



HAL
open science

Standardized tests and the development of argumentation competency: Tensions perceived by Chilean mathematics teachers

Manuel Goizueta, Constanza Ledermann, Helena Montenegro

► To cite this version:

Manuel Goizueta, Constanza Ledermann, Helena Montenegro. Standardized tests and the development of argumentation competency: Tensions perceived by Chilean mathematics teachers. Twelfth Congress of the European Society for Research in Mathematics Education (CERME12), Feb 2022, Bozen-Bolzano, Italy. hal-03746870v2

HAL Id: hal-03746870

<https://hal.science/hal-03746870v2>

Submitted on 17 Oct 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Standardized tests and the development of argumentation competency: Tensions perceived by Chilean mathematics teachers

Manuel Goizueta¹, Constanza Ledermann² and Helena Montenegro³

¹Pontificia Universidad Católica de Valparaíso, Chile; mgoizueta@gmail.com

²CMM, Universidad de Chile, Chile; cvlederm@uc.cl

³CMM, Universidad de Chile, Chile; helena.montenegro@uchile.cl

Chile's standardized tests SIMCE and PSU have been criticized for their unintended effects in the classroom, for contributing to growing elitism in education, and for perpetuating social inequality. However, teachers' perceptions about these tests and their impact on teaching practice have seldom been considered. We focus on teachers' perceptions about the tension between the development of argumentation competency proposed by the national curriculum and the standardized tests. Resorting to in-depth interviews, we account for the systemic lack of alignment that mathematics teachers perceive between curriculum and tests. Results show that this leads to an impoverished learning experience, suggesting that curriculum and test design might not be enough to foster curricular intended teaching, and that teachers' perceptions must be taken into account.

Keywords: Argumentation, argumentation competency, standardized tests, mathematics curriculum.

Introduction

Argumentation is key to constructing participative, democratic societies. Notions such as *mathematical literacy* (PISA, 2018) express the expectation that the student-citizens engage in argumentation, drawing on scientific knowledge, attitudes, and competencies to understand reality, and participate in their communities in critical, reflective, and value-based ways. This has driven many countries to consider argumentation – under this or other denomination – to be a competency to be developed as part of citizenship education. The Chilean curriculum proposes argumentation as one of four competencies to develop at all levels of mathematics education (MINEDUC, 2019, 2015, 2012).

In the course of the last two decades, citizenship education has gained momentum through international standardized tests, such as the Programme for International Student Assessment (PISA), or the Trends in International Mathematics and Science Study (TIMSS). These tests have stimulated the need to assess competency development in students as part of educational system quality assessment efforts. Argumentation being one of the competencies the Chilean curriculum aims to develop, both the Education Quality Measurement System test (SIMCE) –used nationally to assess the quality of compulsory education at various school levels – and the University Selection Test (PSU) – given nationwide in the 12th grade to rank students for access to higher education, are expressly meant to be measuring it.

However, what is stated in the curriculum and assessment systems does not directly translate into classroom practice. Crucially, it is teachers who interpret them and make the pedagogical decisions that shape students' learning experience. They make these decisions while experiencing the tension

between what is explicitly demanded by the educational system and what they perceive could best help students succeed in standardized testing. In contexts of high-stakes standardized testing – such as Chile – the resolution of this tension has enormous consequences.

Regardless of the subject, studies on the impact of high-stakes testing consistently report narrowing of the curriculum, increase of teacher-centered pedagogies, and abandonment of high-quality forms of instruction in favor of routine practices (Au, 2007; Blazar & Pollard, 2017; Meckes, 2007; Valli & Buese, 2007). Poor-quality “teaching to the test” has often been considered a consequence of standardized testing (Darling-Hammond & Adamson, 2014), even though test advocates insist that cognitively demanding tests may motivate teachers to improve and enrich instruction (Au, 2007, Peery, 2013). In the case of mathematics, previous studies at various school levels have found that test preparation activities predict poorer-quality and less ambitious mathematics instruction (Blazar & Pollard, 2017; Cohen & Hill, 2008; Meckes, 2007).

Most of these studies have resourced to classroom observation to evaluate teaching quality. Although such studies are crucial to describing high-stakes testing effects in the classroom, they typically neglect teachers’ voices. Besides, they tend to consider classroom activity as a whole and do not focus on specific curricular aspects. In the present study, we draw on in-depth interviews, exploring the perceptions of Chilean mathematics teachers with respect to the tension between development of argumentation as a curriculum-proposed competency and the national standardized tests SIMCE and PSU, and its consequences in classroom teaching practice. We first present aspects of the Chilean curriculum and programs, and of the SIMCE and PSU tests. Then we explain the methodological design of our study. Finally, we present and discuss our results to account for the systemic lack of alignment that teachers perceive between the national curriculum and the standardized tests in terms of the pedagogical work necessary to develop students’ argumentation skills in the classroom.

The argumentation competency in the Chilean curriculum

The national Curricular Bases (CCBB) designed by the Ministry of Education (MINEDUC) constitutes the cornerstone of the educational system. They provide a mandatory, standard framework for education in all schools, presenting educational principles, and establishing learning goals (LG) for each level of school instruction. In the case of mathematics, the CCBB establishes specific and progressive LG for the development of argumentation (MINEDUC, 2019, 2015, 2012). Table 1 presents the LG for grades 8 and 11-12.

8 th grade (MINEDUC, 2015, p. 112)	11 th -12 th grade (MINEDUC, 2019, p. 108)
<ul style="list-style-type: none"> ● Describing mathematical relations and situations verbally and using symbols. ● Explaining and justifying: (i) Students’ own solutions and procedures; (ii) Results using definitions, axioms, properties, and theorems. ● Backing conjectures using examples. ● Assessing the arguments of others, giving reasons. 	<ul style="list-style-type: none"> ● Making decisions based on statistical evidence and/or results obtained from a probabilistic model. ● Using symbolic language and other representations to justify the truth or falsehood of conjectures and assessing the reach and limits of arguments.

Table 1: Learning goals for argumentation, grades 8 and 11-12

Within the CCBB, argumentation is understood to be part of a dialogue, expression of ideas, and collaboration aimed at convincing others of the validity of mathematical results. Students are expected to transit from an informal/intuitive to a formal/mathematical style of argumentation (MINEDUC, 2015). However, the CCBB does not provide teachers with a clear operational definition of argumentation that allows them to identify instances of the competency and to progressively assess its development.

SIMCE, PSU, and assessment of the argumentation competency

The Quality Education Agency (ACE) designs and implements the SIMCE test as part of the National System of Quality Education Assurance (SAC). The objective of SAC is “to ensure access to quality and equity in education for all students in the country, through comprehensive evaluation, pertinent inspection, and constant support and guidance to educational establishments” (CNED, 2021). SIMCE is administered in the 2nd, 4th, 6th, 8th, 10th, and 11th grade, and encompasses several subjects, including mathematics. As a basis for the SIMCE test, the Ministry of Education develops Learning Standards, which “describe the degree of curricular achievement required for evaluation and monitoring of students’ performance in order to provide feedback to educational policy and to the school system” (MINEDUC, 2017, p. 5). The Learning Standards should thus allow teachers to monitor and plan for their students’ progress following the CCBB. However, and despite declaring that students must develop “problem-solving, representation, modeling and argumentation competencies” (MINEDUC, 2017, p. 50), no specific competency indicators are presented, nor are milestones in expected competency development.

The Department of Educational Assessment, Measurement and Registration (DEMRE) is responsible for developing the PSU test for the process of admission to Chilean universities. On its website, DEMRE declares that the test will assess contents and competencies in line with the current curriculum. The mathematics test “will have 65 multiple-choice questions with a single answer among 4 or 5 choices, (...) it will assess the Understanding, Applying and Analyzing, Synthesizing and Evaluation competencies, (...) and will contain 13 questions that assess the competencies specified in the Curricular Bases: Problem-Solving, Representing, Modeling, and Argumentation” (DEMRE, 2020). On its website, DEMRE refers to argumentation as “the ability of the applicant to evaluate procedures, deductions, solution strategies, and inferences in various problems; to distinguish and detect erroneous arguments; and to understand chains of logical implications”. However, no further information is presented about how the test allows argumentation assessment.

We must also consider the consequences that PSU and SIMCE results have, turning them into high-stakes measures for students, schools, and teachers. PSU results determine access to higher education and career options. A university degree is seen as a strong socioeconomic mobility factor in Chile, hence test results are crucial both for students and their families. Although the SIMCE test does not have direct consequences for students, results are clustered by schools and made public on official and media websites, *de facto* functioning as a school ranking tool. This affects families’ preferences – hence student enrollment – economically impacting state-financed schools through a per-pupil voucher system based on attendance (Taut et al., 2009). Teachers’ income is also often affected through school economic incentives linked to students’ test performance (Botella & Ortiz, 2018).

Methods

The study used a qualitative methodology with a multiple-case study approach (Yin, 2009) to explore recurrent themes addressed by participating teachers during semi-structured interviews. A purposive sample of mathematics teachers (Neuman, 2006) was obtained (n=21) using two sampling criteria: i) the subjects were teaching in educational establishments located in Chile's Metropolitan and Valparaíso Regions, and ii) they taught mathematics at the 6th and 12th grade (12 and 17-year-old students respectively), in which the SIMCE and PSU tests are applied. The teachers' ages ranged from 24 to 49, and their teaching experience from 1 to 20 years.

The interview protocol included questions to investigate teachers' perceptions about argumentation and its teaching in the mathematics classroom; decision-making processes aimed at developing the argumentation competency; the relationship between argumentation competency development, and SIMCE and PSU; and the impact of these tests on teaching practice. The interview time with each participant ranged between 32 and 109 minutes. The set of basic questions was common to all, but follow-up questions differed based on answers, explaining the differences in interview duration. All interviews were recorded and transcribed for analysis.

A thematic analysis of the interviews was carried out by two researchers in order to identify recurring patterns and themes in the data and allow for triangulation (Braun & Clarke, 2006). The first stage involved open codification (Strauss & Corbin, 1998) of each of the interview transcripts. Each researcher independently identified conceptually similar codes and interview excerpts, grouping them into preliminary analytical categories. Initial coding was compared and discussed among the researchers for purpose of validation and adjustment, thereby achieving greater reliability. Then axial codification was performed on the whole data corpus (Strauss & Corbin, 1998). Key themes related to main categories were identified through an iterative process of inductive/deductive analysis, enabling elaboration of a thematic map of the analysis (Braun & Clarke, 2006). A systematic data review was conducted in order to identify negative cases, thus avoiding any confirmation bias. Finally, results were prepared in the form of narrative themes (van Manen, 1990) addressing teachers' perceptions about tensions between the development of argumentation competency and SIMCE and PSU tests, and their consequences in the mathematics classroom.

We present three main related themes and include as many excerpts from teachers' interviews as possible to allow their voices to elaborate on them. The original excerpts are in Spanish. We have tried to convey the sense along with idiosyncratic nuances when translating them into English. In the following section, secondary teachers are referred to as ST, while elementary teachers as PT.

Findings

The importance of argumentation in the mathematics classroom

There is consensus about the importance of argumentation in the mathematics classroom among participating teachers. Teachers associate argumentation with self-regulation and understanding, as opposed to mechanical and non-reflective execution of techniques. In the words of ST1:

ST1: [Argumentation] has to do with realizing how you are thinking. That also allows you to analyze your strategies, to improve and complement them, and to understand others'. Bottom line,

if I understand how I reason, I am also able to explain it, and being able to explain it means that I have an understanding of what I am doing. [...]. I think that is what we expect of students.

Teachers unanimously emphasize that student's argumentation allows them to assess learning and make pedagogical decisions, since knowledge and reasoning become "visible" in argumentation.

PT5: When asking the child to explain how he solved a mathematics problem, what he did, the calculations he performed, why he did this or that, that is the way I can become aware of the mistakes he makes. I realize how he is thinking, in what way he is reasoning mathematically. That way, I can address his difficulties.

Obstacles to development of argumentation in the classroom

Despite the importance attributed to argumentation, teachers point out various obstacles to implementing activities for its development related to PSU and SIMCE tests. Their answers can be grouped into two recurring and complementary themes, i) they perceive covering curricular contents in mathematics to be the central requirement of the SIMCE and PSU tests – and of the educational system in general, having priority over development of competencies and attitudes; and ii) they feel pressured to dedicating time to activities aimed at improving students' SIMCE and PSU test scores. These two themes refer to demands perceived by teachers that impact their pedagogical decisions, particularly when selecting learning activities. Below, we elaborate on these themes together, since teachers refer to them when explaining the tensions between the development of argumentation competency and SIMCE/PSU tests, and their consequences in the mathematics classroom.

The pressure to cover curricular mathematics content emerges from teachers' interpretation of the curriculum and from the demands placed on them by senior management staff in schools. Moreover, teachers consider the time devoted to covering mathematics content to be an impediment to planning and implementing learning activities related to argumentation, as expressed by PT5 and ST5.

PT5: The Ministry mandates delivery of all contents. They say "you have to teach all the contents within the year," which is impossible, but there is no focus on competency development. [...] So, I have always been focused on delivering content, but not on competency development.

ST5: Both regulations and your school demand content coverage from you, and also that students be prepared to confront the PSU test answering questions correctly and doing well. Thus, one starts cutting away. Are you going to work on argumentation? Really? How much? Once? Twice a month? Damn, twice!?! You will be neglecting practice and application. They (senior management staff) are going to question you. Hey, what's going on?

Teachers perceive great importance in PSU and SIMCE results and explain that in a multifactorial way. As already mentioned, student scores are grouped by school and made public. Local media present this information as a measure of a school's quality and of the chances offered for university admission. Teachers perceive that these results are presented publicly as an indicator of the quality of their teaching, impacting their job stability. The perceived importance of these tests translates into pressure to dedicate time to tasks aimed at improving test scores, pressure that is openly or implicitly exerted by school mid-level supervisors (academic coordinators, department heads, etc.). This leads teachers to adapting classroom activity, particularly those years in which students must take the tests.

ST6: I think the [pressure] mechanisms are not visible or explicit. [...] Because a coordinator won't tell you to privilege [test] results over learning, but one way or another you get the message.

PT7: We have some autonomy in distributing our work and students' work, but always within a framework. Because, at the end we are subject to the standardized assessments, so we make decisions on that basis. One way or another, you lean towards them.

Consequences for classroom mathematical activity

When questioned about the relationship between argumentation and SIMCE/PSU tests, teachers agree that the tests do not measure argumentation competency and that their focus is on memorization, calculation, and mechanization of procedures. Crucially, they believe that in order to improve test scores, it is necessary to implement activities that are similar to the tests in the classroom. This leads them to frequently implement multiple-choice exercises similar to those that appear in the tests, leaving little to no time for other activities. In their view, the tasks they feel compelled to implement do not stimulate argumentation or are even detrimental. In summary, these teachers perceive a systemic lack of alignment between the tests and the curricular requirement to develop argumentation. Moreover, they do not just consider that argumentation is neglected in this process, but also that the overall student learning experience is impoverished.

ST6: The school where I work is focused on scores. Programs are pretty bulky, loaded with content, and focused on standardized tests, SIMCE, PSU. [...] For this reason, this competency (argumentation) has not been developed. In the end, mathematics work has been centered on closed-ended question tests. [...] I even feel that this forces teachers to constantly think of multiple-choice tasks, and perhaps also that justification is unimportant.

PT2: I was talking with my colleagues the other day. Why should I care whether students practice argumentation in class? The school is going to assess me just by SIMCE results, and SIMCE does not require argumentation. They just ask you multiple-choice questions.

Discussion

The unintended consequences of Chile's standardized tests have been discussed in previous studies. It has been observed how the use of test results to rank schools within a free-market rationale perpetuates the correlation between socio-economic situation and school achievement, contributing to social inequality (Botella & Ortiz, 2018; Inzunza, 2014; Monarca, 2012). Although teachers are expected to use SIMCE test results as a formative, diagnostic tool to improve their pedagogical practice, Taut et al. (2009) have shown that just "a third of the teachers could accurately tell 'good' results from 'bad', and even fewer were able to accurately evaluate change over time and relative status comparing similar schools" (p. 135). With regard to the impact in the classroom, in line with international research (Au, 2007; Darling-Hammond & Adamson, 2014), previous studies have reported narrowing of the curriculum (Taut et al., 2009) and decreased teaching quality (Botella & Ortiz, 2018; Meckes, 2007). However, such studies have seldom taken into account teachers' perceptions of tests and their impact on argumentation. We focused on the tensions between explicit curricular demands regarding development of argumentation competency in the mathematics classroom and the implicit demands arising from teachers' perceptions of PSU and SIMCE tests.

The lack of alignment between the curriculum and the standardized tests perceived by the participating teachers is expressed in two interrelated but well-differentiated ways. On the one hand, they perceive that SIMCE and PSU tests do not allow the assessment of argumentation competency,

since it is not brought into play in the tests in ways that provide evidence of its development. On the other hand, teachers perceive that the best strategy to improve students' scores is to implement multiple-choice tasks similar to those found in the tests, which do not require argumentation. That is, according to this group of teachers, not only is argumentation not measured by the tests, but its development does not help students perform well. Together with the intense pressure to achieve good scores, this results in classroom activities that marginalize argumentation and, in general, impoverish students' educational experience, particularly during the years in which they are tested (Bellei & Muñoz, 2021). Despite being aware of this impoverishment and of the fact that all of them declare that they do so reluctantly, teachers participate in strategies orchestrated by schools to improve test scores. This agrees with other studies that suggest that high-stakes testing often drives teachers "to abandon methods and materials that had been successful with their students" (Bailey, 2000, p.118) and even "engage in practices that were antithetical to their beliefs about a good instructional environment" (Valli & Buese, 2007, p. 552).

It is crucial to note that we did not interview a representative teacher sample, nor we attempted to determine whether SIMCE and PSU tests do measure students' argumentation competency and its development, as declared by test designers. Nevertheless, we have shown that this group of teachers does not perceive this to be the case and that such perception negatively impacts actual teaching practice regarding development of argumentation competency. In line with what Blazar and Pollard (2017), and Cohen and Hill (2008) say, these results suggest that curriculum and test design might not be enough to foster curricular intended teaching, as hoped by test designers and policy makers (Peery, 2013). Our results suggest that teachers' perceptions must be taken into account.

Acknowledgment

This research was supported by Project FONDECYT 11190135, and by Centro de Modelamiento Matemático (CMM), ACE210010, and FB210005 BASAL, all financed by Chile's National Agency for Research and Development (ANID).

References

- Au, W. (2007). High-Stakes Testing and Curricular Control: A Qualitative Metasynthesis. *Educational Researcher*, 36(5), 258–267. <https://doi.org/10.3102/0013189X07306523>
- Bailey, B. (2000). The impact of mandated change on teachers. In N. Bascia & A. Hargreaves (Eds.), *The sharp edge of educational change: Teaching, leading, and the realities of reform* (pp. 112-128). Routledge Falmer.
- Bellei, C., & Muñoz, G. (2021). Models of regulation, education policies, and changes in the education system: a long-term analysis of the Chilean case. *Journal of Educational Change*, 22(3), 1-28. <https://doi.org/10.1007/s10833-021-09435-1>
- Blazar, D. & Pollard, C. (2017). Does test preparation mean low-Quality instruction? *Educational researcher*, 46(8), 420-433. <https://doi.org/10.3102/0013189X17732753>
- Botella, M. T., y Ortiz, C. (2018). Efectos indeseados a partir de los resultados SIMCE en Chile. *Revista Educación, Política y Sociedad*, 3(2), 27-44.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>

- CNED (2021, September 9). Sistema Nacional de Aseguramiento de la Calidad de la Educación Escolar (SAC). <https://www.cned.cl/sistema-nacional-de-aseguramiento-de-la-calidad-de-la-educacion-escolar-sac>
- Cohen, D. K., & Hill, H. C. (2008). *Learning policy: When state education reform works*. Yale University Press. DOI:10.12987/yale/9780300089479.001.0001
- Darling-Hammond, L., & Adamson, F. (2014). *Beyond the bubble test: How performance assessments support 21st century learning*. John Wiley & Sons.
- DEMRE (2020). *Contenidos de las Pruebas de Admisión Transitorias: Prueba obligatoria de matemática*. MINEDUC. Retrieved from: <https://demre.cl/publicaciones/pdf/2021-20-04-temario-matematica-p2021.pdf>
- Inzunza, J. (2014). Estandarización en educación: anatomía de una deformación. *Docencia*, 52, 4-13.
- Meckes, L. (2007). Evaluación y estándares: logros y desafíos para incrementar el impacto en calidad educativa. *Pensamiento educativo*, Vol. 40(1), 351-371.
- MINEDUC (2019). *Bases curriculares III° y IV° Medio*. MINEDUC.
- MINEDUC (2017). *Estándares de aprendizaje. Matemática. Sexto básico*. MINEDUC.
- MINEDUC (2015). *Bases curriculares para 7° básico a 2° medio*. MINEDUC.
- MINEDUC (2012). *Bases Curriculares para la Educación Básica*. MINEDUC.
- Monarca, H. (2012). La racionalidad de las políticas de evaluación de la calidad de la educación. *Revista Iberoamericana de Educación*, 59(1), (1-9). <https://doi.org/10.35362/rie5911407>
- Neuman, W. (2006). *Social research methods: Qualitative and quantitative approaches* (6th ed.). Pearson/Allyn and Bacon. <https://doi.org/10.2307/3211488>
- Peery, A. B. (2013). Reading for the future: How the Common Core will change instruction. *The NERA Journal*, 48(2), 1-9.
- PISA Governing Board (2018). *PISA 2021 Mathematics Framework (second draft)*. OECD.
- Polikoff, M. S., & Porter, A. C. (2014). Instructional alignment as a measure of teaching quality. *Educational Evaluation and Policy Analysis*, 36(4), 399-416. <https://doi.org/10.3102/0162373714531851>
- Strauss, A., & Corbin, J. (1998). *Basics of qualitative research: Techniques and procedures for developing Grounded Theory*. Sage. <https://doi.org/10.4135/9781452230153>
- Taut, S., Cortes, F., Sebastian, C. & Preiss, D. (2009). Evaluating school and parent reports of the national student achievement testing system (SIMCE) in Chile: Access, comprehension, and use. *Evaluation and Program Planning* 32, 129-137. <https://doi.org/10.1016/j.evalprogplan.2008.10.004>
- Valli, L., & Buese, D. (2007) - The changing roles of teachers in an era of high-stakes accountability. *American Educational Research Journal*, 44(3), 519-558. <https://doi.org/10.3102/0002831207306859>
- van Manen, M. (1990). *Researching lived experience: Human science for an action sensitive pedagogy*. Althouse. <https://doi.org/10.4324/9781315421056>
- Yin, R. (2009). *Case study research*. Sage. <https://doi.org/10.33524/cjar.v14i1.73>