

Changing paradigms could be a relevant aspect of cultural changes that would support a fairer system, as teachers’ mind-sets seem to be closed as they are used to the traditional practice that they have been doing for years. Therefore, even when the new assessment decree pretends to promote formative assessment, the educational system where the change must take place is still labelled as being archaic, revealing teachers’ hopes and aspirations:

“We come back to the question of what have we formed [as students], what has been taught to us, or educational system... clearly umm, and educationally, it doesn’t point to the century where we are in, I mean it’s extremely archaic... I hope that from now on, with this thing of the changes and the assessment decrees, we can advance further, I mean that the student starts to realize once and for all, and teachers as well, that, there will be more formative spaces, than summative. But I think it’s a risk because as I was telling you, we have a student formed that maybe will need years or simply months or can be understand quickly this, this formative instance, I mean that the student can take off this concept like ‘it’s graded, I will do it’” (T10).

Teachers identify the challenge of changing old patterns where the grade has been historically the system’s main objective and motivator and keeps being. However, as has been seen, the combination of formative requirements with summative ones makes teachers concentrate on the latter, without open spaces for formative or where formative is only a strategy to keep grading students maybe just with lower frequency. In that sense, formative for them is more a focus on the process of assigning grades rather than a process to analyse and improve learning.

Only one teacher referred to feedback as part of the formative process, although she recognised that it is a practice that teachers do not use properly. This goes in line with the findings about the lack of quality feedback presented in the previous chapter.

Interestingly, one teacher reflected on the cultural appropriateness of formative assessment within the school: *“I’ve researched a bit, in other countries, how the assessment*

process is done... is way more formative, that works there, but would it work here, I ask? Thus, that is, that is the problematic" (T08).

This teacher identifies that different countries face different challenges to implement formative, questioning the appropriacy of the strategy in the country. Especially when, as was shown in Chapters 5 and 6, formative is always linked to grades, and where if there is no grade, there is no students' work. This also exhibits that different cultures have different opportunities to successfully implement formative assessment practices.

This section, following the tone of the international literature on assessment fairness detailed in Chapter 2, has set out different ideas around the concept. Fairness for teachers and students is related to the students' possibilities to obtain learning, especially when the inequality of the country is considered, and also to have a real space to show that learning is taking place. The sense of fairness is also related to the changes proposed by the Ministry of Education within a culture where structures and summative requirements have not changed significantly. Therefore, formative seems to be a foreign idea that is suggested without careful consideration of the systemic conditions where the educational process takes place within the schools. Teachers' ideas about formative assessment are impacted by their lack of time and also the collision of different intentions of each assessment purpose, especially the summative demands. This is especially true when the cultural meaning historically impregnated within the 'assessment' concept is considered and even more when the stakes of the system are considered. Inevitably, fairness ends up being related to the impact of grades on university entrance, which has been explored in Chapter 5. As the consequences of the assessment are maintained and linked to grades, grades are the focus of the assessment dynamic inside the classroom as the only language that their actors fully understand and apply to their daily practices. Finally, the assessment culture seems to be resistant to change easily and without further intervention with regard to the systemic characteristics impacting upon the process that were not changed by the last change in the assessment decree made on 2020. Teachers seem not to know how to implement assessment procedures in other ways, and their practice is highly influenced by the high stakes, therefore they guide the process rather than formative objectives formulated by Mineduc in the national assessment policy. It leaves open the question about the fairness of asking teachers to implement policies that bring ideas from other contexts and countries without the guidelines, information, or adequate structural conditions for their success within high stakes that determine students' futures and possibilities.

7.4 A Structural condition of teachers' lack of time and work overload: the constrained freedom

The findings have been presented considering that Chilean teachers have, in theory, pedagogical freedom to make several decisions about planning, teaching, and assessment. Previous chapters have shown that they can make decisions about teaching and assessment. In practice, these decisions are mainly guided by the lack of time that made them opt for practical choices, usually 'quick and easy', more than what can be pedagogically better and more meaningful for their students' learning. In this way, teachers' freedom is constrained by structural conditions imposed on the learning process, especially by the lack of time prompted by the multiple tasks that teachers must perform and the work overload that they have, as will be detailed here. This aspect of the Chilean educational system was highlighted by both teachers and students. Students are fully aware of the work overload of their teachers and understand that they do not have time for simple activities such as talking with their students:

"In my case, there was no chance [to talk with them] because the teachers that I could talk to, they didn't have time because they were full of, shifts, there they had a class, after that class another one, they had to have lunch..." (S04).

Students were also aware that the lack of time impacted upon their grades and feedback provision: *"Usually they delayed [the results] a lot, but I mean, it's understandable, there are 45 students, there are two classes per level... and it's understandable"* (S07). Students seem to understand the high demand that teachers have, which in any case invites us again to the question about the fairness of this conformity: teachers cannot be there, for them. Even when students identify the lack of time impacting upon teachers' performance, to them the matter is more complex and is inextricably linked to ethical aspects and the fairness of their working conditions:

"I think that what happens is that ethically my bigger conflict is related to that, that teachers don't get to be the real channelling force of a given student's abilities or... or learning mediators because we don't have the time needed to do that. I think that it has to be with time... you have achievement percentages, that you can't only have I don't know, 20% of your students failing, that there has to be one assessment per month, that you have to give the assessment results in two weeks... so everything is related to external rhythms that are imposed, independent of your subject, of your

weekly hours, of the nature of what you are teaching... Thus umm... everything is so artificial, so structured, so mechanised, and you have to adapt yourself to that... I feel that, umm, I mean we should have to reset the complete system, starting from pre-service teachers' education in the university" (T02).

As can be seen, the various demands that teaching and especially the assessment procedure imply are seen as following external schedules imposed by the school that ignores several aspects of the nature of teaching inside the classroom, which is unfair and creates a sense of ethical conflict for teachers. Therefore, the systemic nature of the space where learning is taking place negatively affects its possibilities and also evidences that teachers' freedom is hugely constrained by the educational structures. Once again, this invites to question if it is fair to ask teachers to implement assessment reforms and changes when they do not have time and must respond to several administrative requirements. This also invites a more rigorous and critical questioning of similar international experiences of assessment reforms.

Time constraints also are impacted by the class size and the demands of these numerous groups of students, which in the case of assessment promote an overwhelming amount of work for teachers if tests and their frequency are considered, as has been detailed in Chapters 6 and 7:

"Understanding that the human resources are not enough... not enough, and that, we have to be really aware that even if you have got a strong will, a lot of times it can be... It's only a few classes that have less than 40 students... and the work requires time and time... that one, sadly, we don't have" (T05).

When the educational system is viewed with the fairness lens, several changes are proposed by its actors. For students, time is a crucial fact by the simple reason that it can help teachers to teach everyone: *"The time... because sometimes there is no time to explain to everyone in detail"* (S08). Therefore, it is interesting to know how students solve their questions if teachers do not have time, along with the fairness of the situation. That answer also commonly appeared in the interviews: students recur to their classmates more than ask their teachers directly, showing a sociocultural sense of the learning process through peer mentoring and learning. Students seem to help each other considering and respecting that they understand things in different ways, and that way is not always the one used by the teacher:

"I have had classmates that I've helped in Maths, and they do well with how I explained to them. So, you need to find the way in which that person understands... [Female classmate name] I have taught her, she got a 5 in Maths with a lot of effort,

and it was like two days for the semester test where I helped her, she got a 6.1, I mean, it's a lot of difference, only because I explained to her as I knew she would understand" (S01).

This kind of personalised explanation seems to be neglected by their teachers because of the lack of time, adding a sense of unfairness to the students' experience. Unfairly, students help each other and seek support from their classmates when needed, because the teacher, who is supposed to do that work, does not have time:

"I ask [male classmate name], or some classmate that had understood, I ask, I don't know, did you understand this? Can you explain it to me? And we start like playing, for example, let's see, maybe it is in this way, and this is the one I understood, this is correct..." (S01).

This capability of explaining the ideas collaboratively learnt to other classmates is also identified by students as 'real learning' instead of merely memorizing contents. However, students also reflect on why help between classmates is needed when that task should be done by the teacher:

"You excuse yourself by saying 'it's because I don't understand the teacher', that 'the teacher doesn't know how to teach', but maybe we don't do our part, and neither do we look for help in another place, although I know that shouldn't be because the teacher is in the classroom to teach you, but I think that is a mutual responsibility, teachers' and students'. If the teacher sees that one student is not understanding, or is not going with the rest of the class, they should say how can I teach you? What don't you understand? How can I help?" (S03).

Teachers' lack of time seems to create a shared sense of which teaching actions should be developed, which in practice, are difficult because of the constrained opportunities to do it, such as monitoring their students' understanding in a classroom where usually they have 45 students. This questions about the fairness of teachers' possibilities to teach in such complex conditions.

The activities that teachers have also make them reflect on the need for spaces to work collaboratively, for instance in school meetings, to have a common line with colleagues, and even for essential tasks such as talking about what they are understanding by education:

"I think that there is a lack of space for teachers to talk about education, what vision each of us has for it because that is the thing, we can unify criteria but maybe your vision about education doesn't go in that line, do you see? I think that those spaces

are absent like to talk, to dialogue, to discuss, to be asked about such things, what do you think of the assessment, why do you grade so much..." (T09).

As can be seen, the lack of time has an impact, not only on pedagogical decisions that teachers make, but also it negatively impacts upon their possibilities to work collaboratively with their colleagues. This was especially highlighted by teachers as they did not have time even to talk with their subject colleagues about crucial decisions on the teaching provision. When they have time to work collaboratively with colleagues, the situation is promoted by personal interest and organisation, several times sacrificing free time rather than the school giving them time to do so. For students, there is conformity about this lack of time, generating different strategies to cover key tasks that teachers had not had time to develop. Once again, the fairness of asking teachers to perform assessment practices through a new decree, knowing that they do not have time to implement it can be questioned.

It is worth mentioning that despite teachers' lack of time, they seem to be aware of the emotional support that their students need and provide by close, meaningful, and empathetic relationships. In that sense, several interviewed teachers considered their students as "human beings" in an integral way, namely, individuals with emotions, problems, and concerns. However, teachers felt that this clashes with the constant summative assessment conducted and the numerical reporting of grades, where the human quality of the student is absent. Within the classroom space, and defying the lack of time, conversation seems to be an important part of the educational process analysis and adjustment. This conversation is linked with formative aspects, as well as the focus on the long-term educational process rather than on grades' production through a test, which are relevant elements of learning promotion. Therefore, formative is happening, only embedded under a strong summative culture, and taken away from the assessment process itself to become rather a process that is done in other spaces of interaction between teacher and students (after or outside the assessment instances). Finally, it is worth saying that teachers and students do not blame each other for the educational conditions that they face: they just keep going, following the educational structures with conformity and trying to help each other as much as they can while navigating the rules and the stakes of the school system.

Chapter 8: Discussion

*"I want simple things for you, that you draw a world of justice and joy.
And I know that outside all shine. But if you look, you will see the lie".*
Ana Tijoux, Chilean singer.

8.1 Introduction

This discussion chapter will be organised into three main sections that will describe an assessment system that from the outside seems to perfectly work but from the inside has several disconnections and contradictions. The first section will discuss results on RQ1: What are the SBA purposes and methods used in one secondary school in Chile and what are teachers' and students' perspectives about them? A second section discusses results on RQ2: What are teachers' and students' perspectives about the information provided by the assessment, in terms of grades and feedback, in one Chilean secondary school? Finally, there is a discussion on the results of RQ3: What aspects of school assessment are linked by teachers and students to the fairness concept in one Chilean secondary school? The discussion will be followed by the final Chapter, 9, with conclusions, limitations, and some recommendations for future research on SBA.

8.2 Addressing RQ1 on assessment purposes and methods: policy and practice disconnection on SBA due to lack of attention to the school conditions and needs.

A main feature of the findings on RQ1 is that there is a disconnect between what is stated in policy and school documents and what is done in practice by teachers inside the school, for both, assessment purposes and assessment methods. The new assessment policy gives insufficient attention to school-level conditions and needs at different levels. This aspect of the policy was an unexpected finding, as any of the research questions specifically directed to it. However, it is understandable as the policy acts as a regulatory structure under which the assessment occurs. As Smith et al (2009) stated, the introduction of new ideas is allowed in qualitative research to enrich the discussion. Therefore, the concept of "policy translation" (Maroy & Pons, 2021) will be brought to the fore. Briefly, the authors explain that policy must be shaped to the scales of where it will be situated but also to the contexts of the actor who enacts it, especially when they are based on ideas that circulate at a transnational scale (Maroy & Pons, 2021). The "policy translation" concept

seems useful to explain why the assessment policy exhibits a disconnection from what is found in the practice in the school: it has not been translated considering the (socio-cultural) reality where should be implemented. Firstly, assessment purposes and methods are deeply influenced by teachers' lack of time as a structural condition that has been historically present in their practice and has an influence on teachers' pedagogical choices and practices (Martínez, 2019; Torres-Olave, 2022). Secondly, the policy has presented formative assessment ideas within a school context where historically curriculum and grades have configured the dynamic not only for teachers and students but also for the school community. In this way, teachers' and students' perspectives around SBA tell a story of persistent cultural scripts. Following Elwood and Murphy (2015), this study has understood an assessment script as beliefs about what assessment is and is used for, impacted by cultural and social mediation. In the school, assessment is made by following the curriculum and assessed by a test to obtain a grade. Therefore, the policy requirements proposed by Mineduc do not include a proper consideration and balance of the school conditions needed to break a well-known and dominant script, to be replaced by another one with completely new and sometimes opposite ideas and practices, such as the ones that formative assessment purposes. This is because the new policy has not been able to change what is understood by assessment and how is used for. This section will present two further subsections to discuss the differences this exerts on assessment purposes (RQ1a) and assessment methods (RQ1b).

8.2.1 RQ1a. What are the assessment purposes used by teachers following the school's documents and teachers' and students' perspectives?

In theory, the assessment purposes in the country should be guided by formative ideas promoted by the Ministry of Education. However, agreeing with international experiences on SBA and Chilean studies on assessment discussed previously, results indicate that curriculum is the main philosophy of work for teachers (Gysling, 2017; Hernández-Nodarse et al., 2018; Martínez, 2019; Mujica-Johnson, 2020; Rasooli et al., 2023; Sánchez & Jara). This could be explained for several reasons:

1. The implementation of formative assessment is relatively new if compared to the work with curriculum contents. If we go back to the cultural scripts idea (Elwood & Murphy, 2015), this focus on curriculum contents could be explained because

assessment purpose is embedded with epistemological beliefs: to know (and learn) is to manage the curriculum contents, this is, information. Learning is focused on contents rather than abilities: the object (curriculum) rather than the subject (student), and this is the dominant script. Therefore, an object, the curriculum sets the purpose of the assessment, the learning focus, and how time is used inside the classroom. The focus is not on the human aspect but rather on the technical aspects of teaching. Contents mean knowledge that is valued in society and also allows students to succeed in the assessment methods that they have, which are mainly tests. Tests also have high stakes for students, as they will have an impact on their possibilities to access university.

2. The same curriculum acts against formative strategies such as coming back to one specific content if something is not well understood: you must simply advance covering contents, there is no time to monitor learning and return to reinforce things. Here, memory appears as a key element for allowing the management of those curriculum contents. The need to cover the curriculum (because this is “to know”) is perceived as a huge pressure that restrains teachers and students from reflecting on their learning because there is no time to do so, or reduced spaces where teachers attempt to reflect on learning with their students. In any case, the speed of the learning process is set by curriculum contents and the high stakes of the system.
3. Related to the formative and summative purposes usually present in SBA, in the school they have a complex interaction, as has been previously reported in the literature (Black & Wiliam, 2009; Earle, 2014; Flitcroft & Woods, 2018; Irving et al., 2011; Johnston et al., 2019; Mäkipää & Ouakrim-Soivio, 2019; Remesal, 2011; Yates & Johnston, 2018). SBA features in the school agree with sufficiently reported conflicting and mixed messages received by schools where they are demanded to provide numerical results but also formative elements of assessment (Arzola-Franco, 2017; Santos Guerra, 1988; Vera et al., 2017). Hernández-Nodarse (2017) highlights in this regard that the evaluative practices have not been organised and systematised to be coherent, efficient, and formative. Thus, the ‘formative’ is not extended into practice. The difficulties and tensions of implementing formative assessment in summative-dominated spaces are not acknowledged by the Chilean policy, leaving schools unaware, and ultimately powerless but also responsible to

act through their school assessments. The internal school decree tends to imitate the spirit of the Mineduc decree, without being able to consider the complex interaction of internal and external elements that constrain the possibilities of SBA.

All these elements talk about a different culture with characteristics that are not considered when expecting to implement policies that are not socio-culturally responsive as brought from different Western countries. Proof of that is that the documents created by MINEDUC to promote formative assessment are highly based on European literature, with no “policy translation” efforts made (Maroy and Pons, 2021). In that sense, the policy diminishes the possibility for Chilean assessment to be “authentic”, as demanded by Chilean researchers for a long time, as it finds itself constantly imitating and transplanting foreign ideas (Condemarín & Medina, 2000). To be authentic is a relevant element to be socio-culturally responsive. The Latin American literature presented in Chapter 3 exhibits that foreign ideas have promoted ‘pathologies’ of the assessment, in the sense of Santos Guerra (1988; 1993, 1996) that have not been considered or solved when they have been historically denounced in the zone. Therefore, both in Latin America and in Chile we have felt uncomfortable with some assessment ideas, and that could explain why the socio-culturally responsive assessment made sense for this study. Teachers’ resistance to changing their evaluation practices because reforms do not recognise the specific context has been also highlighted in Latin America by authors such as Pizarro and Gómez (2008) and Nodarse et al (2018). Chilean students’ resistance to change as they are used to certain evaluative practices, as stated by Martínez (2019) could be also focused on what is regarded as being important by the educational system. This resistance is also explained by this study as that context involves structural conditions that are not acknowledged by assessment policies, such as teachers’ lack of time and/or students’ concerns about their future because of the high stakes of the system. Thus, the intentions of this formative assessment could be seen, following Mitchell (2017), as foreign criteria that cannot happen in a specific context (in this case, Latin America). However, this also reinforces the notion that Western approaches cannot be always taken as the norm (Kennedy et al., 2008). These ideas combined, call for an urgent socio-culturally responsive approach in the sense that assessment could not be a universal recipe to be copy pasted to every country in the world. Because the assessment process must be designed (and translated) always acknowledging the diversity of countries, societies, communities, cultures, and human beings where it is applied.

8.2.2 RQ1b What are the assessment methods used, following school documents and teachers' and students' perspectives?

Here once again, the policy and practice disconnection is evidenced, as the school assessment decree promotes the use of a variety of assessment methods while teachers use more frequently traditional tests and formats, such as multiple-choice worksheets. This is because teachers have scarce time to respond to specific summative requirements for promotion and certification of learning. The analysis of assessment methods used by teachers (RQ1b) as per the schoolbooks gave an overview of the rigidity of strategies used within the secondary school level and the guidelines (or confines) under which they had been created, which is regarded by the literature as impacting mostly in a negative way on students' results as they cannot fully show their variegated abilities (Bonner et al., 2018; Bota & Tulbure, 2015; Mäkipää & Ouakrim-Sovio, 2019; Puad & Ashton, 2021; Smyth & Banks, 2012). This goes against the central idea of SBA allowing for a wider range and variety of tasks that should be integrated during the learning process (ARG, 2002; Yates & Johnston, 2018). Agreeing with Bota and Tulbure's study in Romanian schools (2015), Chilean students already identified alternative assessment methods (other than tests) as an opportunity to have better results and allow them to show different abilities, such as attitudes, personal traits, and behaviour. However, it has been shown that abilities are not a central purpose/focus of the assessment process: the curriculum is. The assessment methods included summative use of formative:

1. An example of this is rubrics. As rubrics are used to grade, they become another summative assessment method, where students can check if they have complied with assignment requirements and its points-associated, which duly generates a grade. This cannot be underestimated; the summative cultural script learnt by this community is so strong, that formative when it appears, is to be converted to summative. That does not mean that the use of this set of criteria, with descriptions of the level of performance as defined by Brookhart (2013) is not supporting learning; but only that as suggested by Pope et al., (2009) students could tend to focus on pass (and the grade) rather than on learning improvement. In this way, the language in the rubric guiding the assessed activity takes second place because finally, the result is a grade.
2. Another formative aspect used to grade was corroborated by the accumulative grade used by Chilean teachers in the school: they labelled as formative the use of several activities (worksheet type) that receive points which later are converted to a grade, in a process occurring in several classes more than just once assessment instance. Even when teachers can support and

provide formative feedback on these activities, formal feedback is not detailed beyond points or a grade assignment. It could be, only if the teacher marks/reviews the work in the classroom, accompanied by oral explanations and comments about the students' work and advice for improvement. As such, there is a cultural appropriation of the formative concepts that seem not to follow the international tendencies shown in Chapter 2: Chilean teachers adapt and attribute distinct meanings to formative, which respond to the local summative cultural reality. Formative "made in Chile".

The incongruence is not only between policy and practice but also, as stated by studies such as James and Pedder (2006) and Monteiro et al. (2021), to the extent of teachers' perspectives and their real practice. Even when they criticise these formats, they assess their students mostly using them within a space where teachers and students seem to accept that with the historical norm with generalised conformity. It is what has always been done by everyone. The ethical standpoint of this choice could be also criticised: they know that the test is not the best format for students. They provide little information about their learning and results are not always the best, but they keep massively using them because are quicker/easier. Once again, this can be framed within the structural conditions where they do their work and their historical lack of time to do it.

Despite the intentions of the MINEDUC decree and the school decree, assessment methods evidenced a lack of variety and the use of standardised test formats (SIMCE, PAES). PAES formats are used, however, to help students to be prepared to take the university entrance test used in the country, evidencing a social impact recognised by the assessment and in that way, being responsive to the students' needs. This is especially relevant when the inequality of the country and the social context of the district and the students of the school are considered.

Following Elwood and Murphy (2015), the artefacts constructed by teachers to assess their students follow here, inevitably, the impact of society and culture, and could be added to its pressures. What is knowledge and learning here? Knowledge is the information you can put into a test to later, forget. What is learning in this format? Learning occurs when you obtain high grades/scores. And those scores could lead you to university. What is valued by society as a dominant culture is information management and high grades. Grades will be further explored in the following section on the information provided by the assessment.

8.3 Addressing RQ2 on the information provided by assessment: numbers, humans, creativity, and words.

Discussion on RQ2 What are teachers' and students' perspectives about the information provided by the assessment, in terms of their grades and feedback, in one Chilean secondary school? This will be separated into two different sections, to cover the results associated with the summative aspect in terms of grades, and in the formative aspect in terms of feedback provided to students. Once again, assessment information is impacted by Chilean teachers' lack of time as a structural condition, especially non-lecture time in their contracts highlighted also by Chilean researchers such as Gysling (2017) and Martínez (2019). The high stakes of the system have a relevant impact on how the results are focused: namely the grades are all that matter. In combination with teachers' lack of time, they make possibilities for formative feedback to happen difficult. With this, it can be also said that policy and practice disconnection also reaches the assessment results, as the policy promotes feedback that is scarcely found in the practice. However, teachers use creative strategies for advancing students' longer-term interests within social structures and elements larger than the school that have an impact on their decisions on assessment. These topics will be presented in the following three sections.

8.3.1 SBA and grades: The Power of numbers vs. formative spaces

The summative culture lets be reminded here, must necessarily be understood then as a result of learnt and shared behaviour, values, customs, and beliefs (Frierson et al., 2010) found within the school. A policy cannot simply change that overnight, even less so when the sociocultural context and structural conditions where education takes place are not duly taken into account. For all the RQs of this study, it has been shown that the summative culture is the dominant "cultural script" (Elwood and Murphy, 2015). Numbers are at the core of the assessment process and hold sway over teachers and students as once they are assigned, they have an impact that goes beyond the classroom. And this is because they are embedded in a system with social consequences attached to them. Numbers in Chile matter, as grades decide not only when students are promoted from one level to another and if they obtain a certificate from secondary school (MINEDUC, 2018a) but also determine their future possibilities, a relevant factor highlighted by studies such as Sivenbring (2019) and Smyth and Banks (2012). This is even more

relevant for students from medium and low socio-economic backgrounds, such as the ones participating in this study.

As has been corroborated by other studies worldwide, rather than a source of guidance about their learning, in this study grades are a source of negative consequences, such as encouraging high achievers and discouraging and overwhelming low achievers (Harris & Brown, 2009; Sivenbring, 2019). Grades also impact students' self-esteem (Irving et al., 2011). This study has also found that teachers treat differently and stereotype good and bad achievers, using grades in a punitive way, used to control and exert power over students for being lazy, sometimes deserving to repeat the academic year. These ideas were found also in studies such as Gysling, 2017, Londoño-Restrepo, 2015 and Sivenbring, 2019. However, in this school, these persistent ideas are still present even when the new assessment decree tried to reduce the instances of this happening. This could be explained because the formative aspects of grading are separated from the process because of a lack of consideration of the school's reality and needs. Including these needs is the representation of grades on NEM, a way in which school grades will have an impact on university entrance: within these high-stakes structures it is difficult to focus on something different, or new, such as formative assessment. Besides, there are no formal requirements for it.

8.3.2 The creativity of grades manipulation

The impact of assessment in terms of inequality and social justice makes us recognise that without justice, assessment is helping us to reproduce inequalities (Elwood, 2013; Murillo & Hidalgo, 2015). And those abound in Chile, as was stated in Chapter 3. It is really interesting how teachers create strategies attempting to help their students succeed in their studies despite their cultural capital. This is because as Smyth and Banks (2012) suggested, they are aware of the effects of grading on low-income students and the effects of a lower cultural capital on working-class young people and their academic results. In this way, elements such as socioeconomic factors and even cultural backgrounds are integrated into the grading process. Within the school setting, all students have the opportunity to access tenths and extra points for class activities, for instance. However, some other opportunities are given only to good-performance students and also, it is more likely that best-performance students will comply with activities that allow them to increase grades. Therefore, through these activities, both teachers and students make efforts to balance the resulting grades by using a raft of different strategies that elsewhere might be labelled

as 'score pollution'. Hence, grades as products of the assessment scores obtained by the students, are not objective but rather an emergent social creation. Following Elwood and Murphy (2015), it is important to ethically understand the impact of this practice in terms of what is rigorous and valued in the educational system. And if we ask what is assessment here? It is a set of strategies used for grade production. What is used for? Certification, promotion, and university entrance. Then I will help you with that. Santos Guerra in 1988 already identified that evaluation is a moral phenomenon that must consider who is serving whom. Grades are, therefore, mediated and agreed upon rather than an objective numerical indicator that cannot be changed. Grades must serve students. This agrees with a study in Iran by Rasooli et al (2023), where teachers also compensated students from disadvantaged backgrounds by adjusting grades. Therefore, all the objectivity that this summative culture has from the outside is manipulated from the inside, leaving (at least) two reflections:

1. 'Score pollution' as labelled in the literature on assessment (Green et al., 2007; Martínez, 2019; Peters & Contreras, 2018; Pope et al., 2009; Remesal, 2011), was an interesting way in which teachers help their students improve their grades and, in this way, navigate the rules of the high stakes system while also minimising its consequences within a neoliberal context and creating a fairer school assessment environment. In that sense, grades are used to do 'justice' by helping students by increasing or inflating them. Score pollution here defies what can be labelled as good or bad, and it is open to questions of being right or wrong to students (Estaji, 2011; Pope et al., 2009), as an example of practice ethically rooted. This reality requires, as suggested by Brookheart (2013), a different understanding of the meanings and purpose of the grading strategies used by teachers within the school. Following Colnerud (2006) score pollution could be a way in which teachers balance between justice and care, showing ethical standpoints when teaching. The practice also highlights the relevance, as stated by Contreras and Prieto (2008) of understanding the contentions and intentions of teachers' practices concerning the context where they are being used. In any case, the Chilean case score pollution at the secondary level is a strategy that helps students to navigate the rules of university entrance and improve their opportunities to get an offer, especially if their socioeconomic backgrounds are being considered. For these reasons, this practice is an example of a socio-culturally responsive assessment.
2. The inclusion of social behaviours such as effort and attitude into the grading procedures highlights the complex nature of grading, which has been stressed by authors such as Stiggins and Conklin since 1992, also with more recent warnings on awarding points beyond the learning

environment (Shepard, 2019). Even when the lack of validity is problematized in SBA for the impact of grades on university entrance (Tong & Adamson, 2015), here they are used to helping students in a way that defies the concept 'validity' itself. Teachers are aware of the profound social inequalities rooted in Chilean society and how this duly affects their students. Therefore, the 'score pollution' attempts to mitigate those differences by helping students who show great levels of effort and commitment to their studies: grades are a way to do justice. In doing so, it is shown how teachers balance institutional versus students' needs (Hirsh, 2020; Rasooli et al., 2023), deciding how to help their students best.

Following the idea from Roth and Radford (2011), teachers and students modify what is understood as achievement and success, producing, reproducing, and transforming their interaction through assessment, and this practice is an example of that. In any case, as reinforced by Green et al (2007) this section highlights the need for a more consistent sense of ethical assessment practices. Moreover, a profound need for teachers to receive explicit training on fairness in assessment, as stated by Rasooli et al (2023). One that could balance different teachers' and students' needs from a point of view of doing good and being aware of unfairness prompted by different elements of the assessment system, but knowing how to face it.

8.3.3 SBA and feedback: numbers vs. words

What opportunity could be within an educational environment where grades have such a high impact on students' lives for feedback? Not a big one. What space could a teacher have with no time to provide feedback for numerous classes where all the formal requirements are placed only on grading? Not very much. Formative feedback in the school is occasional and non-systematised. The focus is placed on grades and measurement. Agreeing with Santos Guerra (2003) and Gysling (2017), what is important in the school is to pass, rather than to learn. This fact invites us to question the real possibilities of an assessment system, such as the SBA, of being formative if it does not imply constant or quality feedback. It should be said that the system cannot possibly be SBA. Thus, this study cannot sufficiently detail the feedback characteristics, in terms of being task/process/person related, private/public, timely/delayed, positive/negative, or oral/written as they are described in studies presented in Chapter 2, such as Zhao et al (2017).

In the school, comments associated with feedback are reduced to short positive words that are not really useful for students' progression. Students deal by themselves with their grades and

their meaning individually, without having sufficient written information about how to keep learning or improving, limiting the information obtained to right and correct answers and a grade. This prevents students from making corrections and thus it is ineffective for their learning, contrary to the spirit of their formative assessment (Black & Wiliam, 1998; Hattie & Timperley, 2007). Consequently, a key element of formative assessment is not systematised, as feedback is neglected by the teachers' lack of time and the summative culture that contains elements that are not changed either by the national policy or by the school assessment decree. There are no formal requirements for feedback provision or guidelines, leaving feedback in an informal conversation place with no evidence for students or parents to check for later reflection or work. As stated by Florez Petour (2014), grades are the only language between them. To change that way of communication requires more than the mere intention of a policy: to talk with words and not only with numbers. Mainly because the long-term impact of psychometrics on assessment cannot be overestimated. And this is a contradiction in society, as one could ask why precisely the school is the place where the language and the script is focused on numbers and no words, even when the real interaction between students and teachers and even parents, and teaching itself, is constantly a world of words.

8.4 Addressing RQ3. What aspects of the school assessment are linked by teachers and students with the fairness concept in one Chilean secondary school? The overwhelming sense of unfairness

In the Chilean context, Gysling (2017) has highlighted that fairness of the assessment is a key element to focus on within the assessment process. This idea was included in this study, showing that the main feature of RQ3 was the sense of unfairness about the assessment process in the school that both teachers and students shared. Their perspectives on unfairness were based on the need for changes to different elements of school life for assessment purposes, methods, and results (RQ1 and RQ2). In this sense, they were linked also to structural conditions found in the country, such as social inequality, school segregation, and teachers' lack of time and work overload situation. When these features are looked at in the light of the new policy requirement, the statement that is not fair to request changes that could not happen because of the existing conditions is more than evident. This is because, as stated by Rasooli et al. (2023) in Iran, what is fair is shaped by the social structures of the school, which will be also the case for Chile. Here, the

relevance of considering assessment as a social product appears in the sense that values and culture are important to define what matters in assessment in specific contexts (Broadfoot 1999; Elwood & Murphy, 2015). In the case of Chile, the socioeconomic context is key to understanding how the assessment process is lived in the country. Therefore, fairness will be discussed following a socio-culturally responsive approach presented in section 2.3, integrating theoretical elements of this study with the inductive process of analysis into this discussion.

8.4.1 An unfair society, an unfair school

Chile was labelled in this study as an example of both school segregation and educational inequality (Bellei & Vanni, 2015; García-Huidobro, 2007; Gysling, 2017; Oliva, 2008; Oliva & Gascón, 2016; Puga, 2016; Zancajo, 2019). Both features have been corroborated by the participants of this study as elements that carry important consequences for the school life. Thus, any feature of fairness will necessarily go across school segregation and educational inequality. In this study, teachers and students were raised, educated, and/or part of a district (San Bernardo) that embodies the contradictions of the country in terms of poverty, crime, and unequal income distribution and opportunities. This impacted shaping the acute sense of unfairness for both, teachers and students as they knew beforehand, as has been stated by studies such as Cabalin (2012) and Espinoza and González (2015), that they would have lower results and opportunities to access university because of their socioeconomic status. Both the purposes of the assessment and practices used are dictated and influenced by cultural pressures from where they are created and where meaning is acquired (Elwood & Muphy; 2015; Kennedy et al., 2008).

1. Neither teachers nor students identify as unfair that assessment purposes are guided by curriculum contents, thus relegating formative opportunities to the background and prompting a focus on students' abilities. The epistemological tone of the assessment purpose, i.e., to know is 'to manage contents', is not even regarded as unfair, even when it takes away the focus from the students, their learning, needs, and rhythms.

2. Regarding assessment methods, the experience of numerous classes where teachers assess all students in the same way is labelled as unfair, as low achievers are left behind when a single test format must fit everyone, which questions its moral appropriateness (Beets, 2012; Harrison et al., 2020). Teachers know this, but they do not have the time to create more effective instruments for their students. However, there is a small number of experiences where teachers

used other assessment strategies. This demonstrates that formative evaluation is possible, although this typically takes place under constraining rather than enabling conditions and within summative confines. Students also know this and are aware of the unfairness of the process, as even when the test routine is not effective, teachers keep doing it. If the “what students need” idea is recalled, it could be argued that the assessment should always be adapted to the students, to their characteristics and needs, a guideline that lies at the very core of SBA.

Consequently, the assessment practice inside Chilean schools could not be labelled as a socio-culturally responsive assessment (Brown et al.,2022), as it is not good for all students and it is not responsive in allowing students to show their accrued knowledge in different ways. Therefore, only some of them will succeed, and consequently, only some of them will have good grades and opportunities to compete to access university. In this way, as Villalobos and Quaresma state (2015), inequalities are then reproduced rather than attended with scarce spaces to acknowledge the need for ethical discussions about the competence logic and market orientation of the educational system. Even when the PAES test has had changes in the last three years following national demands, the score gap has remained stable, especially for those from the wealthiest backgrounds and private schools obtaining higher results (DEMRE, 2023).

8.4.3 A socio-culturally responsive assessment made in Chile

What would it take for the assessment system in the country to be fair and socio-culturally responsive? Here, the concept of fairness is linked back to the idea of “what students need” (Brown-Jeffrey & Cooper, 2011), following the culture where the assessment is conducted (Burns, 2019).

- *The assessment policy must acknowledge the school's reality and students' needs.* This refers to the need for policy translation processes (Maroy & Pons, 2021), and to consider the complex interaction of the local context and transnational policies such as SBA. Local actors are key: head school principals, academic coordinators, teachers, students, and parents. These are the specialists that can help to adapt the policy and connect it to specific realities and, because of that, make it applicable and effective. Thus, formative assessment has a real possibility to be integrated into SBA in a meaningful way, thus improving students' learning and promoting good teaching practices.

- *Teachers need time to adapt the assessment following their students' best interests.* The assessment policy must acknowledge the structural social conditions where the educational process is taking place in order to mitigate or reduce its impact. These are teachers' lack of time, work overload, and social inequality. The educational process begins with inequalities from the start, which are not attributable to the teacher nor to the student or the school. Therefore, schools and teachers need time to design and diversify their strategies depending on the context of the school and their students, recognising that some of them will need more resources and support than others. This planning is not possible without time. Time is also connected to the assessment decree, as it should protect the elements relevant to learning improvement (e.g., feedback provision) by formally describing and requiring a register of its characteristics.

- *Students need an assessment process based on their learning as human beings.* The wide curriculum contents that guide class planning and its deep impact on the learning process should be analysed, since it takes away the focus from students' abilities to manage information. This has other consequences associated, such as grades based on that management information, and the fact that students do not work if the activities are not graded. Changing that culture requires modifications in the legal and external structures that rule the assessment culture inside the school: assessment decrees, plans and programs, internal decrees, and strategies used. By extension, these changes could transform how the school communicates with families. Therefore, they imply a new language that can be found beyond numbers and data, and attempts to implement and/or recover the human and formative aspects of the educational process. Therefore, these changes depend on the national policy level (Mineduc).

- *The assessment policy and internal assessment decree must include ethical guidelines.* Because the focus of assessment is students' learning and what is better for them following their needs, ethical guidelines should help teachers support their students on how to respond to the diverse situations and demands generated within the educational space. While some teachers are attempting to help their students, in other settings any strategy is used. However, this should not be left as a personal dilemma to be solved by each teacher; clear guidelines should help teachers guide the learning process with fair values and spaces to discuss their ethical standpoints and how to respond to the ethical demands of the assessment system. The discussion will be followed by a conclusion chapter, implications, and recommendations for research on assessment.

Chapter 9: Conclusion

9.1 Introduction

After discussing the results of this study concerning the three research questions that guided this thesis and making several important connections to previous research on SBA, this final chapter will provide a general conclusion for the study, highlighting its contribution to knowledge, the future implications for policy on assessment, schools, and research on the area, along with an acknowledgment of its limitations, a reflective account of the research process and some recommendations for future research on the topic.

9.2 General conclusion

This doctoral thesis aimed to understand the SBA process at the secondary school level in one school in Chile, by exploring school-produced documents, such as schoolbooks, decrees, and staff meeting presentations, and mainly, teachers' and students' perspectives by the use of semi-structured interviews. The focus of the study was placed on four vital elements of the assessment process: purposes, methods, information provided, and fairness. From the findings of this research, it can be stated that SBA exhibits a disconnect between national policy and local assessment practices in the one school examined, which is impacted by external factors that have been historically present in the Chilean educational setting. In this sense, the intentions of the policy are not found practised in the school reality, where despite the promotion of both, formative and summative purposes, the school is focused on the latter. This obeys the historical summative and grading culture found in the country which is characteristic to Latin American and also in Asian countries where grades and measurement have a key role. This summative culture assessment is also impacted by teachers' lack of time and work overload and by the acute social inequality found in the country. Combined, they configure a SBA approach that is not such, as it is mainly focused on summative assessment, leaving the formative as a limited possibility for being focused on grades production, but also featured as a "cultural illusion" (Elwood & Murphy, 2015) which responds to the social reality where it is supposed to happen. This is because the assessment system presents mainly negative features that affect and even prevent its execution: curriculum's content coverage purposes, classic assessment methods (impacted by teachers lack

of time), a summative culture and grading system where feedback barely exists, and finally, a sense of the unfairness of a system that is assessing young people within high stakes places in a country with one of the worst income distribution of the world (Zancajo, 2019). The impact of the socioeconomic reality where the school is placed transversely affects teachers' and students' perspectives alike, in a combination where they react against it but also display a level of conformity as something that has always been like that – that is to say, which they believe to be intractable. In that reality, the formative assessment concept seems foreign and does not fit with their actors' concerns and what is requested of them by the school system. In a country with specific sociocultural characteristics, such as neoliberal policies and social inequality, it is simply unfair to ask for the implementation of formative assessment as suggested by the policy, when summative action carries so many important consequences for schools, teachers, students, and their families. This is particularly telling when despite guidelines provided by Mineduc, schools continue associating formative assessment with grades. The strong summative and grading culture found in the school indicates that all the work that students do is graded or marked by points/tenths, otherwise, students do not complete the tasks. Besides, the assessment dynamic also suggests that grading is used as a strategy to control behaviour and a threat: if you want to pass, you need to work. Formative assessment is indeed seen as an aspect of good quality teaching, and in that sense is not unfair. However, its implementation needs to be protected by guidelines and requirements that allow formative to be in its real sense and not overshadowed by summative strategies.

Along the same line, it is unfair to ask teachers to implement policies that they do not have time to implement, as all the summative requirements that they have to comply with are constantly acting against their execution. Finally, this study shows that it is necessary to harmonise the language used in assessment, avoiding labels that can make difficult to understand what is going on inside a school. It is not respectful to insert foreign ideas within a particular context, ignoring the sociocultural reality in which teachers and students live the school's everyday reality, especially when that reality acts against the execution of changes. Those changes should always respond to local needs rather than international recipes about what is better in foreign realities.

9.3 Contributions to knowledge

This study aimed at better understanding and describing the SBA practices in one Chilean secondary school centring teachers' and students' voices, and placing an emphasis on assessment purposes and methods used, grading and feedback as information provided by the process, and, finally, the justice of the overall assessment system. Epistemologically, the study promotes a knowledge that is school-rooted and examines participants' perspectives about the assessment reality that teachers and students face in their day-to-day lives with an interpretivist approach.

Following the literature review gaps, a key contribution that this study makes is to advance previous knowledge by including students' voices and also considering low-income students from subsidised schools, which has generally been absent from research in Chile.

Additionally, this study suggests the profound need to include several SBA elements, beyond the classroom reality, such as family, students' future opportunities, policy creation, and socioeconomic reality. Only in this way, being respectful of the real context, assessment policies can have a real impact that can benefit what is lived inside the schools in Chile.

This study has also contributed innovatively by considering and reflecting on the sociocultural context in which SBA occurs, which in this case, has been done by balancing evidence presented in the international literature but acknowledging local literature, both Latin American and Chilean, that outlines differences that are worth to consider to better understand the assessment process in its real practice. This implied a profound respect and consideration that knowledge production in some places of Latin America does not necessarily follow the patterns of Western research on assessment. It also has involved not assuming that what is reported internationally must necessarily happen in Latin America or Chile, nor focusing on results as mistaken only because they do not fit with what is presented as occurring in other contexts.

The study has also advanced previous knowledge on SBA by opening a space to analyse the fairness of the assessment system and being critical about the changes needed to respond to students' needs. The contribution of a socio-culturally responsive assessment will be discussed from the policy implications perspective, as the elements associated with the discussion are mostly related to it, evidencing the relevance of the policy but also the complexity of its implementation.

9.3.1 The relevance of teachers' and students' voices within educational research

Methodologically, this study aimed to contribute to new knowledge by highlighting, as others have, the need to include teachers' but especially students' voices in educational research (Baird et al., 2014; Brown et al., 2009; Brookhart, 2001; Carles, 2005; Cheng et al., 2011; Marshall & Drummond, 2006; Monteiro et al., 2021; Sivenbring, 2019; Tapan, 2001; Torrance & Pryor, 2001; Willis, 2008). Advancing this knowledge, this study has highlighted low-income students from a district and country exhibiting high levels of inequality and segregation. Students should always be foregrounded as key participants of the assessment process rather than merely the recipients of others' decisions (Cheng et al., 2011). It is easy to think about how an educational system should be without being involved and engaged in one. As a teacher myself, I can say that happens more than it should within a field with experts who barely know the reality inside the school or a classroom. And, for that reason, policies on assessment end up being so distant from reality and thus undoable. People's perspectives are an irreplaceable link between policy and reality that can inform, rephrasing Mitchell (2017), about what is happening and could happen within an educational setting. They are also the reason that shapes every decision that configures the practice in the school (Pizarro & Gómez, 2018). Remesal (2011) highlighted the importance of studying these perspectives before even trying to change the summative practice into formative. This has not been sufficiently done in Chile and this study contributes to that new and much-needed knowledge, that could result in a key element of future policy implementations on assessment.

9.4 Implications

This section will present a number of implications of the study findings for research, policy and also for schools in their day to day activities.

9.4.1 Implications for research

Research on assessment is abundant worldwide. However, it lacks a consensus not only on vocabulary but also on key definitions, such as what will be understood as education or fair assessment. Research on SBA must include these definitions to have a common standpoint to progress on. Additionally, the field contains the historical difficulty of concepts such as summative

and formative, that cannot be avoided but can also obscure the discussion (Newton, 2007). Research must also acknowledge that there is no universal recipe for conducting good SBA. This study challenges the conventional logic of SBA by being respectful of other ways of seeing and conducting the assessment process and being honest about the conceptual difficulties. The concept of SBA itself is not widely used in many countries with assessment systems that classify as SBA, as they are focused on both, summative and formative aspects even when this is not explicitly stated, such as Scotland (Scottish Government, 2011). Therefore, it is important for research, as this study did, to revise how SBA is studied by assuming that assessment is a polysemous concept that carries a different conceptual basis (Maldonado-Fuentes, 2021). This could be linked to the idea of cultural scripts, that reveal different ideas about what is assessment in different sociocultural contexts (Elwood & Murphy, 2015). To understand this, research needs to include teachers' and students' voices on SBA research, as these are key to understanding what is happening within school systems from a sociocultural and respectful standpoint, creating socio-culturally responsive approaches to assessment (Brown et al., 2022).

9.4.2 Implications for Schools

Even if policies on assessment are not well located in the reality lived in the schools, they have possibilities of harmonising by making the most of their internal assessment decrees. Thus, by using the freedom that is allowed to them when creating their internal assessment decrees they can mitigate the elements where national policies are disconnected from reality. This can be especially done by including teachers' and students' participation and constructing their decrees with them. In this sense, schools can implement strategies to improve teachers' lack of time and to protect the formative intentions of monitoring students' learning to improve it. Or, as would be said in Latin America, opening and protecting spaces for the qualitative aspect to appear, the one that considers the human aspects of the educational process. This can only be done by using a more flexible approach to grading and by allowing time and space for creating assessment methods that suit the student's needs. Moreover, this goes for every school using SBA, as the features of the school studied agree with the Chilean literature on assessment and curriculum implementation, and therefore, with their difficulties and summative focus. In this way, it is a generalised suggestion as this reality could be found in other schools.

Time is a key resource for the complete assessment process, including real possibilities for feedback provision, and promoting a new way of communication that is not solely based on grades and SIMCE/PAES scores. To open and protect collaborative spaces to reflect on the fairness of the assessment system is also a main contribution of this study, as teachers need to balance the multiple demands that the educational process has keeping always in mind their and their students' wellness. However, the system should provide and protect the spaces to do so. Therefore, recommendations go down the line of focusing on students' abilities and reducing the summative requirements, as well as making formative requirements. This also must be reinforced by moderation processes that help ensure the quality and pertinence of the assessment methods used by teachers. However, the school also must work with the summative and grading culture and language, by communicating not only with students but with parents/carers in a new language not exclusively based on grades but on students' abilities, and progression and providing ideas about how to improve. In any case, improvements on SBA should always consider the process not only as a technical one but as a human process that must necessarily include an ethical aspect and the question about who is being good and adequate.

9.4.3 Implications for Policy

Even when this study focused on SBA practices at the school level, the way in which assessment should be planned to be socio-culturally responsive in Chile must consider policy guidelines. Therefore, a completely new challenge is highlighted by this study, which requires that the complexity of the assessment policy and implementation processes for SBA be reviewed. The evidence shows a strong impact between assessment policies and the assessment practices.

This study challenges evidence on SBA by presenting one Chilean secondary school as being impacted by copy/paste policies borrowed from Western countries, that even considering their benefits and possibilities, need to be adapted or following Maroy and Pons (2021), translated. In this study, this can be done by considering the sociocultural reality where they are intended to be applied (Elwood & Murphy, 2015), before their implementation occurs. Only in that way, ideas such as formative assessment can be more than just a "cultural illusion", as stated by Elwood and Murphy (2015) for countries such as Chile, but also others that face similar difficulties, such as Indonesia (Puad & Asthon, 2021), Iran (Rasooli et al., 2023), Finland (Mäkipää & Ouakrim-Soivio, 2019), and Hong Kong (Cheng, 2019). Although this appears to be a logical step, school systems

have the need but also the right to adopt assessment policies that respond to what is needed by students and respect not only their reality but also teachers'. This study highlights the profound need to consider and respect the sociocultural reality of the school without thinking that one strategy fits all. Moreover, policy makers should be prompted to consider that international trends in assessment may not yield the same results under different contextual circumstances

This study stresses a much-needed acknowledgment of historical external conditions, such as teachers' lack of time and work overload, imposed on teaching within schools which are not considered when proposing new policies on assessment for the educational community. Without this necessary step, the introduction of new policies will always exist on paper but will be difficult to implement for schools and the people in them, promoting a disconnection between policy and practice, as evidenced by this study. Following the experience described by studies in Peru and Mexico (Falcón-Troncoso et al., 2021; Monzón, 2015), the promotion of formative assessment in Chile has no possibilities for a real application.

Additionally, policymakers should always consider exploratory information about what is going on inside different types of schools (public, subsidised, private) before implementing changes. In the case of Chile, assessment policies need to take responsibility for the historical conditions where education has been taking place: neoliberal policies, social inequality, school segregation, and especially, teachers' lack of time and work overload. It is precisely these final elements that can be solved by the implementation of new policies that protect the time that teachers need to implement changes but mainly to conduct a teacher practice, understanding that assessment is one part of it, where they can think about what is better for their students' learning and, following Torres-Olave (2022), to reflect on their practice. Therefore, policy prescriptions must be based on studies that describe the local reality, considering the wide range of schools and students that the country has. In that sense, this study contributes by showing the reality of a subsidised school, where the educational provision of the country is concentrated, that embodies the social inequality and segregation of the country. These types of studies were not conducted before the creation and implementation of the new assessment policy by Mineduc.

Finally, policies should include ethical guidelines and compulsory training on SBA throughout the country and not just for public schools, assuring that schools' staff get in touch with assessment guidelines. Most important, the assessment policy must include systematised spaces for the formative assessment to systematically happen, which include increasing teachers' non-lecture time and changes in the summative requirements that obstruct the execution of real

formative assessment. Hence, the policy must not only be connected but derived from the sociocultural reality of the country, acknowledging it. It is not the school that must adapt to the policy. It is the policy that must be adapted to the school in the first place.

9.5 Limitations and Reflexive Account

A number of methodological, social, and personal limitations have been identified in this doctoral research. Reflecting on the thematic analysis conducted, my experience and the Chilean literature reviewed made it difficult to find positive aspects within the data set and present a nuanced vision of SBA in the school. However, this study intertwines with Chilean literature on SBA that shares negative aspects of the system and tends to highlight them, reinforcing them as the main features of the Chilean educational system. The Chilean educational system is an international example of a neoliberal educational system, carrying with it strong negative meanings.

Therefore, to ensure that findings will be truthful to the school's reality, despite the deductive approach used to work the literature review and research design of this study, the inductive approach taken for data analysis allowed me to be respectful of the findings, despite theoretical trends that could have forced the discussion towards elements that were not relevant for the school reality or for the study's participants.

Adding difficulty to the research process, the impact of COVID-19 cannot be dismissed. The data collection process of the study was planned for 2020, the year that school provision was moved indefinitely online. In Chile, that time was extended by two years. On the one hand, I used schoolbook registers from the previous and 'normal' academic year (2019), and on the other hand, the reality that teachers and students were living when they were interviewed could have also had an impact on the results. This is because several changes were implemented around SBA at that time because of the pandemic, generating difficulties and also reflections about grading reduction and formative strategies that could be useful in the pandemic times. Participants, both teachers and students, were stressed and scared but provided useful reflections on the educational process which, under these circumstances, invited them to think about what was important. This was extended to what was important about assessing, grading, learning, and teaching. Moreover, in that sense, it could have a positive impact on the depth of the reflections provided by the study's

participants. The focus on what is essential for SBA, following teachers' and students' voices is a main contribution of this study.

Hence, the absence of parents' voices in the interviews must be acknowledged as well. Even when the assessment process is lived directly by teachers and students, how the process is seen, lived, and informed by parents represents an interesting focus of attention as they are part of the sociocultural context. This is especially true when the social inequality of the country is considered and foregrounded. It would be relevant to know how they live the grading system, the consequences for the students' progression and future, and also what it means for them attempting to access higher education and to pay for its costs. It also is important to consider how the summative culture and the grades communication can provide more useful information to parents about the learning process of the student. This is key for parents/carers to be able to support the learning process and be involved in it.

At a socio-political level, in addition to the difficulties imposed by COVID-19, Chile was facing an unprecedented social uprising in democratic times that started in October 2019. I travelled to Chile at the end of that month and conducted the preliminary study amid the social uprising where people went to the streets to claim better life conditions, education was one of the topics. I lived through the restrictions, protests, and violence. It has been difficult to calculate to what point the country's situation impacted my data set, analysis, and/or results. This is because the social uprising challenged key elements of the assessment process, such as NEM, ranking, and especially PAES. Grading processes were also challenged by the COVID situation during 2020. Teachers and students were fully aware of the tiredness of the social inequality and school segregation consequences by the time I conducted the interviews. However, they and I have always been aware of that reality and the sense of unfairness of our society and country.

The country had not recovered from the social uprising that brought freedom restrictions and even deaths when the pandemic started and the classes were suspended to be online for almost two years. We also had a presidential election where in December 2021, elected the youngest left-wing president in the world. After that, we had a constituent convention, which was in charge of writing a new constitution to replace the one written in the dictatorship. The proposed text was rejected in September 2022. A second proposed text was once again rejected in December 2023. Looking back, I think that all those complex and sometimes painful social events had an impact on this thesis, especially when thinking about the fairness of the assessment system

within the educational environment and social conditions that were unique and unrepeatable. I only had one 'normal' year as a PhD student and Chilean... then all of us became just survivors.

Finally, I want to acknowledge that my in-between positionality could have limited how this doctoral thesis has been conducted and written. This is because I had difficulty being able to neutrally balance the information learnt through these years of intense reading and research with the powerful knowledge coming from the school and the classroom that I have in me as a teacher. I struggled with the sense of honesty about what it is really like in schools in my country and district, and also struggled with the responsibility of telling this story carrying with me all the learning obtained from years of practice in different schools. I have been teaching in the last three years in different secondary schools in Chile, therefore I keep witnessing, questioning, living, and struggling with the school reality with no intermediaries. However, that is who I am as a teacher-researcher.

9.6 Recommendations for future research on SBA

Future research on SBA should retain a focus on teachers' perspectives without excluding those of students' and parents'/carers' voices to have a more holistic view of the assessment process. This is, to consider the educational process as part of a sociocultural space where different elements have an impact on how assessment is conducted in the school. Research considering this sociocultural approach, as suggested by Elwood and Murphy (2015) should be carried out to improve our understanding of how the assessment, their actors, and the structures where they interact are impacted by the sociocultural reality where education takes place and where assessment meanings are negotiated. Additionally, future research on SBA should be also undertaken by including the ethical and human aspects of SBA beyond technical concepts and recipes about what is better, namely, focusing on who is better and why. Ethical discussions around SBA are still focused on several aspects of the process that show evidence of the profound need to unite visions and widen research in the area. This necessarily involves research on teachers' training on ethical aspects of the assessment process, which is still unknown and not focused on research on SBA.

References

- Acción Educar (2021). Aporte de la educación particular subvencionada en Chile. <https://accioneducar.cl/wp-content/uploads/2021/09/Aporte-de-la-educacion-particular-subvencionada-en-Chile-Accion-Educar.pdf>
- Adamson, B. (2011). Embedding Assessment for Learning. In R. Berry & B. Adamson (Eds.) *Assessment reform in education: policy and practice*, pp.197-203. Dordrecht: Springer.
- Ahmadi Safa, M., & Sheykhholmoluki, H. (2023). An impact study of the Iranian National University Entrance Exam from students and parents' perspectives. *Language Testing in Asia*, 13(1), 40.
- Arnold, J. (2022). Prioritising students in Assessment for Learning: A scoping review of research on students' classroom experience. *Review of Education*, 10(3), e3366.
- Arzola-Franco, D. M. (2017). Evaluación, pruebas estandarizadas y procesos formativos: experiencias en escuelas secundarias del norte de México. *Educación*, 26(50), 28-46.
- Arzola, S., & Collarte, C. (2009). Sustentabilidad del cambio: las prácticas docentes. *Pensamiento Educativo, Revista de Investigación Latinoamericana (PEL)*, 44(1), 99-117.
- Assessment Reform Group. (ARG). (2002). Assessment for learning: 10 principles. *Research-based principles to guide classroom practice*.
- Assessment Reform Group. (ARG). (2006). The role of teachers in the assessment of learning.
- Atkins, L., & Wallace, S. (2012). *Qualitative research in education*. SAGE.
- Avendaño C., William R., & Parada-Trujillo, Abad E. (2012). EL MAPA COGNITIVO EN LOS PROCESOS DE EVALUACIÓN DEL APRENDIZAJE. *Investigación y Desarrollo*, 20(2), 334-365.
- Ávila, M., & Paredes, I. (2015). La evaluación del aprendizaje en el marco del currículo por competencias. *Omnia*, 21(1), 52-65.
- Baird, J., Hopfenbeck, T. N., Newton, P., Stobart, G., & Steen-Utheim, A. T. (2014). State of the field review: Assessment and learning. *Oslo: Norwegian Knowledge Centre for Education*.
- Baker, S. E., Edwards, R., & Doidge, M. (2012). How many qualitative interviews is enough?: Expert voices and early career reflections on sampling and cases in qualitative research. *National Centre for Research Methods*.
- Ball, D. L., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of teacher education*, 60(5), 497-511.
- Barnes, G. J. (2021). Researcher positionality: the liquid inbetweenener. *Practice*, 1-8.
- Bassey, M. (1999). Case study research in educational settings. *McGraw-Hill Education (UK)*.
- Bazeley, P. (2009). Analysing qualitative data: More than 'identifying themes'. *Malaysian journal of qualitative research*, 2(2), 6-22.
- Beaumont, C., O'Doherty, M., & Shannon, L. (2011). Reconceptualising assessment feedback: a key to improving student learning?. *Studies in higher education*, 36(6), 671-687.
- Beets, P. A. (2012). Strengthening morality and ethics in educational assessment through ubuntu in South Africa. *Educational Philosophy and Theory*, 44(sup2), 68-83.
- Bellei, C., Muñoz, G., Pérez, L. and Raczynski, D. (2003) Effective schools serving the poorest: Who said it is not possible?, in: Universidad Diego Portales (ed.) *The Education in Chile Today*, pp. 347-373.
- Bellei, C., & Vanni, X. (2015). Chile: The evolution of educational policy, 1980-2014. *Education in South America*, 1980, 179.
- Bennett, R. E. (2023). Toward a theory of socioculturally responsive assessment. *Educational Assessment*, 28(2), 83-104.

- Bernstein, B. (2001). From pedagogies to knowledges. In A. Morais, I. Neves, B. Davies, & H. Daniels (Eds). *Towards a sociology of pedagogy. The contribution of Basil Bernstein to research* (pp. 363-368).
- Biblioteca del Congreso Nacional (BCN). (2021). Reporte estadístico San Bernardo; Spanish for National Congress Library, statistical report San Bernardo.
https://www.bcn.cl/siit/reportescomunales/comunas_v.html?anno=2021&idcom=13401
- Biblioteca del Congreso Nacional (BCN). (2023)., Reporte estadístico San Bernardo; Spanish for National Congress Library, statistical report San Bernardo.
https://www.bcn.cl/siit/reportescomunales/comunas_v.html?anno=2023&idcom=13401
- Biesta, G. J. (2004). Education, accountability, and the ethical demand: Can the democratic potential of accountability be regained? *Educational theory*, 54(3), 233-250.
- Biesta, G. (2009). Good education in an age of measurement: On the need to reconnect with the question of purpose in education. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1), 33-46.
- Biggs, J. (1998). Assessment and classroom learning: A role for summative assessment? *Assessment in Education: Principles, Policy & Practice*, 5(1), 103-110.
- Black, P. (1998). Formative assessment: Raising standards inside the classroom. *School Science Review*, 80(291), 39-46.
- Black, P., Harrison, C., Hodgen, J., Marshall, B., & Serret, N. (2010). Validity in teachers' summative assessments. *Assessment in Education: Principles, Policy & Practice*, 17(2), 215-232.
- Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice*, 5(1), 7-74.
- Black, P., & Wiliam, D. (2003). 'In praise of educational research': Formative assessment. *British educational research journal*, 29(5), 623-637.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability (formerly: Journal of Personnel Evaluation in Education)*, 21(1), 5.
- Bloom, B., Hastings, T., & Madaus, G. (1971). Formative and summative evaluation of student learning.
- Bonner, S. M., Torres Rivera, C., & Chen, P. P. (2018). Standards and assessment: coherence from the teacher's perspective. *Educational Assessment, Evaluation and Accountability*, 30, 71-92.
- Boote, D. N., & Beile, P. (2005). Scholars before researchers: On the centrality of the dissertation literature review in research preparation. *Educational researcher*, 34(6), 3-15.
- Bota, O. A., & Tulbure, C. (2015). Aspects regarding the relationship between teaching styles and school results. *Procedia-Social and Behavioral Sciences*, 203, 285-290.
- Boud, D. (2000). Sustainable assessment: rethinking assessment for the learning society. *Studies in continuing education*, 22(2), 151-167.
- Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative research journal*.
- Box, C., Skoog, G., & Dabbs, J. M. (2015). A case study of teacher personal practice assessment theories and complexities of implementing formative assessment. *American Educational Research Journal*, 52, 956-983.
- Brannick, T., & Coghlan, D. (2007). In defense of being "native": The case for insider academic research. *Organizational research methods*, 10(1), 59-74.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101.
- Braun, V., & Clarke, V. (2012). Thematic analysis. *American Psychological Association*.

- Briggs, J., & Ellis, V. (2013) Assessment for learning, on Learning and teaching in secondary schools. Learning matters. SAGE.
- Brill, F., & Twist, L. (2013). Where have all the levels gone. *The Importance of Shared Understanding of Assessment at a Time of Major Policy Change. (NFER Thinks: What the Evidence Tells Us)*.
- British Educational Research Association. BERA (2018). Ethical guidelines for educational research. <https://www.bera.ac.uk/wp-content/uploads/2018/06/BERA-Ethical-Guidelines-for-Educational-Research-4thEdn-2018.pdf?noredirect=1>
- Broadfoot, P. (1999) Preface. In A. Filer (Ed.), Assessment: Social practice and social product (pp.ix-xii). *Routledge Falmer*.
- Brookhart, S. M. (2001) Successful students' formative and summative uses of assessment information. *Assessment in Education: Principles, Policy & Practice*, 8 (2), 153–169.
- Brookhart, S. M. (2013). How to create and use rubrics for formative assessment and grading. *Ascd*.
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., ... & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, 86(4), 803-848.
- Brooks, R., Te Riele, K., & Maguire, M. (2014). Ethics and education research. SAGE.
- Brooks, V. (2002). *Assessment in secondary schools: the new teacher's guide to monitoring, assessment, recording, reporting, and accountability*. McGraw-Hill Education.
- Brown, G. T. (2006). Teachers' conceptions of assessment: Validation of an abridged version. *Psychological reports*, 99(1), 166-170.
- Brown, G. T. (2011). Self-regulation of assessment beliefs and attitudes: a review of the Students' Conceptions of Assessment inventory. *Educational Psychology*, 31(6), 731-748.
- Brown, G. T. L. & Hirschfeld, G. H. F. (2008). Students' conceptions of assessment: links to outcomes. *Assessment in Education: Principles, Policy and Practice*, 1(15), 3-17.
- Brown, G. T., Hui, S. K., Flora, W. M., & Kennedy, K. J. (2011). Teachers' conceptions of assessment in Chinese contexts: A tripartite model of accountability, improvement, and irrelevance. *International Journal of Educational Research*, 50(5-6), 307-320.
- Brown, G., Kennedy, K., Fok, P., Chan, J & Yu, W. (2009) Assessment for Student Improvement: Understanding Hong Kong Teachers' Conceptions and Practices of Assessment. *Assessment in Education: Principles, Policy & Practice*, 16 (3), 347–363.
- Brown, M., Bruns, D., Mcnamara, G. y O'Hara, J. (2022). Culturally responsive classroom-based assessment: a case study of five secondary schools in Ireland. *Revista de Investigación Educativa*, 40(1), 15-32.
- Bruner, J. (1996). *The culture of education*. Harvard, MA: Harvard University Press.
- Bruner, J. & Elacqua, G. (2003). Capital humano en Chile. *Santiago: Escuela de Gobierno de la Universidad Adolfo Ibáñez, La Araucana y Perchede*.
- Burnard, P., Gill, P., Stewart, K., Treasure, E., & Chadwick, B. (2008). Analysing and presenting qualitative data. *British dental journal*, 204(8), 429-432.
- Burns, D., Brown, M., McNamara, G., O'Hara, J., Altrichter, H., Nayir, F., Nortvedt, G., Fellner, M., Helm, C., Punzenberger, B., Skedsmo, G. and Wiese, E.F. (2019) Aiding Culturally Responsive Assessment. A User-Friendly Handbook for Culturally Responsive Assessment in Educational Settings. https://www.researchgate.net/publication/333508260_A_User-Friendly_Handbook_for_Culturally_Responsive_Assessment_in_Educational_Settings
- Cabalin, C. (2012). Neoliberal education and student movements in Chile: Inequalities and malaise. *Policy Futures in Education*, 10(2), 219-228.

- Cabral, T. C., & Baldino, R. R. (2019). The credit system and the summative assessment splitting moment. *Educational Studies in Mathematics*, 102(2), 275-288.
- Cannella, G. S., & Lincoln, Y. S. (2016). Deploying qualitative methods for critical social purposes. In *Critical Qualitative Inquiry* (pp. 243-264). Routledge.
- Carless, D. (2005) Prospects for the Implementation of Assessment for Learning. *Assessment in Education Principles Policy & Practice*, 12 (1), 39–54.
- Castillo Retamal, F., Mejías Rodríguez, M., & Vásquez-Gómez, J. (2023). Conceptos, estrategias y aplicación de la evaluación formativa en tiempos de pandemia: un caso de la realidad chilena. *Meta: Avaliação*, 15(48).
- Cejas, C., & Alvarez, P. (2006). Evaluación de los resultados del aprendizaje. *Revista Argentina de radiología*, 70(2), 149-155.
- Centro de Investigación Avanzada en Educación. (CIAE). (2021) El 20% de los profesores se retira en los primeros 5 años de vida laboral y las especialidades más críticas son Educación Parvularia y Media. *Comunicaciones CIAE*.
https://www.ciae.uchile.cl/index.php?page=view_noticias&id=2206&langSite=es
- Centro de Investigación Periodística. (CIPER). (2014). Colegios subvencionados: así operan los siete grupos de “megasostenedores” que lideran el negocio.
<https://www.ciperchile.cl/2014/06/06/colegios-subvencionados-asi-operan-los-siete-grupos-de-megasostenedores-que-lideran-el-negocio/>
- Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas. (CPEIP). Ministerio de Educación. (2019). Incremento del tiempo no lectivo. <https://www.cpeip.cl/wp-content/uploads/2019/05/Orientaciones-Horas-No-Lectivas.pdf>
- Centro de Perfeccionamiento, Experimentación e Investigaciones Pedagógicas. (CPEIP). Ministerio de Educación. (2023). Valores 2023 de BRP y asignaciones de la Carrera Docente.
<https://cpeip.cl/asignaciones-carrera-docente-2023/>
- Chan, J. K. S. (2007). We have various forms of assessments but only summative assessments count: Case studies of the implementation of an innovative assessment policy in Hong Kong. *Paper presented at the Redesigning Pedagogy: Culture, Knowledge & Understanding Conference*.
- Cheng, A. (2019) School-based Assessment and Assessment for Learning: An Exploratory Study of Hong Kong English Teachers’ Practices, *THE JOURNAL OF ASIA*, 16, (4), 1368-1375.
- Cheng, L., Andrews, S., & Yu, Y. (2011). Impact and consequences of school-based assessment (SBA): Students’ and parents’ views of SBA in Hong Kong. *Language Testing*, 28(2), 221-249.
- Chimento, A., Rahman, A., Machin, L., & Frith, L. (2018). Mediated research encounters: Methodological considerations in cross-language qualitative interviews. *Qualitative Research*, 18(6), 604-622.
- Cizek, G. J. (1996). Learning, achievement, and assessment: Constructs at a crossroads. In *Handbook of classroom assessment* (pp. 1-32). Academic Press.
- Clarke, V., & Braun, V. (2013). Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *The psychologist*, 26 (2), 120-123.
- Cohen, L., Manion, L., Morrison, K., & Wyse, D. (2010). *A guide to teaching practice*. Routledge.
- Colnerud, G. (2006). Teacher ethics as a research problem: Syntheses achieved and new issues. *Teachers and Teaching: theory and practice*, 12(3), 365-385.
- Condemarín, M., & Medina, A. (2000). Evaluación de los aprendizajes. *Santiago de Chile: Mineduc*.
- Contreras Guajardo, D., Hojman Trujillo, D., Huneeus Lagos, F., & Landerretche Moreno, Ó. (2011). El Lucro en la Educación Escolar. Evidencia y Desafíos Regulatorios. *Universidad de Chile*.
- Cook-Sather, A. (2006). Sound, presence, and power: “Student voice” in educational research and reform. *Curriculum inquiry*, 36(4), 359-390.

- Córdoba, M. (2009). Evaluación vs calificación. *Innovación y experiencias educativas*, 1-10.
- Cormier, G. (2018). The language variable in educational research: An exploration of researcher positionality, translation, and interpretation. *International Journal of Research & Method in Education*, 41(3), 328-341.
- Couture, A. L., Zaidi, A. U., & Maticka-Tyndale, E. (2012). Reflexive accounts: An intersectional approach to exploring the fluidity of insider/outsider status and the researcher's impact on culturally sensitive post-positivist qualitative research. *Qualitative Sociology Review*, 8(1).
- Creswell, J. W. (2007). Qualitative inquiry & research design: Choosing among five approaches.
- Creswell, J. W. (2013). Steps in conducting a scholarly mixed methods study. 48. *DBER Speaker Series*.
- Creswell, J. W., & Creswell, J. D. (2018). Research design: Qualitative, quantitative, and mixed methods approaches. SAGE.
- Crotty, M. J. (1998). The foundations of social research: Meaning and perspective in the research process. *The foundations of social research*, 1-256.
- Cumming, J. J., & Maxwell, G. S. (2004). Profiles of educational assessment systems worldwide: Assessment in Australian schools: current practice and trends. *Assessment in Education: principles, policy & practice*, 11(1), 89-108.
- Dalby, A. (2003). Language in danger: The loss of linguistic diversity and the threat to our future. Columbia University Press.
- Dann, R. (2012). Promoting assessment as learning: Improving the learning process. Routledge.
- Davidson, J. (2009). Exhibitions: Connecting classroom assessment with culminating demonstrations of mastery. *Theory into practice*, 48(1), 36-43.
- Davies, D. (2016). The 'iron gate': high-stakes assessment at age 16 in Nepal and England. *Compare: A Journal of Comparative and International Education*, 46(4), 582-602.
- Davies, D., C. Collier, and A. Howe. 2012. "Assessing Scientific and Technological Enquiry Skills at Age 11 Using the E-scape System." *International Journal of Technology and Design Education*. 22: 247-263.
- Delandshere, G., & Jones, J. H. (1999). Elementary teachers' beliefs about assessment in mathematics: A case of assessment paralysis. *Journal of Curriculum and Supervision*, 14(3), 216.
- DeLuca, C., Chapman-Chin, A. E., LaPointe-McEwan, D., & Klinger, D. A. (2018). Student perspectives on assessment for learning. *The Curriculum Journal*, 29(1), 77-94.
- Denzin, N. K. (1970). The research act: A theoretical introduction to sociological methods. Aldine.
- Denzin, N. & Lincoln, Y. (2008). The Landscape of Qualitative Research, 3rd revised edition. Thousand Oaks. SAGE.
- Departamento de Evaluación, Medición y Registro Escolar. (DEMRE). Ranking requirements for University entrance. <https://psu.demre.cl/proceso-admision/factores-seleccion/puntaje-ranking>
- Departamento de Evaluación, Medición y Registro Escolar. (DEMRE). NEM requirements for University entrance. <https://psu.demre.cl/proceso-admision/factores-seleccion/notas-ensenanza-media>
- Departamento de Evaluación, Medición y Registro Escolar. (DEMRE). (2023). Informe Técnico de las pruebas de Admisión 2023 PAES regular. <https://demre.cl/estadisticas/documentos/informes/2023-informe-resultados-paes-regular-admision-2023.pdf>
- Duncan, C. R., & Noonan, B. (2007). Factors affecting teachers' grading and assessment practices. *Alberta Journal of Educational Research*, 53(1).

- Earle, S. (2014). Formative and summative assessment of science in English primary schools: evidence from the Primary Science Quality Mark. *Research in Science & Technological Education*, 32(2), 216-228.
- Elwood, J. (2012). Qualifications, examinations and assessment: Views and perspectives of students in the 14–19 phase on policy and practice. *Cambridge journal of education*, 42(4), 497-512.
- Elwood, J. (2013). Educational assessment policy and practice: A matter of ethics. *Assessment in Education: Principles, Policy & Practice*, 20(2), 205-220.
- Elwood, J., & Murphy, P. (2015). Assessment systems as cultural scripts: A sociocultural theoretical lens on assessment practice and products. *Assessment in Education: Principles, Policy & Practice*, 22(2), 182-192.
- Emmel, N. (2013). Sampling and choosing cases in qualitative research: A realist approach. *SAGE*.
- Escudero, T. (2003). Desde los tests hasta la investigación evaluativa actual. Un siglo, el XX, de intenso desarrollo de la Evaluación en Educación. *Revista Electrónica de Investigación y Evaluación Educativa*, 9(1), 11-43.
- Espinoza, Ó., & González, L. E. (2015). Equidad en el sistema de educación superior de Chile: acceso, permanencia, desempeño y resultados. *Educación superior en Chile: transformación, desarrollo y crisis*, 12, 517-580.
- Estaji, M. (2011). Ethics and validity stance in educational assessment. *English Language and Literature Studies*, 1(2), 89.
- Falcón Ccenta, Y., Aguilar Hernando, J., Augusto Luy-Montejo, C., & Morillo-Flores, J. (2021). Formative Assessment, Reality or Good Intentions? Case study in primary level teachers. *Propósitos y Representaciones*, 9(1), e1041. <https://dx.doi.org/10.20511/pyr2021.v9n1.1041>
- Fleer, M. (2015). Developing an assessment pedagogy: The tensions and struggles in re-theorising assessment from a cultural–historical perspective. *Assessment in Education: Principles, Policy & Practice*, 22(2), 224-246.
- Flick, U. (Ed.). (2009). The sage qualitative research kit: Collection. *SAGE*.
- Flitcroft, D., & Woods, K. (2018). What does research tell high school teachers about student motivation for test performance? *Pastoral Care in Education*, 36(2), 112-125.
- Flórez Petour, M. T. (2014). Systems, ideologies and history: A three-dimensional absence in the study of assessment reform processes. *Assessment in Education: Principles, Policy & Practice*, 22(1), 3-26.
- Flórez Petour, M. T., & Rozas Assael, T. (2020). Accountability from a social c perspective: Criticism and proposals. *Journal of Educational Change*, 21, 157-182.
- Frierson, H. T., Hood, S., Hughes, G.B. & Thomas, V.G. (2010). A guide to conducting culturally responsive evaluations. The 2002 User-Friendly Handbook for Project Evaluation, 63-73. *National Science Foundation*.
- Gaete Silva, A., Castro Navarrete, M., Pino Conejeros, F., & Mansilla Devia, D. (2017). Abandono de la profesión docente en Chile: Motivos para irse del aula y condiciones para volver. *Estudios pedagógicos (Valdivia)*, 43(1), 123-138.
- García, M. V., Alvarado, M. J., & Jiménez, A. (2000). La predicción del rendimiento académico: regresión lineal versus regresión logística. *Psicothema*, 12(2), 248-252.
- García-Huidobro, J. E. (2007). Desigualdad educativa y segmentación del sistema escolar. Consideraciones a partir del caso chileno. *Pensamiento Educativo, Revista de Investigación Latinoamericana (PEL)*, 40(1), 65-85.
- García-Huidobro, J. E., & Bellei, C. (2003). Desigualdad educativa en Chile. *Santiago: Universidad Alberto Hurtado*, 62.

- Garcia Palomer, C., & Paredes, R. D. (2010). Reducing the educational gap: good results in vulnerable groups. *The Journal of Development Studies*, 46(3), 535-555.
- Gardner, J., Holmes, B., & Leitch, R. (2009). *Assessment and social justice* (Vol. 16). Futurelab.
- Gilson, L., Schneider, H., & Orgill, M. (2014). Practice and power: a review and interpretive synthesis focused on the exercise of discretionary power in policy implementation by front-line providers and managers. *Health policy and planning*, 29(suppl_3), iii51-iii69.
- González, A. (20/01/2022). Expertos advierten que cambios en la PAES no disminuirán desigualdades en el sistema escolar. *Bío-Bío*. <https://www.biobiochile.cl/especial/admision-2022/noticias/2022/01/20/expertos-advierten-que-cambios-en-la-paes-no-disminuiran-desigualdades-del-sistema-escolar.shtml>
- González P., Guell, P., & Márquez, R. (2012). ¿A qué se deben las protestas? [Why the Protests?] *Revista Humanum*.
- Gonzalez, T. E., Hernandez-Saca, D. I., & Artiles, A. J. (2017). In search of voice: Theory and methods in K-12 student voice research in the US, 1990–2010. *Educational Review*, 69(4), 451-473.
- González-Escobar, M., Silva-Peña, I., Gandarillas, A. P., & Kelchtermans, G. (2020). Abandono docente en América Latina: Revisión de la literatura. *Cuadernos de Pesquisa*, 50, 592-604.
- Green, S. K., Johnson, R. L., Kim, D. H., & Pope, N. S. (2007). Ethics in classroom assessment practices: Issues and attitudes. *Teaching and Teacher Education*, 23(7), 999-1011.
- Green, J., & Thorogood, N. (2004). *Qualitative methods for health research*. SAGE.
- Grodsky, E., Warren, J. R., & Felts, E. (2008). Testing and social stratification in American education. *Annual Review of Sociology*, 34, 385–404.
- Gronlund, G. (2003). *Focused early learning: A planning framework for teaching young children*. St. Paul, MN: Redleaf Press.
- Gysling, J. (2016). The historical development of educational assessment in Chile: 1810–2014. *Assessment in Education: Principles, Policy & Practice*, 23(1), 8-25.
- Gysling, J. (2017). La evaluación: ¿dispositivo para promover el aprendizaje de todos o para seleccionar? La formación de profesores en evaluación en Chile. *Leiden University*.
- Haertel, E., Ferrara, S., Korpi, M., & Prescott, B. (1984). Testing in secondary schools: Student perspectives. *American Educational Research Association*.
- Hammersley, M. (2012). What is qualitative research? *A&C Black*.
- Hargreaves, A. (1994). Changing teachers, changing times: Teachers' work and culture in the postmodern age. *Continuum & Teacher's College Press*.
- Harlen, W. (2004). A systematic review of the evidence of reliability and validity of assessment by teachers used for summative purposes. *EPPI-Centre, Social Science Research Unit, Institute of Education, University of London*.
- Harlen, W. (2007). Criteria for evaluating systems for student assessment. *Studies in Educational Evaluation*, 33(1), 15-28.
- Harlen, W. 2013. Assessment and Inquiry-Based Science Education: Issues in Policy and Practice. *Global Network of Science Academies*.
- Harlen, W., Gipps, C., Broadfoot, C., & Nuttall, D. (1994). Assessment and the Improvement of Education' in Moon, B. and Mayes, AS. *Teaching and Learning in the Secondary School*.
- Harris, L. R., & Brown, G. T. (2009). The complexity of teachers' conceptions of assessment: Tensions between the needs of schools and students. *Assessment in Education: Principles, Policy & Practice*, 16(3), 365-381.
- Hartell, E., & Strimel, G. J. (2019). What is it called and how does it work: examining content validity and item design of teacher-made tests. *International Journal of Technology and Design Education*, 29(4), 781-802.

- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research*, 77(1), 81-112.
- Hennink, M., Hutter, I., & Bailey, A. (2011). *Qualitative research methods*. SAGE.
- Heritage, M. (2008). Learning progressions: Supporting instruction and formative assessment.
- Herman, J. L., Haertel, E., & Haertel, E. H. (Eds.). (2005). Uses and misuses of data for educational accountability and improvement (Vol. 104, No. 2). *Wiley-Blackwell*.
- Hermanns, H. (2004). Interviewing as an activity. *A companion to qualitative research*, 1.
- Hernández-Nodarse, M. (2017). ¿Por qué ha costado tanto transformar las prácticas evaluativas del aprendizaje en el contexto educativo? Ensayo crítico sobre una patología pedagógica pendiente de tratamiento. *Revista Electrónica Educare*, 21(1), 1-27.
- Hernández-Nodarse, M., Gregory Casamayor, M., & Tomalá Solano, D. G. (2018). ¿Cómo puede transformarse verdaderamente la evaluación del aprendizaje practicada por los profesores? Una experiencia pedagógica por la mejora del profesorado. *Revista Ciencias Pedagógicas e Innovación*, 6(1), 88-106.
- Hirsh, Å. (2020). When assessment is a constant companion: Students' experiences of instruction in an era of intensified assessment focus. *Nordic Journal of Studies in Educational Policy*, 6(2), 89-102.
- HKEEA. (Hong Kong Examination and Assessment Authority). School Based Assessment. <https://www.hkeea.edu.hk/en/sba/introduction/>
- Holloway, I., & Todres, L. (2003). The status of method: flexibility, consistency and coherence. *Qualitative research*, 3(3), 345-357.
- Honey, N., & Carrasco, A. (2022). A new admission system in Chile and its foreseen moderate impact on access for low-income students. *Educational Evaluation and Policy Analysis*, 45(1), 108-133.
- Hughes, S., Davis, T. E., & Imenda, S. N. (2019). Demystifying theoretical and conceptual frameworks: A guide for students and advisors of educational research. *J Soc Sci*, 58(1-3), 24-35.
- Hult, F. M. (2014). Covert bilingualism and symbolic competence: Analytical reflections on negotiating insider/outsider positionality in Swedish speech situations. *Applied Linguistics*, 35(1), 63-81.
- Imenda, S. (2014). Is there a conceptual difference between theoretical and conceptual frameworks?. *Journal of social sciences*, 38(2), 185-195.
- Irving, S. E., Harris, L. R., & Peterson, E. R. (2011). 'One assessment doesn't serve all the purposes' or does it? New Zealand teachers describe assessment and feedback. *Asia Pacific Education Review*, 12(3), 413-426.
- James, M., & Pedder, D. (2006). Beyond method: Assessment and learning practices and values. *The Curriculum Journal*, 17(2), 109-138.
- Johnson, C. C., Sondergeld, T. A., & Walton, J. B. (2019). A Study of the Implementation of Formative Assessment in Three Large Urban Districts. *American Educational Research Journal*, 56(6), 2408-2438.
- Kanuha, V. K. (2000). "Being" native versus "going native": Conducting social work research as an insider. *Social work*, 45(5), 439-447.
- Kennedy, K. J., Chan, J. K. S., Fok, P. K., & Yu, W. M. (2008). Forms of assessment and their potential for enhancing learning: Conceptual and cultural issues. *Educational Research for Policy and Practice*, 7, 197-207.
- Kiger, M. E., & Varpio, L. (2020). Thematic analysis of qualitative data: AMEE Guide No. 131. *Medical teacher*, 42(8), 846-854.

- Kim, H. W. (2012). Research with children: Challenges and dilemmas as an insider researcher. *Early Child Development and Care*, 182(2), 263-276.
- Klenowski, V. (2009). Assessment for learning revisited: An Asia-Pacific perspective.
- Kohn, A. (2001). Beware of the standards, not just the tests. *Education Week*, 21(4), 52.
- Kožuh, A. (2019). Perspectives on assessment at school. *Journal of Contemporary Educational Studies/Sodobna Pedagogika*, 70(2).
- Kremerman, M. (2007). Radiografía del financiamiento de la educación chilena: diagnóstico, análisis y propuestas. *Juventud y enseñanza media del bicentenario*.
- Kvale, S., & Brinkmann, S. (2009). Interviews: Learning the craft of qualitative research. SAGE.
- Lam, R. (2019). Teacher assessment literacy: Surveying knowledge, conceptions and practices of classroom-based writing assessment in Hong Kong. *System*, 81, 78-89.
- Lambert, D., & Lines, D. (2000). Understanding assessment: Purposes, perceptions, practice. *Routledge*.
- Larroucau, T., Ríos, I., & Mizala, A. (2013). Efecto de la incorporación del ranking de notas en la selección universitaria. *DEMRE*.
- Leung, C. Y., & Andrews, S. (2012). The mediating role of textbooks in high-stakes assessment reform. *ELT journal*, 66(3), 356-365.
- Liehr, P., & Smith, M. J. (1999). Middle range theory: Spinning research and practice to create knowledge for the new millennium. *Advances in Nursing Science*, 21(4), 81-91.
- Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. SAGE.
- Londoño-Restrepo, L. A. (2015). Relaciones de poder en la evaluación de los aprendizajes con docentes en formación. *Entramado*, 11(1), 156-174.
- Lupton, D. (2020). Social Research for a COVID and Post-COVID World: An Initial Agenda.
- Mäkipää, T., & Ouakrim-Soivio, N. (2019). Finnish Upper Secondary School Students' Perceptions of Their Teachers' Assessment Practices. *Journal of Teaching and Learning*, 13(2), 23-42.
- Malaby, M., & Esh, M. (2014). A Review of "The Active/Ethical Professional: A Framework for Responsible Educators."
- Maldonado-Fuentes, A. C. (2021). Representación escolarizada de la evaluación: Un aprendizaje social profesional. *ALTERIDAD. Revista de Educación*, 16(2), 184-197.
- Maroy, C., & Pons, X. (2021). Re-conceptualising education policy trajectories in a globalised world: lessons from a multi-level comparison of accountability in France and Quebec. *Comparative Education*, 57(4), 560-578.
- Marshall, B & Drummond, M J (2006). How teachers engage with Assessment for Learning: lessons from the classroom. *Research Papers in Education*, 21, (2), 133-149.
- Martínez Quiñones, G. (2019). Percepción de docentes y estudiantes sobre implementación de estrategias diversificadas de evaluación. Una mirada a un establecimiento educativo secundario de la comuna de San Miguel. *Universidad de Chile*.
- Mason, J. (2002). Qualitative researching. SAGE.
- Matthews, R., & Ross, E. (2010). Research methods: A practical guide for the social sciences. *Pearson Education Ltd*.
- Maxwell, J. A. (2013). Qualitative research design: An interactive approach. Thousand Oaks. SAGE.
- McMillan, J. H. (2001). Secondary teachers' classroom assessment and grading practices. *Educational Measurement: Issues and Practice*, 20(1), 20-32.
- McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational measurement: Issues and practice*, 22(4), 34-43.
- McMillan, J. H. (2013). Why we need research on classroom assessment. *SAGE handbook of research on classroom assessment*, 3-16.

- McNess, E., Arthur, L., & Crossley, M. (2015). 'Ethnographic dazzle' and the construction of the 'Other': revisiting dimensions of insider and outsider research for international and comparative education. *Compare: A Journal of Comparative and International Education*, 45(2), 295-316.
- McNess, Broadfoot, P., & Osborn, M. (2003). Is the effective compromising the affective? *British educational research journal*, 29(2), 243-257.
- Meng, C. C., Idris, N., & Eu, L. K. (2014). Secondary Students' Perceptions of Assessments in Science, Technology, Engineering, and Mathematics (STEM). *Eurasia Journal of Mathematics, Science & Technology Education*, 10(3).
- Merkens, H. (2004). 4.4 Selection procedures, sampling, case construction. *A companion to qualitative research*.
- Meriño-Montero, I. (2021). Repitencia escolar: análisis y reflexiones en un establecimiento educacional en la comuna de Valparaíso, Chile. *Revista Innova Educación*, 3(1), 20-38.
- Merriam, S. B. (1988). Case study research in education: A qualitative approach. *Jossey-Bass*.
- Merriam, S. B., & Tisdell, E. J. (2016). Qualitative research: a guide to design and implementation (Fourth, Ser. The jossey-bass higher and adult education series). *John Wiley & Sons*.
- Mertens, D. M. (1998). Research methods in education and psychology: Integrating diversity with quantitative and qualitative approaches. Thousand Oaks. SAGE.
- Milligan, L. (2016). Insider-outsider-inbetweener? Researcher positioning, participative methods and cross-cultural educational research. *Compare: a journal of comparative and international education*, 46(2), 235-250.
- Ministerio de Educación. (Mineduc). (2016) "Percepciones, significados y necesidades sentidas en relación a la evaluación de aula en profesores, directivos y estudiantes de educación escolar de tres regiones del país". Unidad de Curriculum y Evaluación. https://www.curriculumnacional.cl/614/articles-70901_archivo_01.pdf
- Ministerio de Educación. (Mineduc). (2017). Evaluación formativa en el aula: Orientaciones para docentes: Integrando el uso pedagógico de la evaluación en la enseñanza. https://bibliotecadigital.mineduc.cl/bitstream/handle/20.500.12365/17448/ORIENTACIONES_EVAL_FORMATIVA_DOCENTES.pdf?sequence=1&isAllowed=y
- Ministerio de Educación. (Mineduc). (2018a). Decreto de calificación y promoción escolar N°67/2018. https://www.curriculumnacional.cl/614/articles-70983_archivo_01.pdf
- Ministerio de Educación. (Mineduc). (2018b). Política para el fortalecimiento de la evaluación en el aula. <https://bibliotecadigital.mineduc.cl/bitstream/handle/20.500.12365/2255/mono-587.pdf?sequence=1&isAllowed=y>
- Ministerio de Educación. (Mineduc). (2020a). Orientación al sistema escolar en el contexto de COVID-19. https://www.mineduc.cl/wp-content/uploads/sites/19/2020/03/OrientacionesContextoCOVID19_2703.pdf#:~:text=en%20contexto%20de%20COVID%2D19,-En%20este%20sentido&text=Asegurar%20que%20el%20establecimiento%20mantenga,derivaci%C3%B3n%20a%20las%20redes%20especializadas.
- Ministerio de Educación. (Mineduc). (2020b). Criterios de evaluación, calificación y promoción de estudiantes de 1° básico a 4° año medio. <https://www.mineduc.cl/wp-content/uploads/sites/19/2020/08/CriteriosPromocionEscolarCalificacionEvaluacion.pdf>
- Ministerio de Educación. (Mineduc). (n.d). Evaluación formativa en el aula. https://www.curriculumnacional.cl/614/articles-89342_archivo_01.pdf
- Ministerio de Desarrollo Social. (2015). Reporte de Desarrollo social y familiar. http://www.desarrollosocialyfamilia.gob.cl/pdf/upload/Libro_IDS_2015_final.pdf

- Mitchell, R. (2017). An ethnographic case study of the agendas, participation and influence of stakeholders at an urban government primary school in Tigray, Ethiopia. *University of Leicester*.
- Mizala, A. & Romaguera, P. (2000) Determinación de factores explicativos de los resultados escolares en educación media en Chile. *Centro de Economía Aplicada, Universidad de Chile*, N° 85.
- Montecinos, C., Cortez, M., & Walker, H. (2015). School administrators' understandings and management of barriers for the school's involvement in the practicum component of initial teacher education in Chile. *International Journal of Educational Development*, 43, 100-108.
- Monteiro, V., Mata, L., & Santos, N. N. (2021). Assessment conceptions and practices: Perspectives of primary school teachers and students. In *Frontiers in Education* (Vol. 6, p. 631185). Frontiers Media SA.
- Monzón-Troncoso, M. Y. (2015). Evaluación del aprendizaje: un recorrido histórico y epistemológico. *Revista de Ciencias de la Educación Academicus*, 6(1), 12-21.
- Morin, E., Ciurana, E. R., & Motta, R. D. (2002). Educar en la era planetaria: el pensamiento complejo como método de aprendizaje en el error y la incertidumbre humana. *Universidad de Valladolid*.
- Morris, P., & Adamson, B. (2010). Curriculum, schooling and society in Hong Kong (Vol. 1). Hong Kong University Press.
- Mujica-Johnson, F. N. (2020). Análisis crítico del currículo escolar en Chile en función de la justicia social Critical Analysis of the School Curriculum in Chile Based on Social Justice. *Revista Electrónica Educare* 24(1), 1-14.
- Muñoz Cuchca, E., & Solís Trujillo, B. P. (2021). Enfoque cualitativo y cuantitativo de la evaluación formativa. *Revista de Ciencias Humanísticas y Sociales (ReHuSo)*, 6(3), 1-11.
<https://doi.org/10.5281/zenodo.5512590>
- Murchan, D. (2018). Introducing school-based assessment as part of junior cycle reform in Ireland: a bridge too far? *Educational Assessment, Evaluation and Accountability*, 30(2), 97-131.
- Murillo, F. J., & Hidalgo, N. (2015). Enfoques fundamentantes de la evaluación de estudiantes para la justicia social. *Revista Iberoamericana de Evaluación Educativa*, 8(1), 43-61.
- Murphy, S. (2010). The pull of PISA: Uncertainty, influence, and ignorance. *Inter-American Journal of Education for Democracy*, 3(1), 27-44.
- Neidhöfer, G., N. Lustig & M. Tommasi (2021). Intergenerational transmission of lockdown consequences: prognosis of the longer-run persistence of COVID-19 in Latin America, Documentos de Trabajo N° 148. *Universidad de San Andrés*.
- Newton, P. (2003). The defensibility of national curriculum assessment in England. *Research Papers in Education*, 18(2), 101-127.
- Newton, P. (2007). Clarifying the purposes of educational assessment. *Assessment in education*, 14(2), 149-170.
- Newton, P. (2014) Validating the cultivation of valuable learning. In State of the field review: Assessment and learning. *Oslo: Norwegian Knowledge Centre for Education*.
- New Zealand Qualifications Authority (2023). External assessment. https://www2.nzqa.govt.nz/ncea/external-assessment/#e12767_heading1
- Nothdurfter, U. (2018). "The street-level delivery of activation policies: Constraints and possibilities for a practice of citizenship." *European Journal of Social Work* 19.3-4 (2016): 420-440.
- Oliva, M. A. (2008). Política educativa y profundización de la desigualdad en Chile. *Estudios pedagógicos (Valdivia)*, 34(2), 207-226.

- Oliva, M. A., & Gascón, F. (2016). Estandarización y racionalidad política neoliberal: bases curriculares de Chile. *Cadernos Cedes*, 36, 301-318.
- Othman, J. (2019). REFORM IN ASSESSMENT: TEACHERS' BELIEFS AND PRACTICES. *Journal of Teaching English for Specific and Academic Purposes*, 501-512.
- Parcerisa, L., & Falabella, A. (2017). La consolidación del Estado evaluador a través de políticas de rendición de cuentas: trayectoria, producción y tensiones en el sistema educativo chileno. *Education Policy Analysis Archives/Archivos Analíticos de Políticas Educativas*, (25), 1-24.
- Pellegrino, J. W., Chudowsky, N., & Glaser, R. (Eds.). (2001). Knowing what students know: The science and design of educational assessment. *Washington: National Academy Press*.
- Penuel, W., & Shepard L. (2016). Assessment and teaching. In *Handbook of research on teaching. American Educational Research Association*.
- Perassi, Z. (2008). La evaluación en educación: un campo de controversias. *Ediciones del Proyecto y Ediciones LAE*.
- Peters, B., & Contreras, G. (2019). Criterios de evaluación interpretados por estudiantes de segundo y tercer año medio en la asignatura de matemáticas en Chile. *Perspectiva Educacional*, 58(3), 46-72.
- Phye, G. D. (1996). *Handbook of classroom assessment: Learning, achievement, and adjustment. Academic Press*.
- Picaroni, B. (2009). La evaluación en las aulas de primaria: usos formativos, calificaciones y comunicación con los padres. *Instituto de Evaluación Educativa. UCU/GTEE-PREAL*.
- Pizarro Gamero, E. A., & Gómez Muskus, S. S. (2019). Concepciones docentes sobre evaluación: de los lineamientos, el discurso y la práctica. *Praxis & Saber*, 10(22), 71-88.
- Plan de Desarrollo Comunal (PLADECO). (2011). Caracterización y diagnóstico comunal. Diciembre 2011.
https://www.sanbernardo.cl/contenidos/transparencia/plan.regulador.comunal/TOMO_1_PLADECO.pdf
- Pohpam, W. J. (2000). *Modern Educational Measurement; Practical Guidelines for Educational Leaders*.
- Poorthuis, A. M., Juvonen, J., Thomaes, S., Denissen, J. J., Orobio de Castro, B., & Van Aken, M. A. (2015). Do grades shape students' school engagement? The psychological consequences of report card grades at the beginning of secondary school. *Journal of Educational Psychology*, 107(3), 842.
- Pope, N., Green, S. K., Johnson, R. L., & Mitchell, M. (2009). Examining teacher ethical dilemmas in classroom assessment. *Teaching and Teacher Education*, 25(5), 778-782.
- Prasad, P. (2017). *Crafting qualitative research: Beyond positivist traditions. Taylor & Francis*.
- Prieto, M., & Contreras, G. (2008). Concepciones de los docentes que informan sobre sus prácticas de evaluación: un problema por revelar. *Estudios pedagógicos (Valdivia)*, 34(2), 245-262.
- Puad, L. M. A. Z., & Ashton, K. (2021). Teachers' views on classroom-based assessment: an exploratory study at an Islamic boarding school in Indonesia. *Asia Pacific Journal of Education*, 41(2), 253-265.
- Puga, I. (2011). Escuela y estratificación social en Chile: ¿cuál es el rol de la municipalización y la educación particular subvencionada en la reproducción de la desigualdad social? *Estudios pedagógicos (Valdivia)*, 37(2), 213-232.
- Puspitasari, M., & Pelawi, M. A. (2023). English High-stakes Testing in Indonesia: Friend or Foe?. *Journal of English Education and Teaching*, 7(4), 1071-1084.
- Race, P., Brown, S., & Smith, B. (2005). *500 tips on assessment (2nd ed.). Routledge Falmer*.

- Rachels, J. (2010). A Short introduction to moral philosophy. 2010) *The Right Thing to Do: Basic Readings in Moral Philosophy. McGraw-Hill International Edition.*
- Raczynski, D., & Salinas, D. (2008). Aportes y desafíos en la evaluación de políticas y programas de mejoramiento educativo: reflexiones desde nuestra práctica de evaluación en Chile. *RIEE. Revista Iberoamericana de Evaluación Educativa.*
- Rahman, M. S. (2023). From fried egg to mashed potato and lentils: navigating positionalities in ethnographic research in a global south context. *International Journal of Research & Method in Education, 46(4), 408-420.*
- Rasooli, A., Rasegh, A., Zandi, H., & Firoozi, T. (2023). Teachers' Conceptions of Fairness in Classroom Assessment: An Empirical Study. *Journal of Teacher Education, 74(3), 260–273.*
- Ravela, P., Leymoníe, J., Viñas, J., & Haretche, C. (2014). La evaluación en las aulas de secundaria básica en cuatro países de América Latina. *Propuesta Educativa, (41), 20-45.*
- Ravitch, S. M., & Riggan, M. (2016). *Reason & rigor: How conceptual frameworks guide research.* Sage Publications.
- Reardon, S. F. (2021). The Economic Achievement Gap in the US, 1960-2020: Reconciling Recent Empirical Findings. CEPA Working Paper No. 21-09. *Stanford Center for Education Policy Analysis.*
- Remesal, A. (2011). Primary and secondary teachers' conceptions of assessment: A qualitative study. *Teaching and teacher education, 27(2), 472-482.*
- Research School Network (2021). Key Pillars of Assessment. What does effective assessment practice really look like? <https://researchschool.org.uk/news/key-pillars-of-assessment>
- Reyneke, M. (2016). School-based assessment in English language teaching: weighing the cow will not fatten it. *Per Linguam: a Journal of Language Learning, 32(2), 1-14.*
- Roco, R. (2010). Caracterización de los establecimientos educacionales en Chile: la necesidad de nuevas consideraciones. *Primer Congreso Interdisciplinario de Investigación en Educación, CIAE, Universidad de Chile y CEPPE, Pontificia Universidad Católica de Chile. Congreso en Santiago, Chile.*
- Rodgers Gibson, M. (2019). Towards a neoliberal education system in Queensland: Preliminary notes on senior secondary schooling reforms. *Policy Futures in Education, 17(8), 983-999.*
- Rodríguez-Garcés, C. & Jarpa-Arriagada, C. G. (2015). Capacidad predictiva de las notas en enseñanza media sobre el rendimiento en pruebas de selección universitaria: El caso chileno. *Aula Abierta, 43, pp. 61-68.*
- Román, M. (2013). Factores asociados al abandono y la deserción escolar en América Latina: una mirada en conjunto. *Revista Iberoamericana sobre calidad, eficiencia y cambio en educación, 11(2), 33-59.*
- Roth, W. M., & Radford, L. (2011). A cultural-historical perspective on mathematics teaching and learning (Vol. 2). *Springer science & business media.*
- Sánchez-Sánchez, G. I., & Jara-Amigo, X. E. (2019). Los docentes en formación y la comprensión de sus experiencias pedagógicas: la evaluación del aprendizaje. *Revista Internacional de Investigación en Ciencias Sociales, 15(1).*
- Santos Guerra, M. Á. (1988). Patología general de la evaluación educativa. *Infancia y aprendizaje, 11(41), 143-158.*
- Santos Guerra, M. A. (1996). Evaluar es comprender. De la concepción técnica a la dimensión crítica. *Investigación en la Escuela, (30), 5-13.*
- Santos Guerra, M. Á. (2003). Una flecha en la diana. La evaluación como aprendizaje. *Narcea.*
- Santos Guerra, M. Á. (2014). La evaluación como aprendizaje: Cuando la flecha impacta en la diana (Vol. 200). *Narcea Ediciones.*

- Schmidt, C. (2004). The analysis of semi-structured interviews. *A companion to qualitative research*, 253-258.
- Scott, J. (1990). *A matter of Record*. Cambridge: Polity Press.
- Scott, G. E. (2018). A quantitative study of the implementation of formative assessment strategies in the classroom. *Ashland University*.
- Scottish Government (2011). Curriculum for excellence building the curriculum 5: a framework for assessment. <https://www.education.gov.scot/Documents/btc5-framework.pdf>
- Scottish Qualification Authority, SQA. (n.d.) National Qualifications: Overview of Quality Assurance. <https://www.sqa.org.uk/sqa/74667.html>
- Shepard, L. A. (2019). Classroom assessment to support teaching and learning. *The ANNALS of the American Academy of Political and Social Science*, 683(1), 183-200.
- Shields, P. M., & Tajalli, H. (2006). Intermediate theory: The missing link in successful student scholarship. *Journal of public affairs education*, 12(3), 313-334.
- Sivenbring, J. (2019). Making Sense and Use of Assessments. *Scandinavian Journal of Educational Research*, 63(5), 759-770.
- Smith, L. T. (2021). *Decolonizing methodologies: Research and indigenous peoples*. Bloomsbury Publishing.
- Smith, J., Flowers, P., & Larkin, M. (2009). *Interpretative Phenomenological Analysis. Theory, Method and Research*. Sage.
- Smyth, E., & Banks, J. (2012). High stakes testing and student perspectives on teaching and learning in the Republic of Ireland. *Educational Assessment, Evaluation and Accountability*, 24(4), 283-306.
- Somekh, B., & Lewin, C. (Eds.). (2005). *Research methods in the social sciences*. SAGE.
- Sondel, B. (2016). "No excuses" in New Orleans: The silent passivity of neoliberal schooling. In *The Educational Forum* (Vol. 80, No. 2, pp. 171-188). Routledge.
- Stiggins, R. J. (1999). Evaluating classroom assessment training in teacher education programs. *Educational measurement: Issues and practice*, 18(1), 23-27.
- Stiggins, R. J. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, 83(10), 758-765.
- Stiggins, R. J., & Conklin, N. F. (1992). In teachers' hands: Investigating the practices of classroom assessment. *Suny Press*.
- Stiggins, R. (2014). *Defensible teacher evaluation: Student growth through classroom assessment*. Corwin Press.
- Stobart, G., & Eggen, T. (2012). High-stakes testing—value, fairness and consequences. *Assessment in Education: Principles, Policy & Practice*, 19(1), 1-6.
- Strandler, O. (2016). Equity through assessment? Teachers' mediation of outcome-focused reforms in socioeconomically different schools. *Scandinavian Journal of Educational Research*, 60(5), 538-553.
- Stromquist, N. P., & Sanyal, A. (2013). Student resistance to neoliberalism in Chile. *International studies in sociology of education*, 23(2), 152-178.
- Stutchbury, K., & Fox, A. (2009). Ethics in educational research: introducing a methodological tool for effective ethical analysis. *Cambridge Journal of Education*, 39(4), 489-504.
- Swanson, C. B., & Stevenson, D. L. (2002). Standards-based reform in practice: Evidence on state policy and classroom instruction from the NAEP state assessments. *Educational evaluation and policy analysis*, 24(1), 1-27.
- System-Wide Analysis of Assessment Practices, SWAAP (2017). *Australian council for Educational Research (ACER)*.

- Tan, K. H. K. (2017). Asking questions of (what) assessment (should do) for learning: the case of bite-sized assessment for learning in Singapore. *Educational Research for Policy and Practice*, 16(2), 189-202.
- Tapan, S (2001) A study of present learning assessment practices to ensure quality of primary education. Basic education studies, 2000-2001. *UNESCO Office*.
- Taras, M. (2005). Assessment—summative and formative—some theoretical reflections. *British journal of educational studies*, 53(4), 466-478.
- Temple, B., & Young, A. (2004). Qualitative research and translation dilemmas. *Qualitative research*, 4(2), 161-178.
- The Organization for Economic Cooperation and Development. (OECD). (2016). Report for countries: Chile.
<http://gpseducation.oecd.org/CountryProfile?primaryCountry=CHL&treshold=10&topic=EO>
- Thomson, P., & Gunter, H. (2011). Inside, outside, upside down: The fluidity of academic researcher 'identity' in working with/in school. *International Journal of Research & Method in Education*, 34(1), 17-30.
- Tong, S. Y. A., & Adamson, B. (2015). Student voices in school-based assessment. *Australian Journal of Teacher Education*, 40(2), 2.
- Toranzos, L. V. (2014). Evaluación educativa: hacia la construcción de un espacio de aprendizaje. *Propuesta educativa*, (41), 9-19.
- Torché, F. (2005). Privatization reform and inequality of educational opportunity: The case of Chile. *Sociology of education*, 78(4), 316-343.
- Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British educational research journal*, 27(5), 615-631.
- Torres, J., Acevedo, D., & Gallo, L. (2015). Causas y consecuencias de la deserción y repitencia escolar: una revisión general en el contexto latinoamericano. *Cultura Educación y Sociedad* 6(2), 157-187.
- Torres-Olave, B. (2022). Strengthening science teacher agency: a Freirean perspective of two teacher-led communities in Chile. University of Bristol.
- Towne, L., & Shavelson, R. J. (2002). Scientific research in education. *Washington, DC: National Research Council*.
- Tuckett, A. G. (2005). Applying thematic analysis theory to practice: A researcher's experience. *Contemporary nurse*, 19(1-2), 75-87.
- UChile newspaper (25/11/2021). Alcalde de San Bernardo se querrela contra ex edil Nora Cuevas por fraude al fisco. <https://radio.uchile.cl/2021/11/25/alcalde-de-san-bernardo-se-querrela-contra-ex-edil-nora-cuevas-por-fraude-al-fisco/>
- Valenzuela, J. P., Bellei, C., & Ríos, D. D. L. (2014). Socioeconomic school segregation in a market-oriented educational system. The case of Chile. *Journal of education Policy*, 29(2), 217-241.
- Vera Sagredo, A., Poblete Correa, S., & Díaz Larenas, C. (2017). Percepción de los docentes chilenos sobre sus perspectivas, habilidades y prácticas evaluativas en el aula. *Estudios pedagógicos (Valdivia)*, 43(3), 361-372.
- Verdinelli, S., & Scagnoli, N. I. (2013). Data display in qualitative research. *International Journal of Qualitative Methods*, 12(1), 359-381.
- Villalobos, C., & Béjares, C. (2017). Políticas de promoción y repitencia en Chile: recomendaciones desde una mirada comparada. CEPPE Policy Brief N°14, CEPPE UC.
- Villalobos, C., & Quaresma, M. L. (2015). Sistema escolar chileno: características y consecuencias de un modelo orientado al mercado. *Convergencia*, 22(69), 63-84.

- Villalobos, C., & Valenzuela, J. P. (2012). Polarización y cohesión social del sistema escolar chileno. *Revista de análisis económico*, 27(2), 145-172.
- Vogt, K., & Tsagari, D. (2014). Assessment literacy of foreign language teachers: Findings of a European study. *Language Assessment Quarterly*, 11(4), 374-402.
- Watson, N. (2014). IWRM in England: bridging the gap between top-down and bottom-up implementation. *International Journal of Water Resources Development*, 30(3), 445-459.
- Wellington, J. (2015). Educational research: Contemporary issues and practical approaches. *Bloomsbury Publishing*.
- Wiliam, D. (2006) Assessment for learning: why no profile in US policy? J. Gardner (Ed.) *Assessment and learning*, 169-185. *SAGE*.
- Wiliam, D. (2011). What is assessment for learning? *Studies in educational evaluation*, 37(1), 3-14.
- Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.). *The future of assessment: Shaping teaching and learning* (pp. 53–82). *Lawrence Erlbaum Associates*.
- Willis, J. (2008) Assessment for learning: a socio-cultural approach. In: 'AARE 2008 International education research conference: Brisbane: papers collection', compiled by P L Jeffrey. Melbourne, *Australian Association for Research in Education*.
- Whale, K., Cramer, H., Wright, A., Sanders, C., & Joinson, C. (2017). 'What does that mean?': a qualitative exploration of the primary and secondary clinical care experiences of young people with continence problems in the UK. *BMJ open*, 7(10), e015544.
- World Health Organization. (WHO). (2020). Coronavirus Disease (COVID-19) Dashboard. https://covid19.who.int/?gclid=Cj0KCCQjwuJz3BRDTARIsAMgHxUAsKYT966OhFFKNz2JVNbKP IV67gzfgPuakbEcO2yfpHak-eCq0UwaAjUGEALw_wcB
- Yates, A., & Johnston, M. (2018). The impact of school-based assessment for qualifications on teachers' conceptions of assessment. *Assessment in Education: Principles, Policy & Practice*, 25(6), 638-654.
- Yin, R. K. (2009). *Case study research: Design and methods* (Vol. 5). *SAGE*.
- Young, V. M., & Kim, D. H. (2010). Using assessments for instructional improvement: A literature review. *education policy analysis archives*, 18, 19.
- Zancajo, A. (2019). Education markets and schools' mechanisms of exclusion: The case of Chile. *Education Policy Analysis Archives*, 27(130).
- Zhao, X., Van den Heuvel-Panhuizen, M., & Veldhuis, M. (2017). Classroom assessment in the eyes of Chinese primary mathematics teachers: A review of teacher-written papers. *Studies in Educational Evaluation*, 52, 42-54.

Appendixes

Appendix A: Some examples of countries using SBA and some of its characteristics.

<i>Country/zone</i>	<i>Description</i>	<i>High-stakes/External testing</i>
<i>Australia</i>	<p>The ACACA set guidelines for the requirements of students' knowledge and skills as defined and required by the syllabus. The guidelines highlight that the set of assessment instruments should be variegated and supervised work, involving a range and balance of assessment types (Cumming and Maxwell, 2004). However, the SBA experience in Australia has raised concerns related to student performance expectations within curriculum guidelines, assessment guidelines/rules and moderation practices (Cumming and Maxwell, 2004).</p>	<p>Incorporated SBA through a combination of internal assessment in addition to external examinations, reporting in the end of the secondary school within a high stakes system (Cumming and Maxwell, 2004). The case of Queensland is insightful, where final subject assessments are entirely based on SBA, as in Chile, without the external examinations.</p>
<i>Finland</i>	<p>The country has set assessment guidelines for secondary education in its National Core Curriculum. It establishes that the assessment needs to be conducted during and at the end of the learning process in a formative and also in a summative way (Mäkipää and Ouakrim-Soivio, 2019).</p>	<p>The system is also high stakes, insofar as students must complete seventy-five courses in three years to participate in the only examination that students have (Mäkipää and Ouakrim-Soivio, 2019).</p>
<i>Hong Kong</i>	<p>Two of the main challenges for SBA are to share rubrics with students which allow them to understand the purposes and requirements of the tasks they are carrying out and also to use students' results to adjust their teaching (Cheng, 2019). It is also highlighted that teachers showed that SBA did not have much effect on their teaching "at all" (Cheng, 2019). The follow up of SBA should allow teachers to take students'</p>	<p>SBA started out being used as part of English assessment and has been increasingly rolled out to core and elective units of the curriculum (HKEAA, 2021). It is combined with external examination at the end of the course.</p>

	performance considered in lesson planning (Cheng, 2019).	
<i>Indonesia</i>	The case of Indonesia is particularly insightful, as they began using a new assessment decree promoting SBA in 2020, at the same time as Chile (Mineduc, 2018a; Puad and Ashton, 2021). The new policy promotes the use of peer and self-assessment. However, results show that assessment in the classroom is carried out with a summative approach rather than a formative one, exhibiting a strong component of graded behaviour and attitudes attached to academic success (Puad and Ashton, 2021). In that way, teachers control students' behaviour through the grading processes. Policies in the country tend to be top-down, namely imposed without consider teachers' voices neither local school conditions, such as numerous classrooms (40 students).	High stakes exams and teacher-centred pedagogies have made the implementation of formative assessment problematic, facing challenges that commonly appear in non-Western countries, such as test-driven learning and passive students (Puad and Ashton, 2021). Standardised high-stakes testing was eliminated in 2021 (Puspitasari and Pelawi, 2023).
<i>Iran</i>	The country uses a combination of summative and formative assessment that is promoted for the national policy. Even when the formative aspect is valued, teachers had historically equated assessment with summative purposes: design, administer and score tests (Rasooli et al., 2023).	Students have national examinations in order to receive a high school diploma that allows them to take a test (INUEE) for university entrance (Ahmadi Safa and Sheykhmololuki, 2023).
<i>Scotland</i>	The Scottish government has promoted the use of SBA, although not explicitly using the concept itself. Instead, they promote the use of both formative and summative making explicit that "the terms 'formative' and 'summative' do not describe a type or form of assessment per se, but instead they describe how the assessments are used" and specifying that any assessment can be used formatively (Scottish Government, 2011, p.23). The Scottish government also promote the use of a range and variety of evidence on students' learning.	The Scottish Qualification Authority (SQA) set three external examinations for secondary students between 14 and 17 years old (National, Highers and Advanced). SQA is involved in both, internal and external examination and allows schools device their own summative assessment to suit local needs (SQA, n.d.).

<i>Sweden</i>	The national curriculum included goal and knowledge-based criteria where teachers could assess their students based on these. The country introduced a new curriculum in 2011, attempting to cover issues on summative assessment while promoting formative assessment, following the promise of great effects on students (Hirsh, 2020).	Teachers are seen as capable and responsible to independently assess their students (with no external examinations) and decide on the grades awarded, which are high-stakes for access to further education (Hartell and Strimel, 2019).
<i>New Zealand</i>	SBA is used to allow teachers to adapt their assessment tasks to cover their students' needs satisfying summative and formative purposes by using a range of activities that provide feedback and guidance to students (Yates and Johnston, 2018). SBA is seen as allowing synergies between formative and summative purposes, as it can promote the development of formative assessment due teachers control of the assessment process (Yates and Johnston, 2018).	New Zealand Qualifications Authority (NZQA) run examinations called National Certificates on Educational Achievement for secondary students. Schools also submit portfolios of students work to NZQA on Technology, Visual arts, Design and visual communication and Education for sustainability (NZQA, 2023).

Overall, these countries have different experiences using SBA which highlight the challenges, possibilities and difficulties of the process. They also show that the process is understood in differently, following the national culture and needs, which makes relevant the study of the assessment in each country in its specificities.

Appendix B: Documents reviewed during preliminary study

<i>Source</i>	<i>Information given</i>	<i>Notes</i>	<i>Access</i>
1. <i>School self-created assessment decree</i>	School assessment regulations. Reviewed during July 2019 and ready in November 2019. It covers the new Decree 67/2018.	To analyse its specifications in relation with Decree 67/2018 (that is to say: is it complete? does it cover all the points included in the policy?) To be used in school context section.	Access granted to be obtained during data collection.
2. <i>Schoolbooks</i>	They allowed me to make registers and take notes about assessment methods and activities registered by hand by teachers in each subject in each class in the Secondary Level.	Some teachers describe with higher detail their activities than others. High presence of tests.	Access granted to take notes. In any case my notes or registers cannot contain students' names.
3. <i>Classes grades reports</i>	They present all the grades averages of each student in all their subjects per class.	I reviewed one, which contained 45 rows for students' identification where only the numbers were legible (names were erased), 17 columns for subjects' grades, and final columns for 'general average', 'Religion' and 'attendance percentage'.	Access granted to obtain printed copies during fieldwork. They will not contain students' names on them. Not used for the final data set.
4. <i>Teachers planning</i>	Class to class objectives for the year, activities and assessment planning, made 3 months before the academic year start.	They are created, sometimes, before the teacher knows their students. They are not updated during the year. There are differences between what teachers plan and what they actually do in the classroom.	Will not be used as they don't give accurate information about classroom assessment (methods or purposes).

Appendix C: Students' interview schedule (showing changes implemented after preliminary study)

General data: name (or pseudonyms), age, class.

Code colour on the basis of piloting: Green for additions or movements within the questionnaire structure, red for eliminated material and black for questions that remains without changes.

Questions' item	Suggested questions
Assessment purposes	<p>In your opinion, what are the objectives of the assessments that you have? Do assessments help you to understand/know about your learning? Do you believe that teachers apply remedial strategies when you don't perform well? Do you have space in the class, to reflect about what you are learning? <i>Do you think this is important?</i> <i>How do you evaluate your classes? How you evaluate the assessment strategies that your teachers use? (classes and assessment were separated into 2 questions)</i></p>
Assessment methodologies	<p>What type of assessment do you have the most? Tests, oral presentations, others? <i>Do you think it allows you to copy?</i> What kind of items do you answer in a typical test or assessment? What are the points distribution? <i>In terms of items, they were clearly written? Did you know the points assigned? Do you receive clear instructions? Is the assessment previously scheduled? What do you think about the assessments' frequency? Why?</i> <i>What do you think your teachers attempt to assess you? Why? Can you give me some examples? Eliminate, it is repetitive.</i></p>
Assessment methods criteria	<p>What do you expect/want to be assessed on in the school? Do you have experiences of co-assessment or self-assessment? Do you know the learning goals/expectations for your subjects? How? Do you think that your teachers modify the class structure and assessment when needed? Why? <i>What do you think about how the points that are assigned in your assessments? Eliminate, it is repetitive.</i></p>
Assessments' feedback	<p>What type of feedback do you receive? (oral, written). Would you assess that feedback as positive/negative? Public/private? Focused just on achievement (high/low) or others? Immediate or delayed? focused on correct/incorrect? <i>What happens when you don't understand something? Moved from grades differences to feedback.</i></p>
Grades differences and outcomes	<p><i>Does the assessment that you undertake measure what you have learnt? Why or why not? Do you think that your grades reflect what you have learnt? Do you think that grades are important? Why?</i> How do you self-assess your achievement in class? (low, medium, high). Why? <i>Do you think that there is different treatment towards students, based on their grades or others personal characteristics, as gender?</i> In your opinion, what are the reasons linked with the classroom settings that could have an impact on grades differences? <i>What aspects of classroom assessment concern you? Why?</i></p>

	<p>Do you receive special strategies or help when need it? Eliminate, repetitive.</p> <p>Do you think that the assessment methods used by your teachers have any impact on your learning and grades? Why or why not?</p> <p>Do you think that is necessary to unify assessment criteria?</p>
<p>Personal perspectives</p>	<p>What do you think of the classroom assessment strategies? Eliminated, repetitive.</p> <p>Would you make any changes to the assessment process? Why?</p> <p>Can you tell me 3 things that you would change? (expansive)</p> <p>What do you assess as 'learning'? Why?</p> <p>Would you evaluate the assessment strategies used in your classes as appropriate? Eliminated, repetitive.</p> <p>What do you think of repeating students?</p> <p>What do you think about the impact of grades on university entrance?</p> <p>What do you think of classroom assessment practices in terms of its justice?</p>

Appendix D: Students' final interview guide for main study

General data: name (or pseudonyms), age, class.

Questions' item	Suggested questions
Assessment purposes and methods	<p>What assessments are you involved in?</p> <p>In your opinion, what are the objectives and purposes of the assessments that you have?</p> <p>Do assessments help you to understand/know about your learning? What kind of classroom assessment method is most helpful to support your learning?</p> <p>Do you believe that teachers apply remedial strategies when you don't perform well?</p> <p>Do you have space in the class, to reflect about what you are learning? Do you think this is important?</p> <p>How do you evaluate your classes? How you evaluate the assessment strategies that your teachers use?</p> <p>What type of assessment do you have the most? Tests, oral presentations, others?</p> <p>Do you think it allows you to copy?</p> <p>What kind of items do you answer in a typical test or assessment? What are the points distribution?</p> <p>In terms of items, they were clearly written? Did you know the points assigned? Do you receive clear instructions? Is the assessment previously scheduled? What do you think about the assessments' frequency? Why?</p> <p>What do you expect/want to be assessed on in the school?</p> <p>Do you have experiences of co-assessment or self-assessment?</p> <p>Do you know the learning goals/expectations for your subjects? How?</p> <p>Do you think that your teachers modify the class structure and assessment when needed? Why?</p>
Assessments' feedback and grades	<p>What type of feedback do you receive? (oral, written).</p> <p>Would you assess that feedback as positive/negative? Public/private? Focused just on achievement (high/low) or others? Immediate or delayed? focused on correct/incorrect?</p> <p>Does the assessment that you undertake measure what you have learnt? Why or why not? Do you think that your grades reflect what you have learnt? Do you think that grades are important? Why?</p> <p>How do you self-assess your achievement in class? (low, medium, high). Why?</p> <p>Do you think that there is different treatment towards students, based on their grades?</p> <p>In your opinion, what are the reasons linked with the classroom settings that could have an impact on grades differences?</p> <p>What aspects of classroom assessment concern you? Why?</p> <p>Do you think that the assessment methods used by your teachers have any impact on your grades? Why or why not?</p> <p>Do you think that is necessary to unify assessment criteria?</p>
Assessment ethics	<p>Would you make any changes to their characteristics? Why?</p> <p>Can you tell me 3 things that you would change?</p> <p>What do you assess as 'learning'? Why?</p> <p>What do you think of repeating students?</p> <p>What do you think about the impact of grades on university entrance?</p> <p>What do you think of classroom assessment practices in terms of its justice?</p>

Appendix E: Teachers' interview schedule (showing changes implemented after preliminary study)

General Data: Name or pseudonyms and subject.

Code colour: Green for additions or movements within the questionnaire structure, red for eliminated questions and black for questions that remains without changes.

Questions' item	Suggested questions
Assessment purposes	What objectives can you identify and tell me about the formative and summative? assessments that you design and conduct with your students?
Assessment methodologies	<p>¿What type of assessment do you use with your students? (this category includes tests with true and false, multiple options, short questions, questionnaires, etc.). Are those assessments formal or informal (spontaneous)?</p> <p>Can you give me some examples of the items used when you prepare assessments?</p> <p>¿How do you decide the points distribution? (Calibration) Do you believe this calibration helps to measure what you want to measure in your assessments?</p> <p>Do you assess students by performance? Eliminated, everything that one students does is considered for pilot teachers as 'performance'.</p> <p>What kind of specific activities do your students develop? (such as writing, speaking, reading, creating products projects, reports. The focus is on assessment that differs from test).</p> <p>What reasons support your decision to use these assessments?</p>
Assessment methods criteria	<p>If it is true that the frequency and the quantity of grades is, in part, decided by the school policy and academic coordination, would you make some comment in this regard?</p> <p>During this year, have you used co-assessment or self-assessment? Why?</p> <p>Did you share the learning objectives and your expectations with students?</p> <p>What are the cognitive levels evaluated in your assessments?</p>
Assessments' feedback	<p>Do you give feedback to your students?</p> <p>Is feedback focused on the positive or negative? Oral or written? How many time do you need to give feedback?</p>
Grades differences	<p>In your opinion, what you would say about the cause of grades differences among your students?</p> <p>How do you analyse students' achievement and if the classes were successful?</p> <p>Do you apply special strategies when your students perform below expectation?</p> <p>Do you think that assessment strategies have an impact on students' grades?</p> <p>What is your opinion about repeating students? and about the school grades impact on university entrance?</p>
Personal perspectives	<p>What changes would you make to how assessment works in the schools? Is there any aspect of assessment that concerns you?</p> <p>What do you think of classroom assessment practices in terms of its justice?</p> <p>Does the assessment system prompt any unintended consequences or ethical concerns in your view?</p> <p>Are there any other comments or additional information that you would like to give and was not included in this questions?</p>

Appendix F: Teachers' interview guide for fieldwork

General Data: Name or pseudonyms and subject.

Questions' item	Suggested questions
Assessment purposes and methods	<p>What objectives can you identify and tell me about the formative and summative? assessments that you design and conduct with your students?</p> <p>¿What type of assessment do you use with your students? (this category includes tests with true and false, multiple options, short questions, questionnaires, etc.). Are those assessments formal or informal (spontaneous)?</p> <p>Can you give me some examples of the items used when you prepare assessments?</p> <p>¿How do you decide the points distribution? (Calibration) Do you believe this calibration helps to measure what you want to measure in your assessments?</p> <p>What kind of specific activities do your students develop? (such as writing, speaking, reading, creating products projects, reports. The focus is on assessment that differs from test).</p> <p>What reasons support your decision to use these assessments?</p> <p>If it is true that the frequency and the quantity of grades is, in part, decided by academic coordination, would you make some comment in this regard?</p> <p>During this year, have you used co-assessment or self-assessment? Why?</p> <p>Did you share the learning objectives and your expectations with students?</p> <p>What are the cognitive levels evaluated in your assessments?</p>
Assessments' feedback and grades	<p>Do you give feedback to your students?</p> <p>Is feedback focused on the positive or negative? Oral or written? How many time do you need to give feedback?</p> <p>In your opinion, what you would say about the cause of grades differences among your students?</p> <p>How do you analyse students' achievement and if the classes were successful?</p> <p>Do you apply special strategies when your students perform below expectation?</p> <p>Do you think that assessment strategies have an impact on students' grades?</p> <p>What is your opinion about repeating students? and about the school grades impact on university entrance?</p>
Assessment ethics	<p>What changes would you make to how assessment works in the schools?</p> <p>Is there any aspect of assessment that concerns you?</p> <p>What do you think of classroom assessment practices in terms of its justice? Does the assessment system prompt any ethical concerns in your view?</p> <p>Are there any other comments or additional information that you would like to give and was not included in this questions?</p>

Appendix G: SoE Research Ethics form

Student Name: Tamara Cepeda

Proposed research project: 'This is not a fair system': Exploring Classroom Assessment Practices at a Secondary School in Chile

Funder: Financial support from the Programme of Advance Human Capital, National Commission of Scientific and Technological Research of Chile (CONICYT)/Ministry of Education (Mineduc), Chile.

Discussant for the ethics meeting: PhD candidate Hugo Parra

Names of supervisors: Professor Sally Thomas, Dr. Shelley McKeown Jones

Has your supervisor seen this submitted draft of your ethics application? Y

Please include an outline of the project or append a short summary:

Classroom assessment procedures helps to acquire and process evidence for analysing and improving students learning, mainly by measuring performance (Briggs and Ellis, 2013; Lambert and Lines, 2000). In the case of Chile, assessment procedures are reflected by internal school assessment regulations and assessment instruments, such as tests, and grades. There are two contextual factors to consider here. Firstly, at government level, the Chilean assessment policy allows schools to create their own assessment regulations, based on the policy guidelines, which have been recently updated (Mineduc Decree N° 68, 2018). Secondly, teachers use self-designed summative assessment instruments to evaluate their students. Assessment actions are led by teachers in order to obtain and interpret information about students learning, adopting decisions to promote their progress and providing feedback about the learning process (Mineduc Decree N° 68, 2018). Consequently, assessment instruments could be school-level designed or teacher-designed, depending on the respective school's specific assessment regulations and practice. In the new Chilean public management orientation, assessment procedures become the centre where resources and results meet, making relevant grades as learning results evidence. In this light, grades become a core topic in assessment procedures and define much of classroom life. At Chilean secondary school level, classroom assessment practices are contingent on several factors. Firstly, the possibility that students' can be forced to repeat the academic year. Beyond this their school grades become a percentage of the final score that will (dis)allow them entrance to university and relatedly access to scholarships. Thus, grade-based assessment is potentially related to the future curtailment of possibilities; with the attendant absolute dichotomous categorisation of students as simply 'high' or 'low' achievers; with the negative impact of grades on school engagement and also on consequences for disadvantaged students (Baird et al., 2014; Bloom, Hastings and Madaus ,1971; Brooks, 2002; Sánchez and Jara, 2019; Poorthius, Juvonen, Thomaes, Denissen, Orobio de Castro and Van Aken, 2015; Stiggins and Conklin, 1992). Despite these assessment procedures being mainly decided by teachers, some researchers claim in the literature that teachers do not take responsibility for students' difficulties, attributing these difficulties to external causes such as students' lack of effort, ability or lack of due parental support (García and Paredes, 2010; Mizala and Romanguera, 2000; Sánchez and Jara, 2019). As stated earlier, the Chilean policy gives space and freedom for schools, directors and teachers to shape their assessment practices. In this regard, the US literature shows

that 'grading methodology, the frequency of grading, homework assignments, types of questions, etc., are largely ungoverned by policy' and depends on school and teachers' decisions (Stiggins and Conklin, 1992, p. 151). Therefore, it is original to examine how assessment policy works in one Chilean school, in the sense of the summative assessment variety and the pedagogical ideas supporting these practices. Specifically, in a context where teachers design their summative assessment instruments using different criteria and ideas, which is increasingly interesting when the new assessment decree demands that schools make changes in March 2020. With this in mind, the ways in which schools and teachers will react and the changes that are decided upon becomes a significant area of policy implementation.

Globally, research on in-classroom summative assessment is regarded as limited and based upon sparse available information, as the focus is placed usually on formative assessment (Baird et al., 2014; Lambert and Lines, 2000; Stiggins and Conklin, 1992; Young and Kim, 2010). This research topic appears to be receiving fewer resources when compared with standardised tests with accountability purposes (Harlen, 2004). Therefore, we know very much about the numbers generated by external standardised assessment procedures conducted within schools, but the same does not apply to what is happening with the classroom summative assessment procedures occurring between students and teachers. This reinforces the need for studies focused on the shifting dynamics that are constantly taking place in the day-to-day school life, namely, how assessment practices and specifically internal summative assessments are integrated into educational systems (SWAAP, 2017).

For that reason, this research aims to analyse in-depth three stages: (1) an overview of the summative and formative assessment practices and methods used in one typical Chilean secondary school as guided by national and school-created decrees, (2) students' and teachers' perspectives of the summative assessment process in the classroom and their attendant positive/negative consequences and finally the (3) ethics involved in the procedures. Therefore, the qualitative methods proposed for this study are designed to explore the summative assessment practices. The stages constitute the research objectives and questions as follows:

- a. To analyse and describe how teachers conduct summative assessment practices in a chosen secondary school and the subsequent results.
- b. To explore students' and teachers' perspectives about the summative assessment practices conducted in one Chilean secondary school.
- c. To describe the implications and challenges related with the assessment practices in terms of fairness and ethics.

RQ1. What are the characteristics of the summative assessment practices that teachers design and conduct in the chosen secondary school in Chile?

RQ2. What are the reasons behind the teachers' assessment methods' decisions and how are assessment practices perceived by teachers and students in terms of their outcomes and future life implications?

RQ3. How is school assessment practice in one Chilean secondary school linked by teachers and students with broader goals of ethics and fairness within the educational system?¹⁰

¹⁰ Some aspects of these research questions were developed and changed during the progress of the study.

Research phases:

1. The pilot study will involve analyse the availability and format of school documents, teachers' annual assessment meeting participation and semi-structured interviews to explore teachers' conceptions of assessment and the nature of school based assessment. This phase will help to identify the information needed and available within the school context in order to answer the research questions and also determine the documentation access for the fieldwork. This pilot will be undertaken in November and December 2019. In this time the academic year is reaching a conclusion and all the tensions of the closure procedures, grades, final outcome (fail or pass) and the grades average impact on university entrance for leavers (Year 12 students) are at their peak level. Document analysis will look for evidence of how schools in Chile enact the national assessment policy and how that is reflected in day-to-day practice (school assessment decree, teachers planning, assessment instruments, grades register books). In this time all the schools in the country must also prepare for the requirements of the new decree of assessment implementation starting in March 2020. Interviews will be conducted with 2 teachers and 2 students and will last approximately 60 minutes. The piloting process has been thought of as bringing attention to the ethical process of this research with participants and stakeholders, considering their points of view and collaboration (BERA, 2018). This will also contribute to guiding and planning the research design and the data collection process for the main study.
2. The data collection process (main study planned for the first semester of 2020) will involve physically obtain documents for analysis, as previously selected during the pilot and also 20 interviews with ten teachers and ten students respectively. The documents will be obtained in formats previously agreed with the Head of School (computer files, pictures, printed versions), protecting students' names. The participation in teachers' meetings and movement/circulation of the researcher around the school during the fieldwork will be also discussed and negotiated with the Head of School during the piloting to clearly demarcate the zones allowed to circulate and work.

Ethical issues discussed and decisions taken:

1. Researcher access/exit:

Data collection for piloting will be in one School in Santiago, Chile, to which access has already theoretically been granted. This school has been selected for its interest in participating in the research process in order to improve its assessment practices, as they have to review their internal assessment decree because of a new decree coming from the Ministry of Education (Decree N° 67, 2018). There is no involvement of the research funder (Conicyt/Mineduc) in enabling or allowing access to this Chilean school. The conditions of access and exit will be revisited and negotiated with the gatekeeper prior to the commencement of the fieldwork (data collection) in November/December 2019 (Brooks, te Riele and Maguire, 2014). In this light, there are multiple responsibilities and sensitivities to consider, such as, for example, the responsibility to colleagues, students and their parents (Brooks, te Riele and Maguire, 2014). The selected school is located in a district well known by the researcher, as she was raised and worked there on a half-time (0.5)

contract for one and a half years in the past. This situation to some extent mitigates against the sense of intrusion, as the community knows the researcher. Some of the teachers were colleagues of the researcher in the past, but the students being part of the secondary level since November 2019 have never been taught by her. Former colleagues could be potentially interviewed, but there is no friendship or any type of relationship with them in the present. Nonetheless, this previous acquaintance will be assessed during the volunteers' selection and during the interviews as a factor which may have an impact in the data gathered.

Access to documents will be done respecting students' anonymity. Grades reports printed will not contain names of students and they will not imply use for individual analysis in any case. Markbooks information used will not be related with students' personal information. Notes will be taken based on teachers' assessment activities, as access to markbooks is confined to the school. This procedure does not entail any conflict with Chilean school regulations and was reviewed with the gatekeeper. However, the researcher and gatekeeper agreed that the grades analysis will require explicit parental consent. Informed consent will provide parents with information explaining the proposed research project and ask them to complete an opt-in form giving permission for the student to take part in the grades analysis stage.

Research in this study is considered more humanistic and thus involves mutual trust (Brooks, te Riele and Maguire, 2014). This will help to conduct protocols and procedures which are socio-culturally appropriate, and respect participants' culture(s) (BERA, 2018). This includes other benefits, such as the researcher's capacity to engage with the 'everyday jargon' and idiomatic language used in the Chilean context (especially by adolescents in certain social contexts), and concomitantly being able to make follow up replies and create richer data in the interview process (Brannick and Coghlan, 2007). There are also benefits regarding access to documentation, data, people and meetings inside the school (Brannick and Coghlan, 2007), creating a school contextual-adequate data collection process, thoughtfully designed, that will subsequently facilitate the main study during 2020 that will involve a study of the changes to the assessment policy implementation prompted by Mineduc's new policy. The researcher will ensure, upon saying farewell, all that participants are duly thanked for their key role in this research. In order to make the researcher's role abundantly clear and explicit, she will introduce herself and will always use the University of Bristol lanyard during fieldwork activity. Contact with participants on social media or social events outside school will be avoided. As has been stated before, the points covered for the Ethics form for pre-fieldwork will be reviewed and negotiated with the Head Teacher, as the main gatekeeper, during an initial meeting in November 2019, which will demarcate formal aspects such as access (Brooks, te Riele and Maguire, 2014). Documents will be obtained detailing fully the conditions for access and formats established by the Head Teacher during the piloting procedure.

2. Information given to participants:

Those directly involved in the research will be informed of the general objectives and procedures of the research both during piloting in November 2019, and during fieldwork in 2020. After a first meeting with the gatekeeper, teachers will be informed of the conditions established by the Head of School, as set out in the following table:

Group/person involved:	Information
1° Gatekeeper (Head of School)	Meeting in November 2019. Review of the research process, ethics and safeguarding procedures. Review of documents: information sheet to participants and informed consent for piloting and fieldwork. Attention of concerns and questions. Second meeting in 2020 for fieldwork, revisiting ethics and procedure. In both situations (piloting and fieldwork) the Head of School will receive a participation information sheet.
2° Teachers	The research project will be explained orally in the terms agreed with the Head of School for phase 1 and phase 2. Interview volunteers will also receive a participant information sheet.
3° Students	They will be orally informed by the researcher at a specific time (providing Head of School allows) in order to disrupt as little as possible their classes. The objective is to identify volunteer participants at the secondary level for the pilot (N=2) and for fieldwork (N=10). Interview volunteers will receive a participant information sheet. Volunteers will be registered on a list.
4° Parents	Piloting (Phase 1) involves interviews with 2 students of a legal age (over 18). Parents/legal guardians will be still informed with the participant information sheet. Volunteers will be registered on a list. Interview volunteers' parents/legal guardians will receive a participant information sheet for fieldwork interviews and grades analysis (Phase 2).

The communication process could differ from this plan, as it requires the gatekeeper's approval. Otherwise, the piloting planning will be strictly following the school requirements. Participants of Phase 1 will not take part in Phase 2. Participants will be selected attempting to be gender balanced and with no additional requirements other than being part of the school community. The list of volunteers will serve in case of participants' withdrawal in order to choose a new participant.

3. Participants right of withdrawal:

All participants will be given clear knowledge and understanding of their right to take part in the pilot or not, as volunteers, and of their right to withdraw from the research piloting and data collection at any point during the research process up until one month after the interview. They will be allowed to withdraw without needing to provide an explanation (BERA, 2018). In the case of withdrawing after the interview process, the right to be forgotten will be considered (GDPR, 2018). This involves the possibility to erase participants' personal data and to cease further dissemination of the data must be allowed

The participants' right of withdrawal will be explained verbally to the teachers and students and will be also contained in the informed consent. The willingness to participate in the project will be considered the main starting point. As the school has several teachers and the students at secondary level, the withdrawal of one participant can be easily solved by finding another volunteer. This will be clearly explained in order to reduce the pressure on possible participants as they can

freely choose to take part in the piloting and data collection or not, knowing that their withdrawal would not have a negative impact on the overall project.

4. Informed consent:

Teachers, students and interview volunteers will be given information relevant to the research topic to enable them to make an informed decision as to whether to participate or not in the piloting and data collection process. Informed consent will be obtained from students, underage students' parents and teacher participants in all cases.

Group/person involved:	Informed consent characteristics
1° Gatekeeper (Head of School)	Informed consent will be required to access the school for document analysis, assessment meetings participation and teachers'/ students' interviews. This will be done for piloting during 2019 and fieldwork during 2020.
2° Teachers	Informed consent will be required for interviews during piloting (Phase 1) and fieldwork (Phase 2) and for the analysis of documents created by them (planning, assessment instruments).
3° Students	Informed consent will be required for all students (over 18) for the interview process and grade analysis. Parents of underage students (under 18 according the Chilean law) will be required to consent only after students express their will to participate. The underage group age range is between 13 to 17 years old in the Chilean secondary school system (Year 9 to Year 12).
4° Parents	The researcher and gatekeeper agreed that the grades analysis will require explicit parental consent. Informed consent will provide parents with information explaining the proposed research project and asking them to complete an opt-in form giving permission for the student to take part in the grades analysis section of the research.

Informed consent for participants will include information about how the data will be stored, used and represented by the researcher, and will be obtained with a participant information sheet about the project. They will also receive a verbal explanation, as, in practice, participants may not fully read or understand the information of the consent form (Hammersley and Traianou, 2012). All the students participating in this research will be able to decide if they want to be accompanied by a chaperone (parent/legal guardian or classmate) during the interview. The interviews will be conducted within the school, following all the internal regulations.

5. Complaints procedure:

Participants will be given information about the complaints procedure, in terms of how they can make any complaint and/or present a doubt or inquiry during the piloting and fieldwork process. Contact information about the Head of School (gatekeeper), the PhD student (the researcher), supervisors and Graduate School of Education will be also provided to the participants on the information sheet, which will be distributed in their Spanish versions. As the participants' first language is Spanish, the researcher will analyse with the Head of School the possibility of translation in the cases of complaints, being supported by an English school teacher. This will help in case any

participant would like to make a complaint directly with the School of Education. This will increase their right to complaint during the piloting process and from that moment forward.

6. Safety and well-being of participants/ researchers:

Suggestions for a safe private place within the school, battery supplies and technical information to secure the success the piloting interview process will be considered. The space to conduct the interviews will be analysed with the Head of School and participants, in order to make them feel comfortable and to be able to speak privately and confidentially in a quiet room within the school. The researcher will try to make participants feel comfortable and supported during the whole interview piloting process, adjusting the schedule to their availability and to school procedures and rules. There are no further safety or wellbeing issues for the researcher, as she is an experienced teacher and manages the security procedures in case of emergency (fire, earthquake, evacuation, etc.). In any case, the safety and well-being of participants and the researcher will be assessed with the School Security Manager, who is a compulsory figure within Chilean schools. The country and zone for piloting and fieldwork are not a conflictive setting and there is no high risk of diseases involved in it.

7. Anonymity/ confidentiality:

The pilot and fieldwork participants will be labelled with pseudonyms chosen by them in the interviews, and their identities will therefore not be identifiable. However, within the school context teachers and students will probably know who is participating in the piloting and fieldwork process, as they will volunteer to take part in a process that will involve different stakeholders. Consequently, anonymity cannot be guaranteed. The stored information will not use names and the information gathered during the piloting and fieldwork will be securely stored by password on personal laptops and University drives. Information will be used only for the purposes given and known by participants. The school's name will remain confidential in the pre-fieldwork and piloting, which will be explained to the Head of School and participants. For the main study, the presentation of contextual information that may enable the school to be identified will be duly minimised. Nevertheless, given some contextual information about the school present in written formats, school anonymity cannot be totally guaranteed. This will be explained and analysed with the Head of School. To sum up, participants will be known and identifiable within the school dynamic, as they will volunteer and participate in the interview process, from signing an informed consent form to the interviewing process. However, they will be not identifiable in the final written thesis. The researcher will make strenuous efforts in order to protect the school name as far as possible.

8. Data collection:

The interview process will ask for teachers' and student's experiences of the classroom summative assessment procedures: methods, results and ethics. Several authors have reported the importance of considering teachers' and students' beliefs about assessment and learning as a very important part of possible changes in assessment strategies (Baird et al., 2014; Torrance and Pryor, 2001; Brown et al., 2009; Brookhart, 2001; Willis, 2008; Tapan, 2001; Carless, 2005; Marshall and Drummond, 2006). Educational settings hold different meanings to different people involved in

assessment practices. \Therefore, the interviews will be used to understand those meanings in teachers' and students' voices. One of the main focuses of the piloting interviews will be to interpret the assessment principles that guide teachers' actions and students' learning judgements based on these principles and strategies. This will be done only in audio-recording and transcribing the content of the interview afterward in their original language (Spanish).

The piloting and fieldwork process will involve the analysis of school documents related to assessment procedures characteristic within the school (internal regulation, teachers planning, register books) and assessment meetings observation (if available). The piloting process will help to identify the sources and information available in order to answer the research questions during the data collection process. As was stated before, in this way the data collection process will be adequately designed. Complexities during the data collection process will be solved with well informed and situated judgements. This process will consider respect for the person, the production of knowledge, democratic values and the quality of educational research (BERA, 2018).

9. Data analysis:

Interview transcriptions and documents will be reviewed in terms of the capacity to help to answer the research questions. Data analysis will involve thematic analysis and the NVivo software in order to guide qualitative analysis. Documents will be content analysed. As can be seen, the piloting and fieldwork processes will contain information that will serve for the effective triangulation of data coming from different sources, thus enhancing the validity and reliability of the future data analysis (Merriam, 1988). In the case of doubts during the data analysis, peer review will be used. This is relevant to consider that the teaching experience of the researcher could potentially be a source of bias. The broad aims of social science research to maximise benefit and minimise harm will be always foremost (BERA, 2018). The research project will be seen as a study that could help the school by allowing their participants an opportunity to reflect on summative assessment procedures. In the case of controversial or sensitive findings, these will be discussed with the gatekeeper in order to agree a fair publication strategy.

10. Data storage:

Data will be stored safely using both the University of Bristol encrypted network and personal laptop with password protection. Data collected will be used and held only for the purposes informed to the participants in the informed consent. As was explained before, the data held by the researcher for research process design and review will be anonymised and will not contain participants' names.

11. Data Protection Act:

All the procedures including the information collected in the study will be adjusted to conditions indicated by the UK Data Protection Act regulations in terms of being used fairly, lawfully and transparently (as follows in <https://www.gov.uk/data-protection>). This research will also consider the General Data Protection Regulation (GDPR, 2018; as follows in <https://eugdpr.org/the-regulation/>). Therefore, the school data protection policy will be considered during all the piloting and fieldwork processes. There is no race, ethnic background, political opinions, religious beliefs,

trade union memberships, genetics, biometrics, health or sex life or orientation involved in this research.

12. Feedback:

The school gatekeepers will not have access to any raw data from interviews or observed meetings as they could identify participants on their personal opinions. However, they will be able to read the final thesis. As set out above, the final thesis will be duly anonymised in order to protect participants' identity, opinions and possible sensitive topics expressed. The results of the research will be informed to the participants in the form of a final report only if required. They will be given access to a pdf file with general research results using e-mail as contact resource. This possibility will be contained in the informed consent form. Feedback conditions about the prohibition of the dissemination of results in the academic community will be considered and stated in the informed consent. Confidentiality in this regard will be strictly protected.

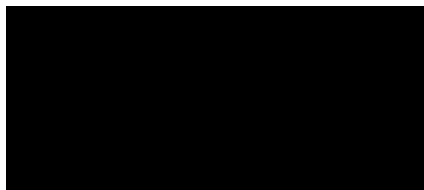
13. Responsibilities to colleagues/ academic community:

The researcher will act all the time in accordance with the University of Bristol 'Ethics Regulations' responding to the commitment of protecting the integrity and reputation of educational research by ensuring high academic standards (BERA, 2018). There is also respect about the excellence involved in the Scholarship granted by the Ministry of Education of the Chilean government. As a result of the different countries involved in this research, a broader structure for ethical considerations will be employed (Bassegy, 1999; Stuchbury and Fox, 2009). However, the Ethics Code for professionals of education will be considered besides BERA guidelines, respecting both regulations in terms of educational research: British and Chilean. It will imply a respect for democracy, in terms of the freedom of expression, giving and receiving ideas and information and considering the responsibilities of participants in terms of veracity (Bassegy, 1999). The piloting and data collection process have been sensitively and thoughtfully designed considering the differences between cultural contexts in a number of ways (BERA, 2018). On the one hand, there are clear guidelines contextualised in the Global North/Western society and in the University of Bristol that are not always entirely transferable to the research field. On the other hand, there are cultural characteristics in the Global South and Chile, such as the common absence of formal ministerial, council or local ethics. Therefore, the researcher considered universal ethical principles that have been considered throughout the creation of this document (Brooks, te Riele and Maguire, 2014). The researcher intends to show high standards of professional integrity, academic skills and professionalism throughout the piloting and fieldwork process.

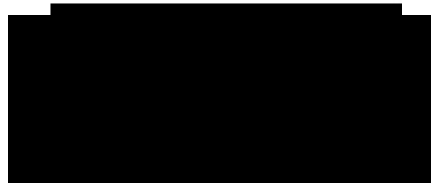
14. Reporting of research:

The results of this pilot study will be guiding the PhD research project and methodology, which will be presented in a progression panel during the first semester of 2020 at the University of Bristol. The findings of the piloting and fieldwork will be used and disseminated with journal publication purposes and/or conferences with no time limitations. Considering the international character of this research, it is expected that publications both in the original language (Spanish) and in English will be subsequently undertaken, in order to make it locally and internationally

available (BERA, 2018). However, this dissemination will not lead to a breach of the agreed confidentiality and anonymity. This use of data will be clearly stated in the informed consent sheet disseminated. The data collected will be also part of the final doctoral thesis, submitted during 2022. Findings will be always published protecting the participants' and school's identity as far as is possible, and being aware of the potential impact of the research findings for the school used in this case study. Findings will not be required by the financial sponsor, the Chilean Ministry of Education (Mineduc) and the National Commission for Scientific and Technological Research of Chile (Conicyt). The financial sponsor has no specific requirements on ethical guidelines or procedures to comply with beyond the elements already considered in this Ethics Form. Mineduc/Conicyt are not contractual sponsors for this research and they are neither related to its topic, nor to how the research will be conducted, or its results. Consequently, this research does not declare any conflicts of interests neither does it compromise the integrity of the research process.



Tamara Cepeda (Researcher)



Hugo Parra Muñoz (Discussant)

Date: October 7, 2019.

References:

- Bassey, M. (1999). *Case study research in educational settings*. McGraw-Hill Education (UK).
- Baird, J., Hopfenbeck, T. N., Newton, P., Stobart, G., & Steen-Utheim, A. T. (2014). State of the field review: Assessment and learning. *Oslo: Norwegian Knowledge Centre for Education*.
- Bera (2018). Ethical guidelines for educational research. https://www.bera.ac.uk/wp-content/uploads/2018/06/BERA-Ethical-Guidelines-for-Educational-Research_4thEdn_2018.pdf?noredirect=1
- Bloom, B., Hastings, T., & Madaus, G. (1971). Formative and summative evaluation of student learning.
- Brannick, T., & Coghlan, D. (2007). In defense of being “native”: The case for insider academic research. *Organizational research methods*, 10(1), 59-74.
- Briggs, J., & Ellis, V. (2013) *Assessment for learning, on Learning and teaching in secondary schools*. Learning matters. Sage publications.
- Brookhart, S. M. (2001) Successful students' formative and summative uses of assessment information, *Assessment in Education: Principles, Policy & Practice*, 8 (2), 153–169.
- Brooks, V. (2002). *Assessment in secondary schools: the new teacher's guide to monitoring, assessment, recording, reporting, and accountability*. McGraw-Hill Education (UK).
- Brooks, R., Te Riele, K., & Maguire, M. (2014). *Ethics and education research*. Sage.
- Brown, G., Kennedy, K., Fok, P., Chan, J & Yu, W. (2009) *Assessment for Student Improvement: Understanding Hong Kong Teachers' Conceptions and Practices of Assessment*, *Assessment in Education: Principles, Policy & Practice*, 16 (3), 347–363.
- Carless, D (2005) *Prospects for the Implementation of Assessment for Learning*. *Assessment in Education Principles Policy & Practice*, 12 (1), 39–54.

- Garcia Palomer, C., & Paredes, R. D. (2010). Reducing the educational gap: good results in vulnerable groups. *The Journal of Development Studies*, 46(3), 535-555.
- General Data Protection Regulation, GDPR (2018). <https://eugdpr.org/the-regulation/> on October 3, 2019.
- Gonzalez, T. E., Hernandez-Saca, D. I., & Artiles, A. J. (2017). In search of voice: Theory and methods in K-12 student voice research in the US, 1990–2010. *Educational Review*, 69(4), 451-473.
- Hammersley, M., & Traianou, A. (2012). *Ethics and educational research*. London: British Educational Research Association.
- Harlen, W. (2004). *A systematic review of the evidence of reliability and validity of assessment by teachers used for summative purposes*. EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.
- Lambert, D., & Lines, D. (2013). *Understanding assessment: Purposes, perceptions, practice*. Routledge.
- Marshall, B & Drummond, M J (2006). How teachers engage with Assessment for Learning: lessons from the classroom, *Research Papers in Education*, 21, (2), 133–149.
- Mineduc (Ministry of Education, Chile) 2018. Decree of qualification and school promotion N°67/2018. https://www.curriculumnacional.cl/614/articles-70983_archivo_01.pdf
- Mizala, A. & Romaguera, P. (2000) Determinación de factores explicativos de los resultados escolares en educación media en Chile, Centro de Economía Aplicada, Universidad de Chile, N° 85.
- Poorthuis, A. M., Juvonen, J., Thomaes, S., Denissen, J. J., Orobio de Castro, B., & Van Aken, M. A. (2015). Do grades shape students' school engagement? The psychological consequences of report card grades at the beginning of secondary school. *Journal of Educational Psychology*, 107(3), 842.
- Sánchez-Sánchez, G. I., & Jara-Amigo, X. E. (2019). Los docentes en formación y la comprensión de sus experiencias pedagógicas: la evaluación del aprendizaje. *Revista Internacional de Investigación en Ciencias Sociales*, 15(1).
- Stiggins, R. J., & Conklin, N. F. (1992). *In teachers' hands: Investigating the practices of classroom assessment*. Suny Press.
- Stutchbury, K., & Fox, A. (2009). Ethics in educational research: introducing a methodological tool for effective ethical analysis. *Cambridge Journal of Education*, 39(4), 489-504.
- System-Wide Analysis of Assessment Practices, SWAAP (2017). Australian council for Educational Research (ACER).
- Tapan, S (2001) A study of present learning assessment practices to ensure quality of primary education. Basic education studies, 2000-2001. Dhaka, UNESCO Office Dhaka.
- Torrance, H., & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British educational research journal*, 27(5), 615-631.
- Willis, J. (2008) Assessment for learning: a socio-cultural approach. In: 'AARE 2008 International education research conference: Brisbane: papers collection', compiled by P L Jeffrey. Melbourne, Australian Association for Research in Education.
- Young, V. M., & Kim, D. H. (2010). Using assessments for instructional improvement: A literature review. *education policy analysis archives*, 18, 19.

Appendix H: SoE Research Ethics form (updated version)

Student Name: Tamara Cepeda

Proposed research project: 'This is not a fair system': Exploring Classroom Assessment Practices at a Secondary School in Chile

Funder: Financial support from the Programme of Advance Human Capital, National Commission of Scientific and Technological Research of Chile (CONICYT)/Ministry of Education (Mineduc), Chile.

Discussant for the ethics meeting: PhD candidate Hugo Parra

Names of supervisors: Professor Sally Thomas, Dr. Shelley McKeown Jones

Has your supervisor seen this submitted draft of your ethics application? Y

Updated sections details after pilot study:

This ethics form was updated following the pilot study conducted between November 2019 and January 2020 in Chile. It contains changes in the project title, objectives, questions, in the sample size for interviews and the requirement for opt-out consent for grades analysis. The quote 'This is not a fair system' has been duly omitted from the project title in the Informed consent and Information sheet to participants in order to avoid biasing participants prior to participation.

The research objectives and questions as been reformulated as follows:

- a. To conduct an international literature review of teachers' summative and formative assessment practices in the classroom.
- b. To review the Chilean research context in terms of the historical development and current policy and practice of classroom assessment.
- c. To analyse and describe how teachers conduct assessment practices in a chosen secondary school and the subsequent results.
- d. To explore students' and teachers' perspectives about the classroom assessment practices conducted in one Chilean secondary school.
- e. To describe the implications and challenges related to the assessment practices in ethical terms.

RQ1. What are the characteristics of the classroom assessment practices that teachers design and conduct in the chosen secondary school in Chile?

RQ2. What are the reasons behind the teachers' assessment methods' decisions and how are assessment practices perceived by teachers and students?

RQ3. How is school assessment practice in one Chilean secondary school linked with broader ethical concepts within the educational system by teachers and students?

1. Additional considerations for information given to participants:

After a meeting with the gatekeeper prior to the commencement of fieldwork, teachers, students and parents will be informed of the conditions established by the Head of School. All students were orally informed about the project and data collection process by the researcher during November 2019. As volunteers were registered on a list and exceeded the participants

needed for the pilot, those outside the required sample for the study will have the possibility to participate in the interviews during the fieldwork. Year 9 students will be invited to participate during the fieldwork, as they started the secondary level in March 2020. All the students will receive a participants' information sheet including opt-out options for grades analysis and for the interview procedure respectively. All the parents/legal guardian of underage students in the secondary level will receive an information sheet, containing the opt-out option for the grades analysis and information about the interview process.

2. Additional considerations for informed consent:

Details about the opt-out written consent were added to all the information sheets that will be used in the project. The texts included were:

“In the case that the student (or parent/legal guardian) does not want to allow the use of its grades for statistical analysis in a contextual chapter, please complete the opt-out form on the following page and give it to the researcher.

- I do not authorise the anonymised use of my grades for statistical and contextual analysis
- I do not authorise the anonymised use of my pupil grades for statistical and contextual analysis.
- I do not agree to take part in the study.
- I do not agree that my pupil take part in the study.”

Opt-out consent was selected in order to avoid the data being unrepresentative for the secondary school level students sample. Parents of students over 18 years (legal age in Chile) will not be asked to sign any consent.

3. Safety and well-being of participants/researchers update:

The country and zone for fieldwork is not conflictive at the moment, after the national uprising which occurred recently in Chile. The risk of COVID-19 will be assessed in time, following the University of Bristol, UK government and Chilean Government's recommendations.

4. Data collection new considerations:

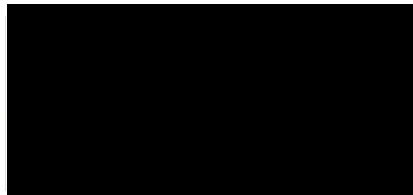
Access to documents will be done respecting students' anonymity. Grades reports printed will not contain the names of students and will not imply use for any individual analysis in any case. Markbooks information used will not be related with students' personal information. Notes will be taken based on teachers' assessment activities, as access to mark books is confined to the school. This procedure does not entail any conflict with Chilean school regulations and was reviewed thoroughly with the gatekeeper during the pilot and again in March 2020. The procedures agreed respect the Ministry of Education (Law 19.628, Protection of personal Data, Chile) requirements pertaining to the protection of personal data, as they are not sensitive data (ideological, religious, health, sexual and/or about disability). The researcher and gatekeeper agreed that grades analysis will require explicit written consent, giving participants the right to opt-out if they so wish. This considers the BERA (2018) requirements about consider the best interests of the students in their right to express their views in the first place, and considering additionally their parents/legal guardians. The information sheet and consent form will provide students and parents/legal

guardians with information explaining the proposed research project and ask them to complete an opt-out form in the case of not taking part in the grades analysis stage. This will comply with the requirements about giving participants information about the researcher, her role as researcher and the use of the data. Participants will have the opportunity to ask for their grades data elimination or cancelation if requested until one month after the data collection process is completed (BERA, 2018; Law 19.628, Protection of personal Data, Chile).

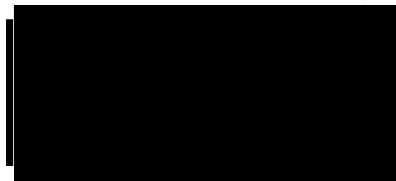
Considering the current worldwide pandemic situation, data collection will imply the possibility of being done using online platforms. This will be further analysed in time, following the University of Bristol, UK government and Chilean Government's recommendations.

5. Data Protection Act additional update:

All the procedures involved in fieldwork for this study have been updated under the Chilean Ministry of Education's recommendations and the Chilean Law for protection of personal data (Law 19.628).



Tamara Cepeda (Researcher)



Hugo Parra Muñoz (Discussant)

Last updated: March 17, 2020.

References

Mineduc (n.d). Protection of personal Data, based on Chilean Law N°19.628, <http://ayudamineduc.cl/politica-de-privacidad>.

Appendix I: Access to School consent

Head teacher access to school consent (English version)



Name of the project: Exploring Classroom Assessment Practices at a Secondary School in Chile

Student name: Tamara Cepeda.

Objective: To identify summative assessment practices conducted in the secondary level of your school as framed by national assessment policy and internal school assessment decrees. The assessment procedures will be analysed in terms of practice within the classroom, results obtained by students and reflected by them and their teachers and the ethical principles involved in the process.

Description: This project involves two phases. The pilot study will involve document analysis, teachers' annual assessment meetings participation and semi-structured interviews during November/December 2019. Document analysis will look for evidence of how schools enact the national assessment policy and how that is reflected in the day-to-day practice (school assessment decree, teachers planning, assessment instruments, grades register books). Interviews will be conducted with 2 teachers and 2 students. Data collection process (first semester 2020) will involve document analysis, as previously selected during the pilot and 20 interviews with teachers (10) and students (10). The documents will be obtained in formats previously accorded with the Head of School (computer files, pictures, printed versions). The participation in teachers' meetings and movement/circulation around the school will be also discussed and negotiated with the Head of School during the piloting.

The school stakeholder, represented by its Head Teacher, give consent for:

<ul style="list-style-type: none">● I give consent for the researcher, Tamara Cepeda, to enter the School as PhD student of the University of Bristol, with a schedule for piloting and data collection specified beforehand. This process will involve to check ethical and safeguarding procedures.
<ul style="list-style-type: none">● I give consent to access documents related with assessment practices within the School in the format previously agreed with the School.
<ul style="list-style-type: none">● I give consent for the selection and interview process of teachers and students of the secondary level, respecting the conditions established by the school.
<ul style="list-style-type: none">● I ask that the school remain anonymous as far as possible in the report of this research.

Mr. [REDACTED]
School Head Teacher signature

Date:

Head teacher access to school consent (Spanish version)



Nombre del proyecto: Explorando las prácticas de evaluación en el aula en un Educación Media en Chile.

Nombre del estudiante: Tamara Cepeda.

Objetivo: Identificar prácticas de evaluación llevadas a cabo en un colegio chileno en educación media, enmarcadas en la política de evaluación nacional y decretos de evaluación internos. Las prácticas evaluativas serán analizadas en término de los procedimientos llevados a cabo dentro de la sala de clases, los resultados obtenidos por los estudiantes y la reflexión de ellos y sus profesores al respecto, así como también los principios éticos involucrados en el proceso.

Descripción: Este Proyecto involucra dos fases. La primera, un estudio piloto, involucrará análisis de documentos, participación en consejos de evaluación anuales y entrevistas semi-estructuradas durante noviembre y diciembre de 2019. El análisis de documentos buscará evidencia de cómo los colegios implementan la política nacional de evaluación y cómo eso se refleja en la práctica evaluativa en el día a día (en el decreto de evaluación interno, en las planificaciones docentes, en los instrumentos de evaluación y en los libros de clases). El proceso de recolección de datos (1° Semestre 2020) involucrará análisis de documentos, según lo previamente seleccionado y autorizado durante el estudio piloto y 20 entrevistas con profesores (10) y estudiantes (10). Los documentos se obtendrán en formatos previamente acordados con el Director del Colegio (archivos computacionales, imágenes, versiones impresas). La participación en consejos de profesores y el movimiento/circulación en el colegio serán discutidos y negociados con el Director durante el estudio piloto. El Director del colegio da, de esta manera, consentimiento para:

<ul style="list-style-type: none">• Doy consentimiento para que la investigadora, Tamara Cepeda, entre al colegio como estudiante de Doctorado de la Universidad de Bristol, con una agenda para estudio piloto y recolección de datos especificada de antemano. Este proceso involucrará la revisión del proceso ético y de seguridad de la investigación.
<ul style="list-style-type: none">• Doy consentimiento para acceder a documentos relacionados con las prácticas de evaluación dentro del Colegio, en formatos previamente acordados con el colegio.
<ul style="list-style-type: none">• Doy consentimiento para la selección y proceso de entrevista de profesores y estudiantes de Educación Media, respetando las condiciones establecidas por el colegio.
<ul style="list-style-type: none">• Solicito que el colegio permanezca anónimo hasta donde sea posible en el reporte de esta investigación.
<ul style="list-style-type: none">• Proveeré en apoyo de un profesor de inglés en caso de la presencia de quejas sobre el proceso de investigación. Esto facilitará la comunicación directa entre participantes y la Universidad de Bristol, considerando que su lenguaje materno es español.
<ul style="list-style-type: none">• Comprendo que cualquier resultado controversial o de contenido sensible será discutido con la finalidad de acordar la estrategia de publicación más justa.


Firma y timbre de Director del Colegio

Fecha:

Appendix J: Information sheets

Participant information sheet for teachers (English version)

Name of the project: Exploring Classroom Assessment Practices at a Secondary School in Chile

Student name: Tamara Cepeda

Objective: To identify summative assessment practices conducted in the secondary level in your school as framed by national assessment policy and internal school assessment decrees. The assessment procedures will be analysed in terms of practice within the classroom, results obtained by students and reflected by them and their teachers and the ethical principles involved in the process.

Characteristics: This research is part of the thesis requirements leading to PhD in Education Degree in the University of Bristol, England. This project involves two phases. The pilot study will involve document analysis, teachers' annual assessment meetings participation and semi-structured interviews during November/December 2019. Document analysis will look for evidence of how schools enact the national assessment policy and how that is reflected in the day-to-day practice (school assessment decree, teachers planning, assessment instruments, grades register books). Interviews will be conducted with 2 teachers and 2 students within the school space. Data collection process (first semester 2020) will involve document analysis, as previously selected during the pilot and 20 interviews teachers (10) and students (10). Interviews will have a duration of 1 hour approximately. The documents will be obtained in formats previously accorded with the Head of School (computer files, pictures, printed versions).

You are asked to participate in an audio-recorded interview about assessment practices within the school, the practice, outcomes and ethics. Your consent implies to answer questions related with your experiences assessing students in your specific subject, the assessment methods used by you and the ideas supporting them, your opinions about students' results and the assessment system in general. The format of the interview is semi-structured, that means that the researcher will have some specific questions guide but the interview will also be conducted following your opinion and interests, and respecting the natural rhythm of the conversation. You will have the opportunity to choose the better schedule, to choose a pseudonymous, to review the interview transcriptions and to request a summary of findings. You will also have the freedom to withdraw until one month after the interview has been conducted by communicating verbally or by e-mail this decision to the researcher. In the case of withdrawal, the right to be forgotten by erasing all data related with your interview will be given.

The information will be used to answer the research questions of this project and will be stored protected by the University of Bristol encrypted system and also with password protection. Results will be used for journal publications, conference presentation and finally, in a written thesis submitted to the University of Bristol during 2022. In any format of presentation, the findings will include neither participants' real names nor the School name. Confidentiality in this regard will be strictly protected.

Complaint Procedure/Information request:

In case of complaint or information request, please contact:

Student	Tamara Cepeda	Tamara.Cepeda@bristol.ac.uk
---------	---------------	--

Supervisors (England)	Sally Thomas Shelley McKeown Jones	S.Thomas@bristol.ac.uk S.mckeownjones@bristol.ac.uk
Head of School (Chile)		

In case of need of translation process to make a complaint directly with the University of Bristol, please refer directly to the Head of School in order to obtain help in this regard.

Participant information sheet for teachers (Spanish version)



Nombre del proyecto: Explorando las prácticas de evaluación en el aula en un Educación Media en Chile.

Nombre del estudiante: Tamara Cepeda.

Objetivo: Identificar prácticas de evaluación llevadas a cabo en un colegio chileno en educación media, enmarcadas en la política de evaluación nacional y decretos de evaluación internos. Las prácticas evaluativas serán analizadas en término de los procedimientos llevados a cabo dentro de la sala de clases, los resultados obtenidos por los estudiantes y la reflexión de ellos y sus profesores al respecto, así como también los principios éticos involucrados en el proceso.

Características: Esta investigación es parte de los requerimientos de tesis para obtener el grado de Doctor en Filosofía(PhD) en Educación en la Universidad de Bristol, Inglaterra. Este proyecto involucra dos partes. El estudio piloto involucrará análisis de documentos, participación en consejos de profesores y entrevistas semi-estructuradas durante noviembre/diciembre 2019. El análisis de documento apunta a identificar la evidencia presente en el colegio que muestre cómo se implementa la política nacional de evaluación y cómo esto se ve reflejado en la práctica en el día a día (en el reglamento interno de evaluación, planificaciones de los profesores, instrumentos de evaluación y libros de clases). Las entrevistas considerarán 2 profesores y 2 estudiantes, dentro del espacio del colegio. El proceso de recolección de datos (primer semestre 2020) involucrará análisis de documentos, según los documentos seleccionados previamente en el piloto, y 20 entrevistas con profesores (10) y estudiantes (10). Las entrevistas tendrán una duración de 1 hora aproximadamente. Los documentos serán obtenidos en formatos previamente acordados con el Director del Colegio (archivos computacionales, imágenes, versiones impresas).

Se te invita a participar en una entrevista sobre prácticas de evaluación en el colegio, práctica, resultados y ética del proceso, que será audio-grabada. Su consentimiento implica responder preguntas relacionadas con sus experiencias evaluando a sus estudiantes en su asignatura, los métodos usados y las ideas que apoyan esas decisiones, su opinión sobre los resultados de sus alumnos y del sistema de evaluación en general. El formato de la entrevista es semi estructurado, eso significa que la investigadora tendrá una pauta con preguntas específicas que guían la conversación, pero ésta también seguirá el ritmo de sus opiniones e intereses, respetando el flujo normal de la conversación. Los profesores que participen tendrán la oportunidad de elegir el mejor horario para la entrevista, elegir un pseudónimo, revisar las transcripciones de su entrevista y solicitar un resumen de los resultados. Los participantes también tendrán libertad de retirarse del estudio hasta un mes después de que la entrevista haya sido conducida, comunicándolo verbalmente o mediante e-mail a la investigadora. En caso de retirarse del estudio, se entrega el derecho a ser olvidado, que consiste en eliminar toda la información relacionada con el participante. La información será usada para responder las preguntas de investigación de este proyecto y será almacenada protegida por el sistema encriptado de la Universidad de Bristol, que también usa protección con contraseñas. Los resultados serán usados para publicaciones en revistas académicas, presentaciones en conferencias y también para la tesis final de doctorado, que se presentará en la Universidad de Bristol durante el año 2022. En cualquier formato de presentación, los resultados no

incluirán ni el nombre real de los participantes ni el nombre del colegio. La confidencialidad en este punto será estrictamente protegida.

Procedimiento de quejas/información:

En caso de queja o solicitudes de información, por favor contactar:

Estudiante de Doctorado	Tamara Cepeda	Tamara.Cepeda@bristol.ac.uk
Supervisores (Inglaterra)	Sally Thomas Shelley McKeown Jones	S.Thomas@bristol.ac.uk S.mckeownjones@bristol.ac.uk
Director Colegio		

En caso de necesitar traducción para dirigir una queja directamente a la Universidad de Bristol, por favor contactar directamente el Director del colegio para obtener ayuda en el proceso.

Participant information sheet for over 18 students (English version)



Name of the project: Exploring Classroom Assessment Practices at a Secondary School in Chile

Student name: Tamara Cepeda

Objective: To identify summative assessment practices conducted in the secondary level in your school as framed by national assessment policy and internal school assessment decrees. The assessment procedures will be analysed in terms of practice within the classroom, results obtained by students and reflected by them and their teachers and the ethical principles involved in the process.

Characteristics: This research is part of the thesis requirements leading to PhD in Education Degree in the University of Bristol, England. This project involves two phases. The pilot study will involve document analysis, teachers' annual assessment meetings participation and semi-structured interviews during November/December 2019. Document analysis will look for evidence of how schools enact the national assessment policy and how that is reflected in the day-to-day practice (school assessment decree, teachers planning, assessment instruments, grades register books). Interviews will be conducted with 2 teachers and 2 students within the school space. Data collection process (first semester 2020) will involve grades and document analysis, as previously selected during the pilot and 20 interviews with teachers (10) and students (10). Interviews will have a duration of 1 hour approximately. The documents will be obtained in formats previously accorded with the Head of School (computer files, pictures, printed versions). Student opt-in consent will be required for grades analysis.

You are asked to participate in an audio-recorded interview about assessment practices within the school, the practice, outcomes and ethics. Your consent implies to answer questions related with your experience being assessed in different subjects, the assessment methods used by your teachers and the ideas supporting them, your opinions about your results and the assessment system in general. You will have the opportunity to choose the better schedule, to choose a pseudonymous, to review the interview transcriptions and to request a summary of findings. You will also have the freedom to withdraw until one month after the interview has been conducted by communicating verbally or by e-mail this decision to the researcher. In the case of withdrawal, the right to be forgotten by erasing all data related with your interview will be given.

The information will be used to answer the research questions of this project and will be stored protected by the University of Bristol encrypted system and also with password protection. Results will be used for journal publications, conference presentation and finally, in a written thesis submitted to the University of Bristol during 2022. In any format of presentation, the findings will include neither participants' real names nor the School name. Confidentiality in this regard will be strictly protected.

In the case that the student (or parent/legal guardian) does not want to allow the use of its grades for statistical analysis in a contextual chapter, please complete the opt-out form on the following page and give it to the researcher.

Complaint Procedure/Information request:

In case of complaint or information request, please contact:

Student	Tamara Cepeda	Tamara.Cepeda@bristol.ac.uk
Supervisors (England)	Sally Thomas Shelley McKeown Jones	S.Thomas@bristol.ac.uk S.mckeownjones@bristol.ac.uk
Head of School (Chile)		

In case of need of translation process to make a complaint directly with the University of Bristol, please refer directly to the Head of School in order to obtain help in this regard.

Opt-out consent form

- I do not authorise the anonymised use of my grades for statistical and contextual analysis.
- I do not agree to take part in the study.

Student name: _____

Class: _____

Signature: _____

Participant information sheet for over 18 students (Spanish version)



Hoja de información al estudiante mayor de edad

Nombre del proyecto: Explorando las prácticas de evaluación en el aula en un Educación Media en Chile.

Nombre del estudiante: Tamara Cepeda.

Objetivo: Identificar prácticas de evaluación llevadas a cabo en un colegio chileno en educación media, enmarcadas en la política de evaluación nacional y decretos de evaluación internos. Las prácticas evaluativas serán analizadas en término de los procedimientos llevados a cabo dentro de la sala de clases, los resultados obtenidos por los estudiantes y la reflexión de ellos y sus profesores al respecto, así como también los principios éticos involucrados en el proceso.

Características: Esta investigación es parte de los requerimientos de tesis para obtener el grado de Doctor en Filosofía(PhD) en Educación en la Universidad de Bristol, Inglaterra. Este proyecto involucra dos partes. El estudio piloto involucrará análisis de documentos, participación en consejos de profesores y entrevistas semi-estructuradas durante noviembre/diciembre 2019. El análisis de documento apunta a identificar la evidencia presente en el colegio que muestre cómo se implementa la política nacional de evaluación y cómo esto se ve reflejado en la práctica en el día a día (en el reglamento interno de evaluación, planificaciones de los profesores, instrumentos de evaluación y libros de clases). Las entrevistas considerarán 2 profesores y 2 estudiantes, dentro del espacio del colegio. El proceso de recolección de datos (primer semestre 2020) involucrará análisis de documentos, según los documentos seleccionados previamente en el piloto, y 20 entrevistas con profesores (10) y estudiantes (10). Las entrevistas tendrán una duración de 1 hora aproximadamente. Los documentos serán obtenidos en formatos previamente acordados con el Director del Colegio (archivos computacionales, imágenes, versiones impresas).

Serás invitado a participar en una entrevista sobre prácticas de evaluación en el colegio, práctica, resultados y ética del proceso, que será audio-grabada. Tu consentimiento implica responder preguntas relacionadas con sus experiencias siendo evaluado en las distintas asignaturas de tu plan de estudios, los métodos usados por tus profesores y las ideas que apoyan esas decisiones, tu opinión sobre tus resultados y del sistema de evaluación en general. El formato de la entrevista es semi estructurado, eso significa que la investigadora tendrá una pauta con preguntas específicas que guían la conversación, pero ésta también seguirá el ritmo de sus opiniones e intereses, respetando el flujo normal de la conversación. Los estudiantes que participen tendrán la oportunidad de elegir el mejor horario para la entrevista, elegir un pseudónimo, revisar las transcripciones de su entrevista y solicitar un resumen de los resultados. Los participantes también tendrán libertad de retirarse del estudio hasta un mes después de que la entrevista haya sido conducida, comunicándolo verbalmente o mediante e-mail a la investigadora. En caso de retirarte del estudio, se entrega el derecho a ser olvidado, que consiste en eliminar toda la información relacionada con la entrevista en la que participaste.

La información será usada para responder las preguntas de investigación de este proyecto y será almacenada protegida por el sistema encriptado de la Universidad de Bristol, que también usa

protección con contraseñas. Los resultados serán usados para publicaciones en revistas académicas, presentaciones en conferencias y también para la tesis final de doctorado, que se presentará en la Universidad de Bristol durante el año 2022. En cualquier formato de presentación, los resultados no incluirán ni el nombre real de los participantes ni el nombre del colegio. La confidencialidad en este punto será estrictamente protegida.

En el caso de que el estudiante no quiera permitir el uso de sus notas para el análisis estadístico para estudio del contexto escolar, por favor complete el formulario al final de la siguiente página y hágalo llegar a la investigadora.

Procedimiento de quejas/información:

En caso de queja o solicitudes de información, por favor contactar:

Estudiante de Doctorado	Tamara Cepeda	Tamara.Cepeda@bristol.ac.uk
Supervisores (Inglaterra)	Sally Thomas Shelley McKeown Jones	S.Thomas@bristol.ac.uk S.mckeownjones@bristol.ac.uk
Director Colegio		

En caso de necesitar traducción para dirigir una queja directamente a la Universidad de Bristol, por favor contactar directamente el Director del colegio para obtener ayuda en el proceso.

Formulario de renuncia a participar

- No autorizo el uso de mis notas para análisis estadístico y contextual.
- No deseo participar en la parte del estudio que involucra el análisis de mis notas.

Nombre del estudiante: _____

Curso: _____

Firma: _____

Fecha: _____

Participant information sheet for underage students (English version)



Name of the project: Exploring Classroom Assessment Practices at a Secondary School in Chile

Student name: Tamara Cepeda

Objective: To identify summative assessment practices conducted in the secondary level in your school as framed by national assessment policy and internal school assessment decrees. The assessment procedures will be analysed in terms of practice within the classroom, results obtained by students and reflected by them and their teachers and the ethical principles involved in the process.

Characteristics: This research is part of the thesis requirements leading to PhD in Education Degree in the University of Bristol, England. This project involves two phases. The pilot study will involve document analysis, teachers' annual assessment meetings participation and semi-structured interviews during November/December 2019. Document analysis will look for evidence of how schools enact the national assessment policy and how that is reflected in the day-to-day practice (school assessment decree, teachers planning, assessment instruments, grades register books). Interviews will be conducted with 2 teachers and 2 students within the school space. Data collection process (first semester 2020) will involve grades and document analysis, as previously selected during the pilot and 20 interviews with teachers (10) and students (10). Interviews will have a duration of 1 hour approximately. The documents will be obtained in formats previously accorded with the Head of School (computer files, pictures, printed versions). Parent/student opt-in consent will be required for grades analysis.

You are asked to participate in an audio-recorded interview about assessment practices within the school, the practice, outcomes and ethics. As you are an underage student, your parent/legal guardian consent will be needed for you to participate in this research. Your consent (and your parent/legal guardian's consent) implies to answer questions related with your experience being assessed in different subjects, the assessment methods used by your teachers and the ideas supporting them, your opinions about your results and the assessment system in general. You will have the opportunity to choose the better schedule, to choose a pseudonymous, to review the interview transcriptions and to request a summary of findings. You will also have the freedom to withdraw until one month after the interview has been conducted by communicating verbally or by e-mail this decision to the researcher. In the case of withdrawal, the right to be forgotten by erasing all data related with your interview will be given.

The information will be used to answer the research questions of this project and will be stored protected by the University of Bristol encrypted system and also with password protection. Results will be used for journal publications, conference presentation and finally, in a written thesis submitted to the University of Bristol during 2022. In any format of presentation, the findings will include neither participants' real names nor the School name. Confidentiality in this regard will be strictly protected.

In the case that the student (or parent/legal guardian) does not want to allow the use of its grades for statistical analysis in a contextual chapter, please complete the opt-out form on the following page and give it to the researcher.

Complaint Procedure/Information request:

In case of complaint or information request, please contact:

Student	Tamara Cepeda	Tamara.Cepeda@bristol.ac.uk
Supervisors (England)	Sally Thomas Shelley McKeown Jones	S.Thomas@bristol.ac.uk S.mckeownjones@bristol.ac.uk
Head of School (Chile)		

In case of need of translation process to make a complaint directly with the University of Bristol, please refer directly to the Head of School in order to obtain help in this regard.

Opt-out consent form

- I do not authorize the anonymised use of my grades for statistical and contextual analysis.
- I do not agree to take part in the study.

Student name: _____

Class: _____

Signature: _____

- I do not authorise the anonymised use of my pupil grades for statistical and contextual analysis.
- I do not agree that my pupil take part in the study.

Parent/legal guardian name: _____

Signature: _____

Date: _____



Hoja de información para estudiantes menores de edad

Nombre del proyecto: Explorando las prácticas de evaluación en el aula en un Educación Media en Chile.

Nombre del estudiante: Tamara Cepeda.

Objetivo: Identificar prácticas de evaluación llevadas a cabo en un colegio chileno en educación media, enmarcadas en la política de evaluación nacional y decretos de evaluación internos. Las prácticas evaluativas serán analizadas en término de los procedimientos llevados a cabo dentro de la sala de clases, los resultados obtenidos por los estudiantes y la reflexión de ellos y sus profesores al respecto, así como también los principios éticos involucrados en el proceso.

Características: Esta investigación es parte de los requerimientos de tesis para obtener el grado de Doctor en Filosofía(PhD) en Educación en la Universidad de Bristol, Inglaterra. Este proyecto involucra dos partes. El estudio piloto involucrará análisis de documentos, participación en consejos de profesores y entrevistas semi-estructuradas durante noviembre/diciembre 2019. El análisis de documento apunta a identificar la evidencia presente en el colegio que muestre cómo se implementa la política nacional de evaluación y cómo esto se ve reflejado en la práctica en el día a día (en el reglamento interno de evaluación, planificaciones de los profesores, instrumentos de evaluación y libros de clases). Las entrevistas considerarán 2 profesores y 2 estudiantes, dentro del espacio del colegio. El proceso de recolección de datos (primer semestre 2020) involucrará análisis de documentos, según los documentos seleccionados previamente en el piloto, y 20 entrevistas con profesores (10) y estudiantes (10). Las entrevistas tendrán una duración de 1 hora aproximadamente. Los documentos serán obtenidos en formatos previamente acordados con el Director del Colegio (archivos computacionales, imágenes, versiones impresas).

Serás invitado a participar en una entrevista sobre prácticas de evaluación en el colegio, práctica, resultados y ética del proceso, que será audio-grabada. Como eres menor de edad, tu participación depende de la autorización de tu apoderado. Tu consentimiento (y el de tu apoderado) implica responder preguntas relacionadas con sus experiencias siendo evaluado en las distintas asignaturas de tu plan de estudios, los métodos usados por tus profesores y las ideas que apoyan esas decisiones, tu opinión sobre tus resultados y del sistema de evaluación en general. El formato de la entrevista es semi-estructurado, eso significa que la investigadora tendrá una pauta con preguntas específicas que guían la conversación, pero ésta también seguirá el ritmo de sus opiniones e intereses, respetando el flujo normal de la conversación. Los estudiantes que participen tendrán la oportunidad de elegir el mejor horario para la entrevista, elegir un pseudónimo, revisar las transcripciones de su entrevista y solicitar un resumen de los resultados. Los participantes también tendrán libertad de retirarse del estudio hasta un mes después de que la entrevista haya sido conducida, comunicándolo verbalmente o mediante e-mail a la investigadora. En caso de retirarte del estudio, se entrega el derecho a ser olvidado, que consiste en eliminar toda la información relacionada con la entrevista en la que participaste.

La información será usada para responder las preguntas de investigación de este proyecto y será almacenada protegida por el sistema encriptado de la Universidad de Bristol, que también usa protección con contraseñas. Los resultados serán usados para publicaciones en revistas académicas, presentaciones en conferencias y también para la tesis final de doctorado, que se presentará en la Universidad de Bristol durante el año 2022. En cualquier formato de presentación, los resultados no incluirán ni el nombre real de los participantes ni el nombre del colegio. La confidencialidad en este punto será estrictamente protegida.

En el caso de que el padre/apoderado del estudiante no quiera permitir el uso de sus notas para el análisis estadístico para estudio del contexto escolar, por favor complete el formulario al final de la siguiente página y hágalo llegar a la investigadora.

Procedimiento de quejas/información:

En caso de queja o solicitudes de información, por favor contactar:

Estudiante de Doctorado	Tamara Cepeda	Tamara.Cepeda@bristol.ac.uk
Supervisores (Inglaterra)	Sally Thomas Shelley McKeown Jones	S.Thomas@bristol.ac.uk S.mckeownjones@bristol.ac.uk
Director Colegio		

En caso de necesitar traducción para dirigir una queja directamente a la Universidad de Bristol, por favor contactar directamente el Director del colegio para obtener ayuda en el proceso.

Formulario de renuncia a participar

- No autorizo el uso de mis notas para análisis estadístico y contextual.
- No deseo participar en la parte del estudio que involucra el análisis de mis notas.

Nombre del estudiante: _____

Curso: _____

Firma: _____

Nombre del padre/apoderado: _____

Firma: _____

Fecha: _____

Parent/legal guardian information sheet (English version)



Name of the project: Exploring Classroom Assessment Practices at a Secondary School in Chile

Student name: Tamara Cepeda

Objective: To identify summative assessment practices conducted in the secondary level in your children/pupil school as framed by national assessment policy and internal school assessment decrees. The assessment procedures will be analysed in terms of practice within the classroom, results obtained by students and reflected by them and their teachers and the ethical principles involved in the process.

Characteristics: This research is part of the thesis requirements leading to PhD in Education Degree in the University of Bristol, England. This project involves two phases. The pilot study will involve document analysis, teachers' annual assessment meetings participation and semi-structured interviews during November/December 2019. Document analysis will look for evidence of how schools enact the national assessment policy and how that is reflected in the day-to-day practice (school assessment decree, teachers planning, assessment instruments, grades register books). Interviews will be conducted with 2 teachers and 2 students within the school space. Data collection process (first semester 2020) will involve grades and document analysis, as previously selected during the pilot and 20 interviews with teachers (10) and students (10). Interviews will have a duration of 1 hour approximately. Your child/pupil will have the opportunity to be accompanied during the interview by you or a classmate. The documents will be obtained in formats previously accorded with the Head of School (computer files, pictures, printed versions). Parent opt-in consent will be required for grades analysis.

Your child/pupil is asked to participate in an audio-recorded interview about assessment practices within the school, the practice, outcomes and ethics. Your consent implies that your child/pupil will answer questions related with his/her experiences being assessed in the different school subjects, the assessment methods used by his/her teachers, opinions about his/her results and the assessment system in general. Participants will have the opportunity to choose the better schedule, to choose a pseudonymous, to review the interview transcriptions and to request a summary of findings. Participants will also have the freedom to withdraw until one month after the interview has been conducted by communicating verbally or by e-mail this decision to the researcher. In the case of withdrawal, the right to be forgotten by erasing all data related with the participant will be given. Parent informed consent will be required for all underage students (under 18 years old). Without your consent, the student cannot be part of the interview process. In case of any doubt, you can e-mail the researcher and be attended personally in the school.

The information will be used to answer the research questions of this project and will be stored in the University of Bristol encrypted system with password protection. Results will be used for journal publications, conference presentation and finally, in a written thesis submitted to the University of Bristol during 2022. In any format of presentation, the findings will include neither participants' real names nor the School name. Confidentiality in this regard will be strictly protected.

In the case that the student (or parent/legal guardian) does not want to allow the use of its grades for statistical analysis in a contextual chapter, please complete the opt-out form on the following page and give it to the researcher.

Complaint Procedure/information request:

In case of complaint or information request, please contact:

Student	Tamara Cepeda	Tamara.Cepeda@bristol.ac.uk
Supervisors (England)	Sally Thomas Shelley McKeown Jones	S.Thomas@bristol.ac.uk S.mckeownjones@bristol.ac.uk
Head of School (Chile)		

In case of need of translation process to make a complaint directly with the University of Bristol, please refer directly to the Head of School in order to obtain help in this regard.

Opt-out consent form

- I do not authorise the anonymised use of my pupil grades for statistical and contextual analysis.
- I do not agree that my pupil take part in the study.

Student name: _____

Class: _____

Parent/legal guardian name: _____

Signature: _____

Date: _____



Hoja de información al padre/apoderado

Nombre del proyecto: Explorando las prácticas de evaluación en el aula en un Educación Media en Chile.

Nombre del estudiante: Tamara Cepeda.

Objetivo: Identificar prácticas de evaluación llevadas a cabo en un colegio chileno en educación media, enmarcadas en la política de evaluación nacional y decretos de evaluación internos. Las prácticas evaluativas serán analizadas en término de los procedimientos llevados a cabo dentro de la sala de clases, los resultados obtenidos por los estudiantes y la reflexión de ellos y sus profesores al respecto, así como también los principios éticos involucrados en el proceso.

Características: Esta investigación es parte de los requerimientos de tesis para obtener el grado de Doctor en Filosofía(PhD) en Educación en la Universidad de Bristol, Inglaterra. Este proyecto involucra dos partes. El estudio piloto involucrará análisis de documentos, participación en consejos de profesores y entrevistas semi-estructuradas durante noviembre/diciembre 2019. El análisis de documento apunta a identificar la evidencia presente en el colegio que muestre cómo se implementa la política nacional de evaluación y cómo esto se ve reflejado en la práctica en el día a día (en el reglamento interno de evaluación, planificaciones de los profesores, instrumentos de evaluación y libros de clases). Las entrevistas considerarán un piloto con 2 profesores y 2 estudiantes, dentro del espacio del colegio. El proceso de recolección de datos (primer semestre 2020) involucrará análisis de documentos, según los documentos seleccionados previamente en el piloto, y 20 entrevistas con profesores (10) y estudiantes (10). Las entrevistas tendrán una duración de 1 hora aproximadamente. Su hijo/pupilo tendrá la oportunidad de ser acompañado durante la entrevista por usted o por un compañero. Los documentos serán obtenidos en formatos previamente acordados con el Director del Colegio (archivos computacionales, imágenes, versiones impresas).

Los participantes serán invitados a participar en una entrevista sobre prácticas de evaluación en el colegio, práctica, resultados y ética del proceso, que será audio-grabada. Su consentimiento implica que su hijo/pupilo responderá preguntas relacionadas con sus experiencias siendo evaluado en las distintas asignaturas de su plan de estudios, los métodos usados por sus profesores y las ideas que apoyan esas decisiones, su opinión sobre sus resultados y del sistema de evaluación en general. El formato de la entrevista es semi-estructurado, eso significa que la investigadora tendrá una pauta con preguntas específicas que guían la conversación, pero ésta también seguirá el ritmo de sus opiniones e intereses, respetando el flujo normal de la conversación. Los participantes tendrán la oportunidad de elegir la mejor planificación de la entrevista, elegir un pseudónimo, revisar las transcripciones de su entrevista y solicitar un resumen de los resultados. Los participantes también tendrán libertad de retirarse del estudio hasta un mes después de que la entrevista haya sido conducida, comunicándolo verbalmente o mediante e-mail a la investigadora. En caso de retirarse del estudio, se entrega el derecho a ser olvidado, que consiste en eliminar toda la información relacionada con el participante. Todos los estudiantes menores de edad deberán tener

consentimiento informado firmado por parte de su apoderado. Sin su consentimiento, el estudiante no puede ser parte del estudio. En caso de cualquier duda, puede contactarse vía e-mail con la investigadora y ser atendido personalmente en el colegio.

La información será usada para responder las preguntas de investigación de este proyecto y será almacenada protegida por el sistema encriptado de la Universidad de Bristol, que también implica protección con contraseñas. Los resultados serán usados para publicaciones en revistas académicas, presentaciones en conferencias y también para la tesis final de doctorado, que se presentará en la Universidad de Bristol durante el año 2022. En cualquier formato de presentación, los resultados no incluirán ni el nombre real de los participantes ni el nombre del colegio. La confidencialidad en este punto será estrictamente protegida.

En el caso de que el padre/apoderado del estudiante no quiera permitir el uso de sus notas para el análisis estadístico para estudio del contexto escolar, por favor complete el formulario al final de la página y hágalo llegar a la investigadora.

Procedimiento de quejas/información:

En caso de queja o solicitudes de información, por favor contactar:

Estudiante de Doctorado	Tamara Cepeda	Tamara.Cepeda@bristol.ac.uk
Supervisores (Inglaterra)	Sally Thomas Shelley McKeown Jones	S.Thomas@bristol.ac.uk S.mckeownjones@bristol.ac.uk
Director Colegio		

En caso de necesitar traducción para dirigir una queja directamente a la Universidad de Bristol, por favor contactar directamente el Director del colegio para obtener ayuda en el proceso.

Formulario de renuncia a participar

- No autorizo el uso de las notas de mi hijo/pupilo para análisis estadístico y contextual.
- No deseo que mi hijo/pupilo participe en la parte del estudio que involucra el análisis de sus notas.

Nombre del estudiante: _____

Curso: _____

Nombre del padre/apoderado: _____

Firma: _____

Fecha: _____

Appendix K: Interview Consent forms

Participant consent form for teachers (English version)



Exploring Classroom Assessment Practices at a Secondary School in Chile

<ul style="list-style-type: none">• I confirm that I have been given, read and understood the information sheet for the above study and have asked and received answers to any questions raised.
<ul style="list-style-type: none">• I understand that my participation is voluntary and that I am free to withdraw until one month after the interview without giving a reason and without my rights being affected in any way.
<ul style="list-style-type: none">• I understand that the researchers will hold all information and data collected securely and in confidence that all efforts will be made to ensure that I cannot be identified as a participant in the study (except as might be required by law) and I give permission for the researcher to hold relevant personal data.
<ul style="list-style-type: none">• I authorize the use of anonymised quotations coming from my interview in the data analysis in journal articles, conferences and the final thesis. I understand that these quotations must be translated to English.
<ul style="list-style-type: none">• I agree to take part in the study.
<ul style="list-style-type: none">• I agree to be audio-recorded.
<ul style="list-style-type: none">• I request a chaperon for the interview procedure YES – NO (circle as appropriate).
<ul style="list-style-type: none">• I understand that I can request a summary of the research findings when completed from the researcher Tamara Cepeda.
<ul style="list-style-type: none">• I understand that my participation involves prohibition of dissemination of the findings in any academic way.

Participant name: _____

Subject: _____

e-mail:

Signature: _____ **Date:** _____

Participant consent form for teachers (Spanish version)



Explorando las prácticas de evaluación en el aula en un Educación Media en Chile.

<ul style="list-style-type: none">● Confirmando que me han entregado la información necesaria respecto a la investigación en la que tomaré parte. He leído y entendido el consentimiento informado de la investigación mencionada en la introducción y he hecho preguntas y recibido respuestas ante cualquier duda.
<ul style="list-style-type: none">● Entiendo que mi participación es voluntaria y que soy libre de retirarme de la investigación hasta un mes después de la entrevista sin dar una razón específica y sin que mis derechos sean afectados de cualquier forma por ello.
<ul style="list-style-type: none">● Entiendo que la investigadora almacenará toda la información y datos recolectados de forma segura y privada y confío en los esfuerzos que serán hechos para asegurar que yo no pueda ser identificado como un participante en el estudio (a menos que sea requerido por la ley) y autorizo a los investigadores a almacenar información personal relevante sobre mí.
<ul style="list-style-type: none">● Autorizo el uso de citas anónimas y fragmentos provenientes de mi entrevista en el análisis de datos conducido por la investigadora en artículos académicos, conferencias y su tesis. Entiendo que las citas deben ser traducidas al inglés.
<ul style="list-style-type: none">● Doy mi consentimiento participar en el estudio.
<ul style="list-style-type: none">● Doy mi consentimiento para ser grabado (audio).
<ul style="list-style-type: none">● Solicito que la entrevista se realice con un acompañante SI – NO (encerrar respuesta).
<ul style="list-style-type: none">● Entiendo que puedo solicitar un sumario de la investigación cuando éstos se encuentren disponibles, a la investigadora Tamara Cepeda.
<ul style="list-style-type: none">● Entiendo que mi participación en este estudio involucra la prohibición de diseminar los resultados en cualquier forma académica.

Nombre del participante: _____

Asignatura: _____

Correo electrónico:

Firma: _____ Fecha: _____

Participant consent form for students over 18 years old (English version)



Exploring Classroom Assessment Practices at a Secondary School in Chile

<ul style="list-style-type: none">• I confirm that I have been given, read and understood the information sheet for the above study and have asked and received answers to any questions raised.
<ul style="list-style-type: none">• I understand that my participation is voluntary and that I am free to withdraw until one month after the interview without giving a reason and without my rights being affected in any way.
<ul style="list-style-type: none">• I understand that the researchers will hold all information and data collected securely and in confidence that all efforts will be made to ensure that I cannot be identified as a participant in the study (except as might be required by law) and I give permission for the researcher to hold relevant personal data.
<ul style="list-style-type: none">• I authorize the use of anonymised quotations coming from my interview in the data analysis in journal articles, conferences and the final thesis. I understand that these quotations must be translated to English.
<ul style="list-style-type: none">• I agree to take part in the study.
<ul style="list-style-type: none">• I agree to be audio-recorded.
<ul style="list-style-type: none">• I request a chaperon for the interview procedure YES – NO (circle as appropriate).
<ul style="list-style-type: none">• I understand that I can request a summary of the research findings when completed from the researcher Tamara Cepeda.
<ul style="list-style-type: none">• I understand that my participation involves prohibition of dissemination of the findings in any academic way.

Participant name: _____

Class: _____

e-mail:

Signature: _____ **Date:** _____

Participant consent form for students over 18 years old (Spanish version)



Consentimiento informado para estudiantes mayores de edad

Explorando las prácticas de evaluación en el aula en un Educación Media en Chile.

<ul style="list-style-type: none">● Confirmando que me han entregado la información necesaria respecto a la investigación en la que tomaré parte. He leído y entendido el consentimiento informado de la investigación mencionada en la introducción y he hecho preguntas y recibido respuestas ante cualquier duda.
<ul style="list-style-type: none">● Entiendo que mi participación es voluntaria y que soy libre de retirarme de la investigación hasta un mes después de la entrevista sin dar una razón específica y sin que mis derechos sean afectados de cualquier forma por ello.
<ul style="list-style-type: none">● Entiendo que la investigadora almacenará toda la información y datos recolectados de forma segura y privada y confío en los esfuerzos que serán hechos para asegurar que yo no pueda ser identificado como un participante en el estudio (a menos que sea requerido por la ley) y autorizo a los investigadores a almacenar información personal relevante sobre mí.
<ul style="list-style-type: none">● Autorizo el uso de citas anónimas y fragmentos provenientes de mi entrevista en el análisis de datos conducido por la investigadora en artículos académicos, conferencias y su tesis. Entiendo que las citas deben ser traducidas al inglés.
<ul style="list-style-type: none">● Doy mi consentimiento participar en el estudio.
<ul style="list-style-type: none">● Doy mi consentimiento para ser grabado (audio).
<ul style="list-style-type: none">● Solicito que la entrevista se realice con un acompañante SI – NO (encerrar respuesta).
<ul style="list-style-type: none">● Entiendo que puedo solicitar un resumen de la investigación cuando éstos se encuentren disponibles, a la investigadora Tamara Cepeda.
<ul style="list-style-type: none">● Entiendo que mi participación en este estudio involucra la prohibición de diseminar los resultados en cualquier forma académica.

Nombre del participante: _____

Curso: _____

Correo electrónico:

Firma: _____ Fecha: _____

Participant consent form for underage students (English version)



Exploring Classroom Assessment Practices at a Secondary School in Chile

<ul style="list-style-type: none">• I confirm that I have been given, read and understood the information sheet for the above study and have asked and received answers to any questions raised.
<ul style="list-style-type: none">• I understand that my participation is voluntary and that I am free to withdraw until one month after the interview without giving a reason and without my rights being affected in any way.
<ul style="list-style-type: none">• I understand that the researchers will hold all information and data collected securely and in confidence that all efforts will be made to ensure that I cannot be identified as a participant in the study (except as might be required by law) and I give permission for the researcher to hold relevant personal data.
<ul style="list-style-type: none">• I authorize the use of anonymised quotations coming from my interview in the data analysis in journal articles, conferences and the final thesis. I understand that these quotations must be translated to English.
<ul style="list-style-type: none">• I agree to take part in the study and I understand that I will need parent/legal guardian consent.
<ul style="list-style-type: none">• I agree to be audio-recorded.
<ul style="list-style-type: none">• I request a chaperon for the interview procedure YES – NO (circle as appropriate).
<ul style="list-style-type: none">• I understand that I can request a summary of the research findings when completed from the researcher Tamara Cepeda.
<ul style="list-style-type: none">• I understand that my participation involves prohibition of dissemination of the findings in any academic way.

Participant name: _____

Grade: _____

e-mail:

Parent/legal guardian name: _____

Student signature: _____ **Date:** _____

Participant consent form for underage students (Spanish version)



Explorando las prácticas de evaluación en el aula en un Educación Media en Chile.

<ul style="list-style-type: none">● Confirmando que me han entregado la información necesaria respecto a la investigación en la que tomaré parte. He leído y entendido el consentimiento informado de la investigación mencionada en la introducción y he hecho preguntas y recibido respuestas ante cualquier duda.
<ul style="list-style-type: none">● Entiendo que mi participación es voluntaria y que soy libre de retirarme de la investigación hasta un mes después de la entrevista sin dar una razón específica y sin que mis derechos sean afectados de cualquier forma por ello.
<ul style="list-style-type: none">● Entiendo que la investigadora almacenará toda la información y datos recolectados de forma segura y privada y confío en los esfuerzos que serán hechos para asegurar que yo no pueda ser identificado como un participante en el estudio (a menos que sea requerido por la ley) y autorizo a los investigadores a almacenar información personal relevante sobre mí.
<ul style="list-style-type: none">● Autorizo el uso de citas anónimas y fragmentos provenientes de mi entrevista en el análisis de datos conducido por la investigadora en artículos académicos, conferencias y su tesis. Entiendo que las citas deben ser traducidas al inglés.
<ul style="list-style-type: none">● Doy mi consentimiento participar en el estudio. Entiendo que necesitaré consentimiento de mi apoderado para participar en el estudio.
<ul style="list-style-type: none">● Doy mi consentimiento para ser grabado (audio).
<ul style="list-style-type: none">● Solicito que la entrevista se realice con un acompañante SI – NO (encerrar respuesta).
<ul style="list-style-type: none">● Entiendo que puedo solicitar un sumario de la investigación cuando éstos se encuentren disponibles, a la investigadora Tamara Cepeda.
<ul style="list-style-type: none">● Entiendo que mi participación en este estudio involucra la prohibición de diseminar los resultados en cualquier forma académica.

Nombre del participante: _____

Curso: _____

Correo electrónico:

Nombre del Apoderado: _____

Firma del estudiante: _____ Fecha: _____

Participant consent form for Parent/legal guardian of underage student (English version)



Exploring Classroom Assessment Practices at a Secondary School in Chile

<ul style="list-style-type: none">• I confirm that I have been given, read and understood the information sheet for the above study and have asked and received answers to any questions raised.
<ul style="list-style-type: none">• I understand that my child/pupil participation is voluntary and that he/she is free to withdraw until one month after the interview without giving a reason and without his/her rights being affected in any way.
<ul style="list-style-type: none">• I understand that the researcher will hold all information and data collected securely and in confidence that all efforts will be made to ensure that my child/pupil cannot be identified as a participant in the study (except as might be required by law) and I give permission for the researcher to hold relevant personal data.
<ul style="list-style-type: none">• I authorize the use of anonymised quotations coming from my child/pupil interview in the data analysis in journal articles, conferences and the final thesis. I understand that these quotations must be translated to English.
<ul style="list-style-type: none">• I agree my child/pupil to take part in the study.
<ul style="list-style-type: none">• I agree my child/pupil to be audio-recorded.
<ul style="list-style-type: none">• I request a chaperon for the interview procedure YES – NO (circle as appropriate).
<ul style="list-style-type: none">• I understand that I can request a summary of the research findings when completed from the researcher Tamara Cepeda.
<ul style="list-style-type: none">• I understand that my child/pupil participation involves prohibition of dissemination of the findings in any academic way.

Student name: _____

Grade: _____

Parent/legal guardian name: _____

Parent/legal guardian e-mail:

Parent/legal guardian signature: _____ **Date:** _____

Participant consent form for Parent/legal guardian of underage student (Spanish version)



Explorando las prácticas de evaluación en el aula en un Educación Media en Chile.

<ul style="list-style-type: none">● Confirmando que me han entregado la información necesaria respecto a la investigación en la que mi hijo/a-pupilo/a tomará parte. He leído y entendido el consentimiento informado de la investigación mencionada en este consentimiento y he hecho preguntas y recibido respuestas ante cualquier duda.
<ul style="list-style-type: none">● Entiendo que la participación de mi hijo/a-pupilo/a es voluntaria y que es libre de retirarse de la investigación hasta un mes después de la entrevista sin dar una razón específica y sin que sus derechos sean afectados de cualquier forma por ello.
<ul style="list-style-type: none">● Entiendo que la investigadora almacenará toda la información y datos recolectados de forma segura y privada y confío en los esfuerzos que serán hechos para asegurar que yo mi hijo/a-pupilo/a pueda ser identificado/a como un/a participante en el estudio (a menos que sea requerido por la ley) y autorizo a la investigadora a almacenar información personal relevante sobre mi hijo/a-pupilo/a.
<ul style="list-style-type: none">● Autorizo el uso de citas anónimas y fragmentos provenientes de la entrevista de mi hijo/a-pupilo/a en el análisis de datos conducido por la investigadora en artículos académicos, conferencias y su tesis. Entiendo que las citas deben ser traducidas al inglés.
<ul style="list-style-type: none">● Doy mi consentimiento para que mi hijo/a-pupilo/a participe en el estudio.
<ul style="list-style-type: none">● Doy mi consentimiento para que mi hijo/a-pupilo/a ser grabado (audio).
<ul style="list-style-type: none">● Solicito que la entrevista se realice con un acompañante SI – NO (encerrar respuesta).
<ul style="list-style-type: none">● Entiendo que puedo solicitar un sumario de la investigación cuando éste se encuentre disponible, a la investigadora Tamara Cepeda.
<ul style="list-style-type: none">● Entiendo que la participación de mi hijo/a-pupilo/a en este estudio involucra la prohibición de diseminar los resultados en cualquier forma académica.

Nombre del alumno: _____

Curso: _____

Nombre del Apoderado: _____

Correo electrónico del Apoderado:

Firma del apoderado: _____ Fecha: _____

Appendix L: Ethics form approval

[University home](#) > [RED](#) > [Ethics On-line Tool](#) > Application Review

Application Review

Logout

ID	Submitted on behalf of	Name	Faculty	Department	Supervisor
95963	Tamara Cepeda	Mrs Tamara Cepeda	Faculty of Social Sciences and Law	Graduate School of Education	Professor Sally Thomas

Status

Signed off

Date added

Nov. 4, 2019

Signed off date

Nov. 20, 2019

Is this a student project?

Postgraduate Phd

Project title

Assessment of policy implementation in a Secondary School in Chile: practice, results and ethics.

Estimated start date

Nov. 11, 2019

Duration (months)

2

Project outline

Globally, research on in-classroom summative assessment is regarded as limited and based upon sparse available information, as the focus is placed usually on formative assessment (Baird et al., 2014; Young & Kim, 2010; Lambert & Lines, 2000; Stiggins & Conklin, 1992). This research topic appears to be receiving fewer resources when compared with standardised tests with accountability purposes (Harlen, 2004). Therefore, we know very much about the numbers generated by external standardised assessment procedures conducted within schools, but the same does not apply to what is happening with the classroom summative assessment procedures occurring between students and teachers. This reinforces the need for studies focused on the shifting dynamics that are constantly taking place in the day-to-day school life, namely, how assessment practices and specifically internal summative assessments are integrated into educational systems (SWAAP Report, 2017). For that reason, this research aims to analyse in-depth three stages: (1) an overview of the summative assessment practices and methods used in one typical Chilean secondary school as guided by national and school-created decrees, (2) students' and teachers' perspectives of the summative assessment process in the classroom and their attendant positive/negative consequences and finally the (3) ethics involved in the procedures.

Files

Full Ethics Application [ethics/95963/ethics-tc-final.docx.pdf](#) (690.2 KB added on Nov. 4, 2019)

Consent Form [ethics/95963/informed-consent.pdf](#) (470.2 KB added on Nov. 4, 2019)

Participant Information Sheet [ethics/95963/information-sheets.pdf](#) (577.5 KB added on Nov. 4, 2019)

Appendix M: Themes, subthemes and topics from interviews' transcriptions (RQ1, 2 and 3)

<i>Theme</i>	<i>Sub-themes</i>	<i>Associated topics</i>	<i>References on interviews</i>
<i>Curriculum contents (RQ1a)</i>	- Impact of curriculum on assessment purposes.		46
<i>Summative culture</i>	- Summative use of formative (RQ1b).	- Graded formative work and rubrics. - Scarce graded co and self-assessment.	161
	- Low variety of assessment/test culture (RQ1b).	- Test as main assessment method. - Use of standardised tests formats (SIMCE, PAES). - Contradictory views on the use of other assessment methods. - The role of memory and reflections on learning.	110
	- Lack of feedback (RQ2).	- Scarce feedback provision impacted by teachers' lack of time, number of students per class and extension of curriculum contents.	65
<i>Complexities of grading (RQ2)</i>	- High frequency of grading. - Score Pollution - Achievement and grades differences. - Strategies for low performance. - Points assignation for grading. - Reflection on learning.		143
	- High stakes of the system (repeating, promotion, university entrance).		38
<i>Justice (RQ3)</i>	- Need for changes towards a fairer system.	- Social inequality. - Injustice of grading. - Lack of teachers' time. - Teachers' work overload.	118

Appendix N: Excerpt of teacher's interview transcription (original version in Spanish).

R: ¿Les gusta? ¿Les gusta cuando aplicas evaluaciones que implican creatividad?

T05: Enganchan, te permite que desarrollen habilidades que no necesariamente son de la asignatura, pero que ayudan a los cabros que de repente no les gusta historia y de repente tienen que hacer otro tipo de cosas para poder equilibrar, pero la prueba escrita sí tiene como preponderancia, porque eh... el análisis de fuentes ahí es más potente, entonces no son cosas, no son pruebas de memoria, sino que también son de análisis...

R: Claro, y de aplicar cosas....

T05: Y de aplicar.

R: ¿Qué estructura tienen generalmente las pruebas?

T05: De alternativa, principalmente...

R: ¿Pero también el análisis de texto lo hacen con alternativa?

T05: Sí, o sea de repente hay análisis de texto con alternativa y hay otros que se hace con desarrollo.

R: Ya, preguntas abiertas...

T05: Sí, preguntas abiertas se usa bastante, sólo en la prueba final que ahí no podemos hacer desarrollo porque, por un tema de formato...

R ¿el formato viene dado?

T05: Claro, ahí son puras alternativas, pero suele ser mixto, mezclado una y otra.

R: oye, ¿y cómo asignas puntaje? Por ejemplo en las pruebas, usas como que todas valen lo mismo, o las que analizan texto valen más...

T05: No, las pruebas suelen ser todas con la misma ponderación, excepto las de desarrollo, las de desarrollo suele tener, tratamos de equilibrar que, que sea ojalá más de la mitad de puntaje que de las alternativas, por ejemplo, si una prueba tiene 14 preguntas alternativas, la de desarrollo tiene que tener 7 puntos, cachai, pa que equilibre, entonces si a un cabro no le fue tan bien en las alternativas pero tiene habilidades de escritura, va a poder equilibrar...

R: Influye ahí, por ejemplo en pensar que si copió todas las alternativas, en la parte del desarrollo vai a cachar si de verdad estudió o no...

T05: Tendría que... no lo pensamos así necesariamente, porque, la verdad que con la [REDACTED] somos bien agujas pa' tomar las pruebas, yo soy bien así como meticulosa, quito celulares, los separo...

R: Ya, tratan de monitorear lo que más puedan...

T05: Sí, cambio de puesto, hay un monitoreo constante entonces yo no hago mis instrumentos evitando que se copien, nunca me ha pasado, entonces la verdad es que no tengo como esa visión,

siento que es más para comprobar si efectivamente, una, hay habilidades de escritura, si hay una comprensión del texto como mayor...

R: Claro... oye [REDACTED] ¿cuáles son los niveles cognitivos que evalúas en las pruebas, emm... esto de analizar, por ejemplo, los textos no es sólo recordar o también es aplicar...?

T05: La verdad es que depende, porque por ejemplo en los cursos inferiores, en séptimo, si hay que evaluar localización de información en el texto, como algo más básico y después pasai a una pregunta más compleja, porque hay un tema de pasos del, del análisis del texto, más un texto histórico, donde los chiquillos se tienen que fijar en el año, quién lo escribió, dónde lo escribió, es que ahí hay como un...

R: O sea hay varias habilidades involucradas en una actividad...

T05: Exacto...

R: no sólo identificar, sino que también...

T05: no, entonces pero las preguntas, al menos las que son de alternativa, tienen que ver con extraer idea principal, emmm... o análisis a partir del texto, analizar el contexto histórico, no tienen que ver con, con eh... buscar información así tan, tan específica, tiene que ver más con la disciplina, el texto para la disciplina.

R: Oye [REDACTED], si bien la frecuencia y la cantidad de notas que tú pones están dadas como de arriba, por UTP, ¿harías algún comentario al respecto, de este ritmo de la nota mensual, o de que de repente, eh... estas fechas estandarizadas atentan con las velocidades de los cursos...?

T05: Sí...

R: a veces un curso va más lento y te piden, te piden la nota, revisan los libros, ¿cómo te sientes con esto?

T05: Mira la verdad que yo, que uno va tomándole ritmo a los cursos y claro, esa medida no está necesariamente acorde a cada curso po, no sé, tenía un séptimo, que era como el séptimo más complejo, que tenía muchas notas porque todo hay que hacerlo evaluado para que los chiquillos lo hagan...

R: sino, no trabajan...

T05: sino, no trabajan, entonces ahí nunca tuve problemas de notas, porque los chiquillos siempre trabajaban, pero con los octavos me pasó por ejemplo que, eh... teníamos el SIMCE encima, y hubo un minuto que llegó Octubre y yo tenía dos notas, porque me había dedicado a pasar contenidos, a pasar contenidos, a pasar contenidos, y no, no había podido detenerme para hacer una evaluación, había hecho dos pruebas escritas, pero no tenía más evaluaciones, eh, y eso también, em... te va dando los ritmos, claro esto de las pruebas, por ejemplo yo en octavo tenía que pasar la materia volando para poder, que los chiquillos cubrieran el SIMCE, y que no necesariamente te da la profundidad que uno quiere, o sea alcanzaste a pasar todo, sí, pero con un nivel, no el adecuado para llegar a primero medio al menos, entonces siento yo que, que hay metodologías que de repente necesitan un poco más de flexibilidad, no sé po, por ejemplo, nos plantearon trabajar en

proyectos, pero el aprendizaje basado en proyectos la misma persona que nos va a hacer la charla decía que necesitai 15 clases, de adónde vas a sacar 15 para hacer un trabajo...

R: Y más si tienes los mismos requerimientos de notas y no cambia nada como en la estructura arriba...

T05: Exacto, claro entonces son ideas súper concretas, que son buenas, pero que tienen que tener un cambio también en la parte administrativa, porque o sino, vas a chocar siempre con la misma vara.

R: claro, porque también me comentaban que pasa eso como con el nuevo decreto de evaluación, que no va a ser necesaria como la nota mensual, pero igual se le da como un foco a lo formativo, pero a la vez siguen pidiéndoles que pongan notas, que haya promedios, que haya no sé, las pruebas semestrales supongo que van a seguir existiendo, entonces igual esto...

T05: yo creo que tiene que ser un cambio, va a ser un cambio paulatino, cuando empezamos a ver que si hay resultados, pero que esos resultados no se condicen con lo que nos están pidiendo, eh...yo siento igual que lo de la nota mensual te da posibilidades de...es democrático pa los chiquillos, porque también sería súper injusto aparecer con 'ya voy a evaluar proceso' pero en ningún momento monitoreaste eso y resulta que al final le aparecen a los chiquillos cuatro notas que ellos no tenían idea, que eso puede ser perjudicial también pa un cabro que a lo mejor no le va tan bien, entonces igual siento yo que la nota mensual ayuda a ir monitoreando y ayuda a que el proceso de evaluación sea democrático, o sea transparente, o sea que no pase que, porque hay profesores que pasan, no sé po, tres meses y los cabros no ven una nota, y después aparecen todas las notas y notas que no sabí bien mucho porqué, no hay retroalimentación, entonces eso igual de repente puede tender a, a afectar más el proceso de evaluación que a mejorarlo.

R: Oye [REDACTED], ¿usaste coevaluación o autoevaluación en el año? ¿tuviste espacio, tiempo como para probar alguna de esas estrategias?

T05: Ehh...espérame, sí, en séptimo hice co-evaluación, hice auto-evaluación en séptimo, con una maqueta que los chiquillos hicieron en la casa, un proyecto familiar, ahí lo hicimos.... Coevaluación hice en octavo, donde hago estos trabajos donde yo te decía, como más flexibles, ahí se da también para hacer....

R: ¿y cómo reaccionaron los chiquillos con las experiencias?

T05: La verdad es que me pasan dos cosas, una que depende mucho de las personalidades de los chiquillos, o sea si hay un cabro que tiene muy buena autoestima, y poca capacidad de reflexión, en ningún minuto se va a poner realmente lo que hizo, cachai, grupos que, que claro se junta mucha gente que tiene egos muy altos y cuesta mucho reconocer errores en los chiquillos, y eso se choca de repente con que ellos creen que lo hicieron muy bien, pero de repente ver la pauta y la nota no refleja el mismo... lo que evaluó uno versus lo que evaluaron ellos. Y me pasa en otros casos que, hay cabros que son muy autodestructivos con sí mismos, que toman la coevaluación como para, hay un, un asumir de que quizás no hubo algunas cosas, pero suele ser que es como muy autodestructivo, es como 'no, hice todo mal, no seguí esto' y claro, versus tú nota, es bien distinto, de repente uno ve, si pero aquí no sé, por ejemplo las maquetas, estaba contextualizado históricamente, me expuso bien lo que, o sea hice tal cosa y me supo explicar cuando había sido

creado, había una investigación detrás y según él habían hecho todo mal, entonces se da mucho, yo creo que falta una apropiación de esa, ese ejercicio evaluativo, que no dependa de que si tú te quieres mucho o no, si tú crees que es muy bacán o creí que...

R: ser capaz de mirar como con más objetividad el trabajo...

T05: Exacto, con objetividad, que es lo mismo que hace uno, o que debiera hacer uno supuestamente, con objetividad, fuera de lo que yo siento de mi trabajo, de mis estudios, de mi persona, porque además que, no sé al menos yo trabajo en edades que son súper complejas, séptimo y octavo los chiquillos, o creen que son adultos y maravillosos o creen que son pencas, autoestimas así muy, muy variables, entonces yo creo que hay que trabajar en eso, hay que, que no dependa de la persona, sino que el instrumento también les permita, ser lo más objetivos posibles con su propio trabajo.

R: ■■■■, ¿tú le comunicas los objetivos de aprendizaje antes a los alumnos y las expectativas que tienes de ellos? Así como esto vamos a aprender, este es el contenido, esto es lo que espero yo de ustedes.

T05: Sí, suelo hacerlo, bueno todas las clases plantear el objetivo, lo que sí me pasa de repente es que no soy..no, no, no hay una coincidencia entre el objetivo que declaro en la planificación y el que te voy a poner en la sala, porque son objetivos demasiado teóricos, que son pal profe digamos, entonces lo, lo que estoy haciendo principalmente es trabajar con los indicadores, que son más aterrizados, son más cortitos, entonces muchas veces te sirve más eso pa' trabajar que el objetivo que es que...como bien amplio. Además que en historia es como habilidad, contenido, metodología, ya una cosa enorme, entonces el indicador es más fácil, sí, y, al empezar unidad siempre digo 'chiquillos, miren, este contenido es, es un poco más complejo, este contenido es más llevadero, vamos a vincular acá, vamos a hacer tal cosa, siempre tratando de hacer como un contexto, cosa de que los chiquillos sepan para donde vamos, qué espero yo de ellos, también ser súper transparente, yo como que no me complica en decir 'saben que chiquillos, a mí esta parte no me es muy agradable (risas) este contenido no me gusta mucho, eh...entonces lo vamos a tratar de hacer más entretenido, porque siento que eso también les da, que lo que a ellos también les puede no gustar un contenido, más que otro, eh... y está bien en el fondo, no es que la profe le tenga que gustar todo tampoco, ni a ellos tampoco, entonces siento yo que eso... a mí no me complica la verdad asumir que hay contenidos que, que no me son tan agradables como otros y hay que transparentarlos porque eso también le da una, a las expectativas que tú tení, de repente hay contenidos que uno de verdad... Grecia y Roma a mí me carga, me carga, me cuesta mucho, lo hago con mucho esfuerzo, versus, no sé, contenidos de primero medio, octavo, que son mucho más... cuando uno ve revolución francesa, son contenidos que, que igual son más agradables pa' uno, porque uno es ser humano, o sea tiene preferencias, al igual que ellos, y a mí no me complica esa parte.

R: Oye ■■■■, me hablaste algo de la retroalimentación, que reciben tus alumnos, eh, ¿qué características tiene la retroalimentación, más allá de la nota que es una forma de retroalimentación?

T05: Yo la verdad que el monitoreo, siento que a mí me funciona bastante bien, el hecho de que los chiquillos tengan, yo lo que hago, principalmente, es plantearles las instrucciones primero, yo les entrego una guía con todas las instrucciones, eso es lo primero, la primera pantalla que yo les doy,

de hecho la trabajamos en clase, como punto por punto, ya, de hecho suele ser un poco latera esa parte, pero siento yo que es, es más transparente hacerlo así, más con cursos que son más chicos que cuesta un poco más. Entonces ya chiquillos, vamos a evaluar esto, esto, yo quiero que hagan esto, se van a reunir... todo clarito, y esa pauta yo después la transformo en una rúbrica, pero lo hago en el proceso, porque eso me permite hacerle ciertas adecuaciones a la rúbrica...

R: ¿vas ajustando de acuerdo a cómo los ves que se van desempeñando?

T05: Sí, porque hay cosas que de repente eh...yo le quería dar este énfasis y resulta que en el camino le di énfasis a otra cosa, y eso también... pero la instrucción ya está, si yo siempre, si, si me guío por esa instrucción, la rúbrica digamos, va a tener coherencia y no le voy a mentir a los chiquillos algo que no estaba en las instrucciones, no, siempre las instrucciones, siempre séptimo y octavo lo hago así, como muy, muy detallado eh... muy a prueba de errores, cosa de que sea, sea lo más transparente posible el proceso, pero sí yo las rúbrica las voy adaptando en el camino.

R: en el camino...

T05: en el momento ya de calificar, de colocar la nota, la verdad es que suele ser súper coherente, nunca he tenido un reclamo de que 'ah, profe usted me está evaluando algo' y no porque el monitoreo es antes, y ellos también saben, 'profe, tiene razón, usted nos dijo' entonces lo que yo hago los monitoreo con las retros, es agarrar la rúbrica y sentarme con los chiquillos, como ya, chiquillos, en cambio con las pruebas escritas, debo reconocer que a mí me cuesta un poco más hacer retroalimentación, siento que es instrumento más servicial, más útil digamos cuando uno está apurado, entonces la verdad es que ahí me caigo...

R: Te das más el trabajo en detalle cuando trabajan con rúbricas...

T05: Sí, la prueba escrita me cuesta bastante, y debo asumirlo con toda humildad...

R: O sea generalmente está más marcadas las malas, las buenas, la nota y ya...

T05: Sí, y en los desarrollos hay comentario obviamente pero no es algo que yo pueda decir, además que porque es individual, en cambio las rúbricas suelen ser trabajos grupales, cuesta menos, por temas de tiempo, cuesta menos retroalimentar que cuando es una prueba escrita, individual, donde tenía que retroalimentar 45...

R: y va con comentarios en el fondo que les sirven a varios, cuando están trabajando en grupo...

T05: Claro, exacto, ahí yo creo que falta, faltan mecanismos quizás, eh... falta darse el tiempo no más de hacerlo, de poder hacerlo.

R: ■■■■, respecto al rendimiento, y las notas, que es una de las cosas que también, también me interesa, es tu opinión, ¿cuáles dirías tú que son las causas de diferencias entre alumnos, entendiendo que todos reciben la misma educación pero resulta en esta distribución de que a algunos les va muy bien, a otros les va mal, a qué crees tú que se debe eso?

T05: yo siento que con el tiempo me he ido convenciendo a mí misma de que, primero en términos súper generales, no sólo de mi asignatura, hay cabros que están más propensos al sistema escolar que otros, o sea, hay niños que por temas familiares, por temas de personalidad calzan mejor en un sistema escolar y eso permite que a los niños les vaya mejor, versus, versus otros alumnos que

lamentablemente no tienen capacidades, tienen personalidades muy fuertes de repente que chocan eh, con el... con este sistema que es mucho más cuadrado, más autoritario y eso genera bajas calificaciones, ahora yo siento que el tema del capital cultural al menos en mi asignatura es fundamental, o sea, se nota demasiado cuando hay niños que traen un capital cultural mayor, que les interesan temas que al resto no le interesan y eso genera una brecha gigantesca, o sea a mí me ha tocado trabajar en otros colegios en donde si tú como profe si no generas el capital cultural el cabro no va a tener capital cultural por muy inteligente y adaptable al sistema escolar que sea, no lo va a tener, porque no lo trae de su casa, su contexto no le permite generar capital cultural, y eso es...emmm...ya te genera una distancia importante.

R: es una responsabilidad heavy también...

T05: Pal' profe, pal profe porque más encima, yo no tengo la posibilidad, el año pasado no tuve ninguna salida pedagógica, o sea ni siquiera las que ya la ■■■ tenía impuestas, ni siquiera esas pudimos realizar por temas de recursos, entonces emm...también me imposibilita no sé, llevar a los chiquillos a un museo...

R: claro que a lo mejor nunca han entrado a uno...

T05: que es lo más probable, y no porque no tengan los recursos, sino porque la familia no trae esa, no trae esa, ese arraigo o dinámica...

R: No se les ocurre, no es parte de los panoramas de familia...

T05: Claro, nos pasaba con la ■■■ del día del patrimonio, 'a profe, qué es eso, y qué podemos ir a ver' y hay cabros que no están ni ahí, aunque tú les ofrezcas una nota, nada, entonces por ahí pasa y otra que yo creo que hay un tema de que eh...lamentablemente el, el modelo de, de calificaciones que tenemos nosotros no siempre está adaptado a los conocimientos que los chiquillos adquieren, eh, la prueba escrita si bien es un instrumento que es favorable pal profe, muchas veces no te refleja lo que los cabros aprendieron, y ha pasado muchas... a mí me ha pasado muchas veces, cabros que tienen pésima nota en mi asignatura, pero que tú les preguntai y a los cabros les gusta la asignatura y saben, pero al... se genera ahí un, una disparidad yo creo por el tipo de instrumento que utilizamos, que un instrumento muy rígido....

R: ¿Tú crees que hay rechazo de ellos también a esta evaluación tan rígida o...?

T05: No sé, porque de repente siento que hay más rechazo a las evaluaciones flexibles que a las pruebas, porque están muy acostumbrados a las pruebas, muy acostumbrados, por ejemplo me pasó en los sextos el año pasado que me decían los chiquillos 'ay tía, a mí me iba excelente en historia' y yo decía ¿pero por qué? 'no sé po tía, porque yo daba las pruebas y me iba bien', cachai, entonces eh... pasa mucho que hay una resistencia a los trabajos, a que me voy a tener que esforzar más, a que voy a tener que crear algo, así como qué lata, que fome, en verdad una prueba es más fácil...

Appendix O: Excerpt of student's interview transcription (original version in Spanish).

S03: Yo creo que en general en todas las materias el objetivo de las evaluaciones era saber si habíamos aprendido algo, siento que era así, aunque en algunos casos no, por ejemplo, en matemática era como pasar materia y hacer una prueba para que nos fuera bien a todos y no era como preocuparse de que hubiéramos aprendido en verdad la materia, sino que era para tener buenas notas, entonces pasaba un contenido y la prueba era al tiro, pa que el contenido estuviera fresco, lo hacían así. En historia, era distinto porque era como mucho contenido y después venía la evaluación, y la profe da un cuestionario para guiarnos y todo, pero en general siento que eso intentaban hacer con las evaluaciones.

R: O sea ver si habían aprendido...

S03: Sí.

R: Tú crees que, o sea, sientes que estas evaluaciones que te hacían en el colegio sí te ayudaban a saber si habías aprendido o si no...

S03: Sí.

R: Ya, y ¿piensas que los profes usaban como estrategias remediales cuando había bajos resultados?

S03: Sí, caleta, siempre, siempre que nos iba mal, usaba alguna estrategia, bajar la escala, hacer otra prueba, hacer alguna actividad.

R: ¿Y sientes que eso es justo? Como que el número que yo veo, si yo voy a mirar las notas de tu curso, en general, y hago un análisis de esas notas, ¿tú sientes que es fiel reflejo de lo que aprendieron?

S03: Emm...no. Para nada porque, bueno, la verdad no sé, porque los mismos que se sacaban malas notas y los hacían hacer una prueba o una actividad, les iba mal de nuevo, pero igual yo encuentro que es injusto para la gente que sí se esforzó en verdad la primera vez, claro.

R: ¿y todos tenían que hacer la misma evaluación?

S03: Si po.

R: No era solo a los que les había ido mal...

S03: O sea era como mínimo cuando te decían que si querías, podías hacerla, pero era obligación, o era todos o nadie.

R: Era para todo el curso, ¿y crees, en las clases, en los distintos ramos que tienes, tienes tiempo de reflexionar sobre el aprendizaje? Es como conversar con el profe, lo que les cuesta, las estrategias...

S03: Ya... eh.... no, no porque la verdad la clase es como muy justa, entonces pasan los contenidos, después una actividad y listo...

R: ¿por una cosa de tiempo?

S03: Sí.

R: ¿Y crees que eso es importante?

S03: Yo creo que sí porque al final uno tiene que también tener una relación con el profe porque así uno se va entendiendo, o sea no de amistad, hablo de entendimiento entre profe y alumno para estar seguro que el alumno está entendiendo lo que uno está enseñando.

R: Y tú cuando no entendías, a nivel personal, ¿buscabas ayuda?

S03: No, yo nunca preguntaba nada, era como lo mínimo, y si no entendía algo le preguntaba a la [REDACTED], o a mis amigas.

R: Entonces pedías ayuda más entre tu grupo, que al profe...

S03: Sí

R: ¿Cómo evalúas en general las estrategias de evaluación que usan los profes en el colegio?

S03: (risas) si es que hay algunas y hay otras que no son tan buenas. Pero encuentro que la verdad eran muy fáciles para lo que yo había vivido no sé, en mi otro colegio, y eso...

R: ¿en qué colegio estabas antes?

S03: En uno de población, pero era súper bueno, se llamaba Santiago de Chile, y cuando yo me cambié...

R: ¿En qué comuna era?

S03: En La Florida.

R: ¿y sientes que era más exigente la evaluación que acá?

S03: Sí, incluso me costaba más, me iba peor que cuando llegué aquí y encuentro que eso es igual un tema porque este colegio igual se paga y el otro era municipal, era gratis, y era mucho más difícil, a mí me costaba mucho las pruebas, mis promedios no subían del 5.

R: y acá, sientes que haciendo el mismo esfuerzo, o teniendo el mismo nivel, ¿te sacas mucho mejor nota?

S03: claro, pero igual cuando me empecé a esforzar para el NEM en la media, ya ahí estudiaba, porque sabía que tenía que tener buen NEM, entonces me iba mejor.

R: ¿En qué curso llegaste al colegio?

S03: En séptimo...

R: Ya mira te voy a hacer unas preguntas como de las metodologías que usan los profes y la primera es ¿qué tipo de evaluación tenías?

S03: Ya, las pruebas de alternativas, pero eso era siempre en matemática, historia, en filosofía es puro escrito, puro desarrollo, escribir en 250 palabras, tal tema.... Era como profe, qué pasa si me paso, si no me paso....

R: Entonces, la mayor cantidad son pruebas...

S03: Sí...

R: ¿Y informes escritos, presentaciones, oral, disertación?

S03: Eh... no la verdad sólo en electivo este año....

R: ¿y en ciencias? Experimentos o algo...

S03: No, nada.... O sea, en los electivos no sé, el único experimento que yo tuve fue el del encéfalo, el del cerebro, pero éramos como un grupo de 10 personas y yo no hice nada, yo solo vi como lo hacían.

R: ¿Y cómo los evaluaron ahí?

S03: Emm...con un informe escrito.

R: Entonces, me lo dijiste recién, pero para confirmar, ¿qué tipo de preguntas o ítem tienes usualmente en una evaluación?

S03: ya las alternativas, claro, si ésta o esta no, solo I solo II, son horribles...

R: ¿Y verdadero y falso o términos pareados?

S03: Casi nunca....

R: Y ¿por qué crees que se usa más la alternativa?

S03: Yo creo que es más fácil para el alumno, o sea cuando un alumno sabe que la prueba es sólo alternativa se confía, entonces ya si no me sé una, la tiro al azar y le achunto capaz, entonces igual no se esfuerzan tanto.

R: Oye y ¿tú que crees que los profes tratan de evaluar, de verdad si aprendiste, sientes que las preguntas que te hacen, así como las formulan, como es una prueba normal común y corriente, sientes que en el fondo evalúa de verdad si aprendiste?

S03: Eh, no. En matemáticas me acuerdo que la profe nos hacía pruebas y ponía ejercicios que hacía en la misma clase, y yo voy a Preu, entonces cuando me pasaban guías con ejercicios de lo mismo que habíamos pasado en el colegio, yo no sabía cómo se hacían, porque ese ejercicio o esa metodología nunca me lo habían enseñado, entonces me hacen un ejercicio fijo y confíate, porque ese ejercicio va a salir en todos lados, va a tener que ser, te lo van a hacer siempre... entonces era como, pasaba el contenido y la prueba venía al tiro, entonces era como, aprendiste en ese momento, bacán, sino....

R: ¿y sientes que después de la prueba se te olvidan las cosas?

S03: Sí, me pasaba mucho, pero en otras materias, como en historia, yo creo que es porque era más contenido e igual es difícil como contextualizarlo todo...

R: Entonces dirías que igual es como más memoria...

S03: Claro...

R: ¿Y crees que debería ser así?

S03: No porque al final no sirve de nada porque después se te olvida, y quedai en blanco igual (risas). Igual a mi después me iba mejor porque como voy al preu y todo y hay más tiempo, las clases duran más, son como 2 horas y media de clases, en el preu entonces es distinto.

R: Oye y si pudieras tú elegir, qué te gustaría que evaluaran en el colegio...

S03: Yo creo que, uhh, qué difícil la pregunta...

R: porque recién me decías 'siento que evalúan más la memoria', entonces qué tipo de....

S03: no, me gusta que hagan informes escritos porque uno puede dar su opinión y puede expresarse así como uno es en verdad...

R: ¿por qué crees que te hacen poquito eso?

S03: No sé, la verdad no tengo idea, yo creo que porque a los niños les cuesta más igual, como escribir, bueno no a todos, pero a la mayoría como que no puede desarrollar una idea escrita, entonces prefiere las alternativas que es como lo más accesible para todos.

R: Oye y respecto al feedback, la retroalimentación que tú recibes, ¿solías recibir como comentarios escritos en tus pruebas o trabajos que te ayudaran?

S03: ya, en las pruebas nada, así en las de alternativa como ningún comentario escrito, sino que sólo la que era buena, la que era correcta, te la marcaban, y en los escritos, de ortografía, de ideas como qué arreglar, qué poner, o lo que estaba mal, eso era más en los informes escritos, que era más filosofía, porque era el único donde nos hacían escribir...

R: ¿y te servía para aprender y para identificar tus errores?

S03: Sí, yo creo que estoy agradecida con el profe Alonso porque gracias a él hice una buena tesis...

R: Oye y tú calificarías esa retroalimentación que te dan como positiva o negativa, en el sentido de que está enfocada en los errores o está enfocada en tus fortalezas, en lo bueno o en lo malo...

S03: Yo creo que había de las dos en ese caso, como en los errores y en las fortalezas, el profe como me decía, ya mira esta idea está buena, pero no me estás dando un argumento claro, entonces como apóyala más... era más ideas positivas que negativas eso sí.

R: Y esta retroalimentación o sea, por ejemplo en la sala de clases, había retroalimentación oral, como que te dijeran eh, bien [REDACTED], no sé, si preguntas algo, que hubiera como un refuerzo...

S03: No, nada, era como entregar las pruebas y listo...

R: Y eso ¿se hace en privado?

S03: mmmmm...ni tanto en privado...

R: ¿Por qué igual todos se enteran la nota que te sacaste?

S03: Sí.

R: ¿Y sientes que esa retroalimentación está basada en tu logro?

S03: sí...

R: ¿sientes que es distinto para alguien que tiene buenas notas y para alguien que tiene malas notas?

S03: demasiado, es súper distinto, o sea yo encuentro que a los que les va peor como que les ponen más comentarios incluso, es como 'esfuérzate, tú puedes' pero a los que nos va mejor, no es por creerme el cuento, pero no nos dicen casi nada... es como que asumen que nos va a ir bien...

R: ¿y tú crees que es importante también ese feedback?

S03: yo encuentro que todos deberíamos recibir eso porque al final si nos va bien, es porque nos esforzamos, no porque seamos inteligentes así ¡oh!...

R: Oye sientes que este feedback, sean comentarios, sea la nota, ¿la recibías como luego o muy retrasada?

S03: no, era dependiendo del profe, había profes que se tomaban el tiempo y nos las entregaban como en 3 semanas....

R: ya, ¿y tú conoces el reglamento del colegio?

S03: Sí, creo que tienen como un periodo para tomar las notas y entregar...

R: ¿y tú sientes que se cumplía?

S03: Con algunos profes sí, con otros profes no.

R: Y ahí reclamaban, pasaba algo, o simplemente esperaban a que les entregaran...

S03: o sea había alumnos que reclamaban, así como 'oiga profe, yapo, queremos saber la nota', pero yo nunca reclamé, era como lo que menos me interesaba en ese momento...

R: ¿crees que de verdad las evaluaciones que tú tienes miden si tú aprendiste?

S03: Si aprendí...

R: ¿Si ese número refleja de verdad lo que tú aprendiste y lo que sabes?

S03: No, para nada, yo encuentro que los profes de este colegio ponen como mucho interés en la nota que tienes y como que te meten el miedo con que si te va mal, tu NEM, no vas a entrar a la U, pero al final no sirve de nada tener buen promedio si no aprendes o si el conocimiento es vago...

R: ¿y sientes que lo que has aprendido en el colegio te ayuda de alguna forma para la PSU?

S03: No.... Es que igual yo tengo la comparación con el preu, como ya he dicho, y en el preu me han pasado tantos contenidos que en el colegio nunca he visto y me dicen 'esto entra mucho en la PSU' y en el colegio nunca lo ví...

R: ¿y las clases acá no estaban enfocadas así?

S03: no, claro, yo encuentro que eso está mal aquí, como que deberían ya desde primero medio empezar a enfocarse en los contenidos que realmente entran en una prueba, porque al final es una prueba que decide tu futuro, aunque eso igual está mal...

R: ¿y sientes que están desconectados?

S03: claro.

R: ¿tú crees que las notas reflejan lo que sabes?

S03: No, para nada, porque al final uno, tenemos habilidades cada uno distintas, en distintas cosas...

R: ¿y sientes que el colegio evalúa en esas diferencias, que evalúa esas habilidades distintas en estas distintas personas que hay adentro de una sala?

S03: No, yo encuentro que es muy general, como que todos tienen que adaptarse a lo mismo, si te va, mal, cosa tuya.

R: ¿y lo describirías como poco variado?

S03: Sí.

R: Porque me decías que era, en el fondo lo recojo porque me decías que eran más que todo pruebas... entonces...

S03: Claro. Sí, poco variado.

R: es difícil que una persona eh, pueda rendir, la idea de que sea variado es precisamente que si a ti te va mal en una cosa, de repente en otra te cueste menos, pero hay poca variedad me dices tú..

S03: Sí, incluso pocos profes como que intentaban hacernos distintas cosas que una prueba, así como infografía, trabajos escritos, presentaciones, pero casi nadie....

R: ¿y te gusta hacer esos trabajos?

S03: no...yo prefiero una prueba....

R: ¿por qué?

S03: porque es más fácil po, y uno se confía, uno dice ya, si no me la sé, la tiro al achunte no más po y puedo contestar. Aunque igual me gusta lo escrito, pero con el límite de palabras era mucha presión, igual no me costaba tanto, pero la de alternativas uno después se acostumbra porque al final es la que siempre te hacen.

R: Oye, y respecto a las diferencias de notas, que hay en un curso o de repente que puedas tener tú misma, en un ramo que te vaya mejor en otro, em, ¿cómo evaluarías tú rendimiento en clases?

S03: mi rendimiento, yo creo que medio porque la verdad no me esforzaba tanto, pero en cuanto a notas y resultados es alto, pero yo encuentro que me esforzaba medianamente para tener esas notas, como le digo, se me hacía mucho más fácil aquí en este colegio entonces me confiaba de que me iba a ir bien, a veces no estudiaba, pero a lo único que le ponía así mucho empeño era para las pruebas semestrales, que eran las finales, porque si me iba mal, me bajaba el promedio, sí o sí...