

Funded by: Iniciativa Bicentenario U. de Chile y FONDECYT Iniciación N° 11170316

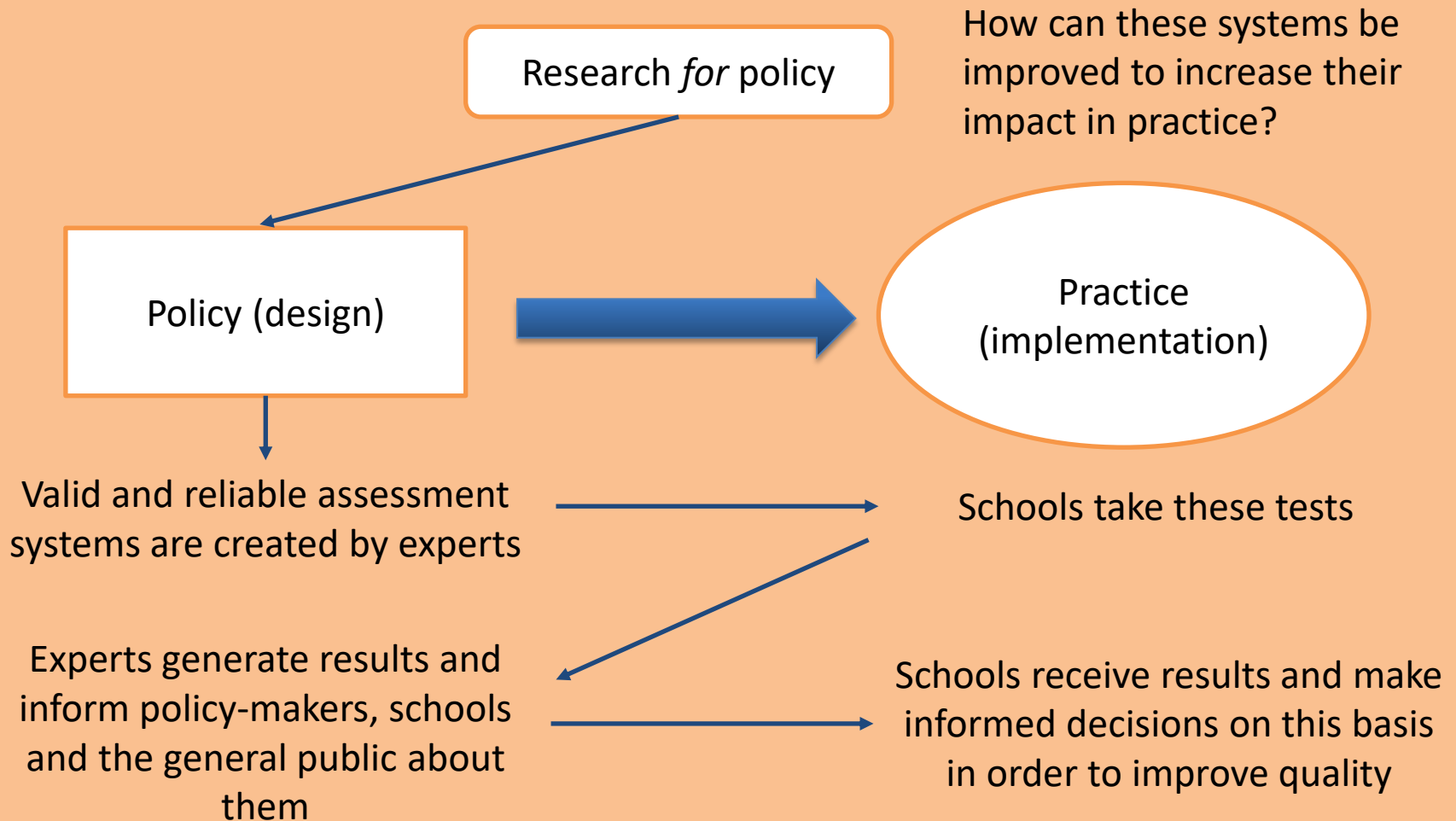
The contested and complex nature of the use of assessment data in decision-making processes: questioning the assumptions of policy



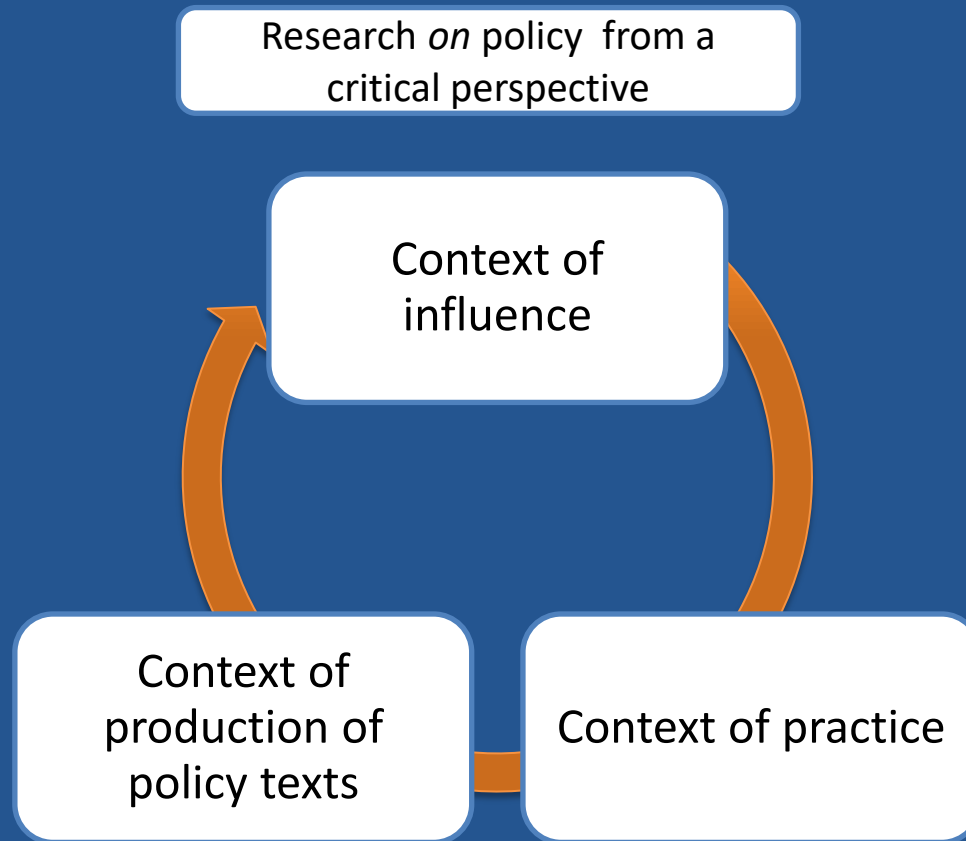
Dr Teresa Flórez Petour
Assistant Professor
Pedagogical Studies Department
University of Chile

Research team: Loreto Montero, Camila Cárdenas, Verónica Carreño, Jennyfer Poblete, Sandra Rivera, Javiera Quiroga, Alexis Echeverría

Assumptions about the use of assessment data in traditional approaches to policy



Questioning assumptions from a complex and interdisciplinary perspective



Who uses (and who doesn't use) assessment data and for what purposes? Whose interests are served by these uses?

An interdisciplinary research programme (2015-2020)

THE PROGRAMME

To develop a systemic, historical and interdisciplinary model of analysis of education policy, in line with the concept of new governance, in order to address education policy from its complexity as a process.

THE PROJECT

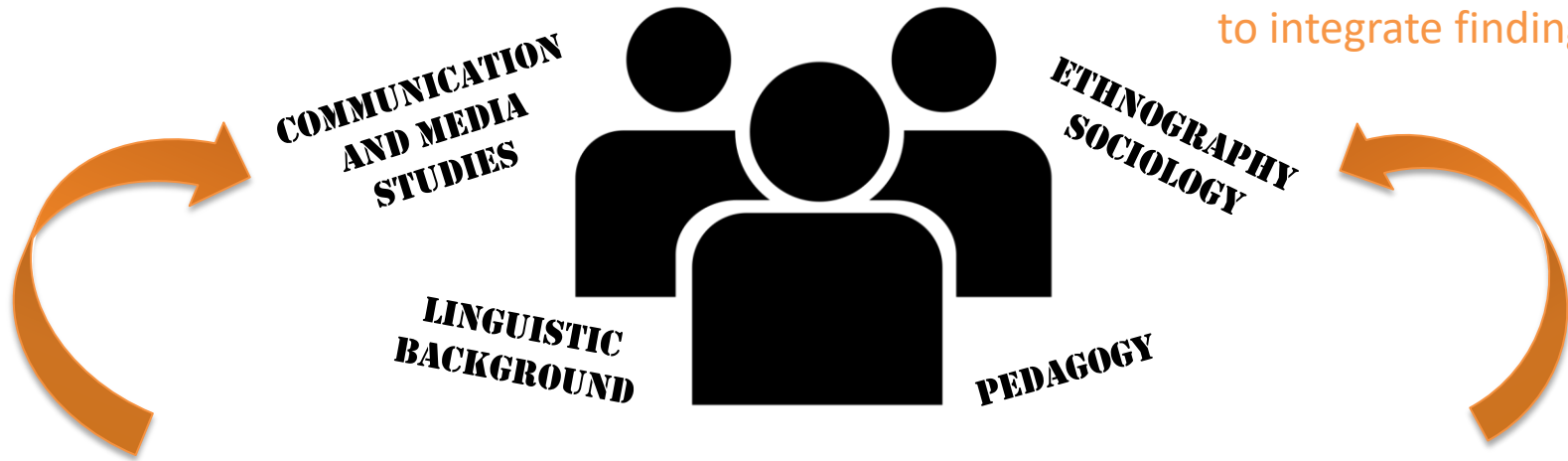
To understand the process of construction of education policy in Chile, taking SIMCE as a case, from an interdisciplinary perspective.

QUESTIONS

- What are the main **actors** and their **interactions** in the definition of policy in SIMCE and the **discourses** they sustain in connection to the test?
- Who **speaks** and **who is spoken for** or (re)presented in policy texts in SIMCE?
- What processes of **interpretation**, **translation** and **resistance** occur in school communities when enacting the SIMCE policy?
- How is the **policy** around SIMCE **constructed** and **produced**, considering the network of actors involved in the whole process, their interactions and the discourses that circulate between them?

MULTIDISCIPLINARY TEAM

Extensive team meetings to integrate findings.



EDITORIALS AND NEWS

- *El Mercurio* and *La Tercera*.
- Timeframe: 2000 – 2014.
- Total sample: 200 texts.
- Framing analysis and CDA.

POLICY TEXTS

- Experts, legal and policy-user documents.
- Timeframe: 1988 – 2014.
- Initial scanning of a sample of 272 texts.
- Sub-sample of 7 texts for detailed CDA.
- Semi-structured interviews with key policy actors.

ETHNOGRAPHIC WORK

- Three vulnerable schools, different performance levels.
- Observation registers, interviews and more informal observations, with emphasis on the process around taking SIMCE in October 2015.

The case: The system for measurement of quality in education (SIMCE)

Context: transformation of education into a market-based system (neo-liberal ideology during the dictatorship of Pinochet)



Created in 1988



Legally instituted in 1990 by the Organic Law for Education (*Leyes de Amarre*)



INITIAL AIMS

- Informing the school market with an indicator of quality that facilitated parental choice and competition for student enrollment between schools.
- Informing policy through its results.

Few critical studies (research for policy)

Growing body of evidence about negative consequences

Accumulation of purposes, increasing frequency, more areas being assessed, increasing incentives/consequences through result-related policies.

1992: administered by the Ministry of Education



2012: administered by the Agency for Quality

SOME FINDINGS

Assumptions about the use of assessment data in policy texts

1988

External standardised assessment systems have a crucial role in the improvement of quality in education, they are necessary and their impact must be constantly improved.

2014



These systems provide information that:

- enables actors who are responsible for results to improve teaching and education quality.
- helps parents to make informed decisions about the school they chose for their children.

In order to generate responsibility in connection to results, the right incentives and consequences must be promoted (competition between schools through a voucher system and funding-related decisions).



These systems are scientifically valid and reliable, they have high credibility and a long-standing trajectory, which legitimates their use. Comparability in time is crucial and technical quality must be constantly improved.

They focus the efforts of actors on learning as established by the curriculum.



MAIN THESES ABOUT SIMCE IN POLICY TEXTS IN TIME

1988	1992	1999	2000	2003	2010	2015
<p>Modernisation of education requires a systematic programme of assessment of education management.</p>	<p>Current measures have not had the expected impact on improving school achievement and equity, and they should be improved in order to achieve this aim.</p>	<p>SIMCE is a necessary assessment system but it needs improvement in order to be able to accomplish its aim of improving quality in education.</p>	<p>SIMCE is an opportunity to know what is not being achieved, it shows weaknesses in order to address them.</p>	<p>SIMCE needs improvement in order to intensify its use and its contribution to the improvement of quality in education.</p>	<p>In the institutional design oriented to ensuring quality in education under discussion, standardised tests are the best quality proxy.</p>	<p>SIMCE is a good assessment system but it needs improvements in the use of its data for its contribution to the improvement of quality in education to increase.</p>

Strategies: «positivisation»



1

Aspects that are portrayed as negative by school communities and some researchers are presented as positive.

«(...) to imprint a sense of urgency for improving the levels of learning in basic areas favours that schools focus their efforts in a sustained manner.» (SIMCE Committee 2014, p. 7)

“(...) I would say it is chaotic at some points. The stress is too much.” (Rosa School Teacher)

“I am worried that my (...) new colleagues, the young ones, won’t have a job, because you know that a school with bad results can even be closed” (Nova School teacher)

“(...) well, they sort of tell you off. The Directorate of Education calls us, arranges a meeting with us with their charts per school. Then it is like discouraging. If one has lowered, (...) One feels horrible. Because one tries to do one’s best.” (Rosa School teachers)



Strategies:

Naturalisation/neutralisation/depoliticisation

2

To present a specific policy/political context for an assessment system as a given, as naturally emerging and, therefore, neutral and potentially unchangeable.

«Since 1980, the successive governments of Chile have (...) introduced multiple large-scale education reforms. In fact, Chile has turned into one of the most mentioned countries as an example of innovative reforms in education. Among the most important reforms are: (i) the decentralisation of state education administration through municipalities; (ii) a change in the funding system towards a grant per student, or voucher, for private and municipal schools; (iii) the implementation of performance-based bonuses for teachers; and (iv) the introduction of compensatory programmes for low-performing schools. (World Bank, 2007, p. 9)

«1. A system framed in a long-standing assessment tradition

(...) The origins of SIMCE go back to the 1960s, when in 1968 the first National Test is taken, aimed at measuring the learning achievements of Year 8 students. (...) In 1988 SIMCE was enacted, a system aimed initially at assessing school performance and the quality of education in Years 4 and 8 of all schools of the country.» (Bravo, 2011, p. 191)

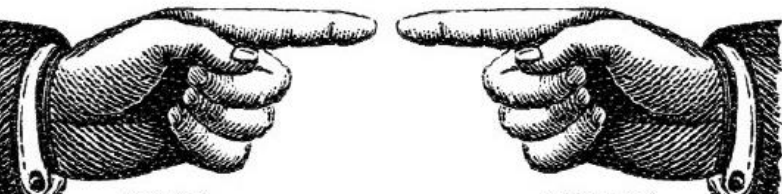
Strategies: Responsabilisation of the Other

3

When something is presented as negative, it is portrayed as the responsibility of someone else outside the assessment system, generally users of data.

«There is a gap in teacher education that makes it difficult for teachers to understand and use measurement results, as the analysis of national and international measurement systems is a subject not sufficiently addressed in initial teacher education and professional development programmes.» (SIMCE Committee, 2003, pp. 11-12)

«Its is important to highlight that the main problems and challenges that were identified by the Committee refer to the uses given to SIMCE by education policy, by the media and schools and to the effects of those uses, rather than the type of instruments of the information that is provided.» (SIMCE Committee, 2014, p. 9)



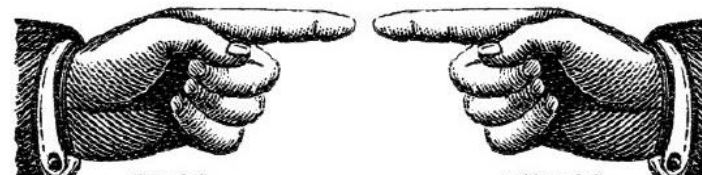
Strategies: Responsabilisation of the Other

«(...) the diagnosis that is being given through SIMCE... [is] worrying due to what occurs inside the classroom. We are with children that have real difficulties and we are being measured by a system that measures all in a general way.» (Sur School Teacher)

«[the school] prepares with fanfare to receive the examiners (...) It all ends up being a show.» (Sur School Teacher)

«The thing is that when results are bad we are questioned but the only thing that the local authority does is <<Who is the teacher responsible for SIMCE>>» (Rosa School Teacher)

«We receive any child with any kind of difficulty (...) Then that is the problem, that in this school they receive any child and that makes it impossible for us to increase SIMCE» (Sur School Teacher)



Strategies: delegitimisation



4

Perspectives that oppose the views of the main voice in discourse are presented as less legitimate.

Presenting resistance in negative terms:

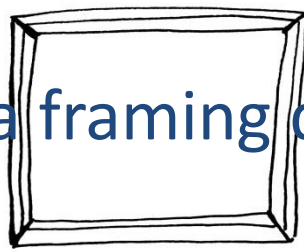
“(...) the attempt of boycott to the test is absurd, as it goes against themselves and their own schools”; “(...) “it is a mechanism that provides parents with useful information in the election of the school for their children; in the same vein, it is used as a factor in the provision of performance-related funding”; “Students must recognise that actions of this sort are counter-productive” (Editorial from La Tercera, referring to students’ resistance to SIMCE)

Resistance = minority:

“very specific cases”, “expressed by minority sectors” (Editorial from La Tercera, referring to students’ resistance to SIMCE)

“There is also a cohort of detractors who state that these assessments are often unfair and have the power of degrading teaching. For many, standardised and transparent assessments are a fundamental tool in improving the quality of education.” (Expert document, 1999)

Strategies: use as a framing device to justify policies



Resultados Simce:
Estancada Educación Municipal

- Mientras colegios municipalizados registraron 238 puntos en lenguaje en 1996 y 1999, en matemáticas bajaron de 245 a 239 puntos. Plantales privados elevaron sus rendimientos.
- Ministra Aylwin dio nuevas acciones para mejorar la enseñanza, pero reconoció falencias en calidad e inequidad en el sistema.
- Magisterio acusó de fracaso en el sector a los gobiernos de la Concertación, especialmente en la reforma. Realidad de los colegios. (C 6)

Cuartos básicos:
Test Simce revela estancamiento de la educación

- Matemáticas registró una baja general, en especial en sectores más pobres, esfumando, quizás, las metas del Gobierno para 2005.

Ministra Jiménez admite que la educación está estancada tras entregar resultados del Simce

- Prueba reveló que 40% de los alumnos de 4° Básico tienen un nivel deficitario de aprendizaje y provienen del sector socioeconómico bajo.
- Expertos coincidieron en la visión de la ministra y sostienen que el Acuerdo de Educación debería mejorar el actual empantanamiento.

Desolador panorama educacional

Se han conocido recién los resultados de la prueba Simce de 4° y 8° básico realizada en 2009. En esta ocasión hubo un esfuerzo interesante por simular dichos resultados a los aprendizajes esperados para cada nivel educativo, y el panorama que emerge de tal esfuerzo es mucho más trágico que el que descansaba sobre puntajes. Pero, al mismo tiempo, es más demecedor.

Quedan en un pie muy relativo algunos avances modestos en los puntajes promedio en 4° básico, en matemáticas y en comprensión de la naturaleza de 8° básico. En las demás asignaturas evaluadas y en 2° medio no se registran demás avances. Y tales puntajes palidecen cuando se contrastan con los niveles de logro efectivo que ellos significan. Así, por ejemplo, en 4° básico solo el 39 y el 29 por ciento de los estudiantes alcanzan, respectivamente, el nivel avanzado en lenguaje y matemáticas. La denominación de este nivel como "avanzado" no debe mover a confusión.

oportunidades. Es más, nuestro sistema escolar se muestra completamente ineficaz a la hora de superar las diferencias iniciales: aprovechando el hecho de que los alumnos que rindieron la prueba el año pasado son los mismos que lo hicieron en 3° básico en 2005, es posible estudiar cómo se modifican las brechas de desempeño entre niños de nivel socioeconómico alto y bajo. En vez de acercarse, éstas aumentaron en ocho puntos en lenguaje y 26 en matemáticas. Estos antecedentes indican que, independientemente de la vara que se utilice, nuestro sistema escolar está lejos de satisfacer las exigencias mínimas que Chile puede plantearse. Esta realidad es especialmente desoladora porque el país ha hecho esfuerzos y destinado importantes recursos para asegurar mayores aprendizajes de sus estudiantes y una distribución más equitativa de los mismos: ha elevado significativamente los salarios de los profesores, establecido la jornada escolar completa,

Nuestro sistema escolar está lejos de satisfacer las exigencias mínimas que Chile puede plantearle. No sólo la calidad está en deuda:



CONSTANT STATE OF CRISIS = NEED FOR CHANGE

Need for increased accountability, economic incentives and consequences, teacher assessment

Economic provision against results (SEP Law), teacher assessment (pending)

Accelerate the approval of the General Law for Education (LGE) and change teaching profession regulations

Incentives against results, teaching career law, accelerate the approval of the Quality Assurance System Law

Strategies: silencing evidence against assumptions

External standardised assessment systems have a crucial role in the improvement of quality in education.



These systems are scientifically valid and reliable, they have high credibility and a long-standing trajectory, which legitimate their use.



Useful information for school communities and parents.



They focus the efforts of actors on learning as established by the curriculum.



In order to generate responsibility in connection to results, the right incentives and consequences must be promoted.



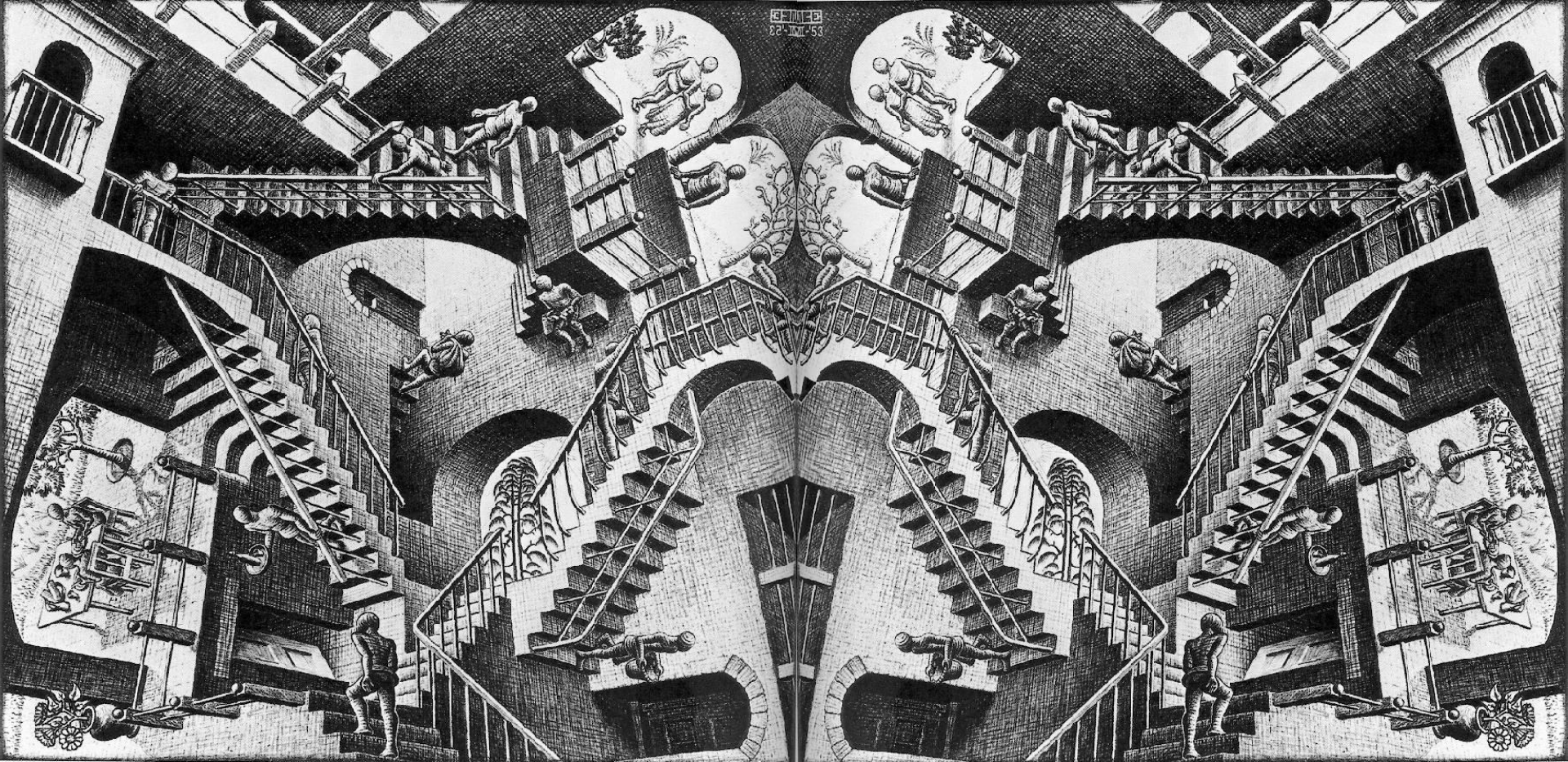
In 30 years of existence, no evidence of SIMCE's contribution to improvement

~~Validity limited to content coverage and~~

Teachers and parents are not using this information (Taut et al, 2009; SIMCE

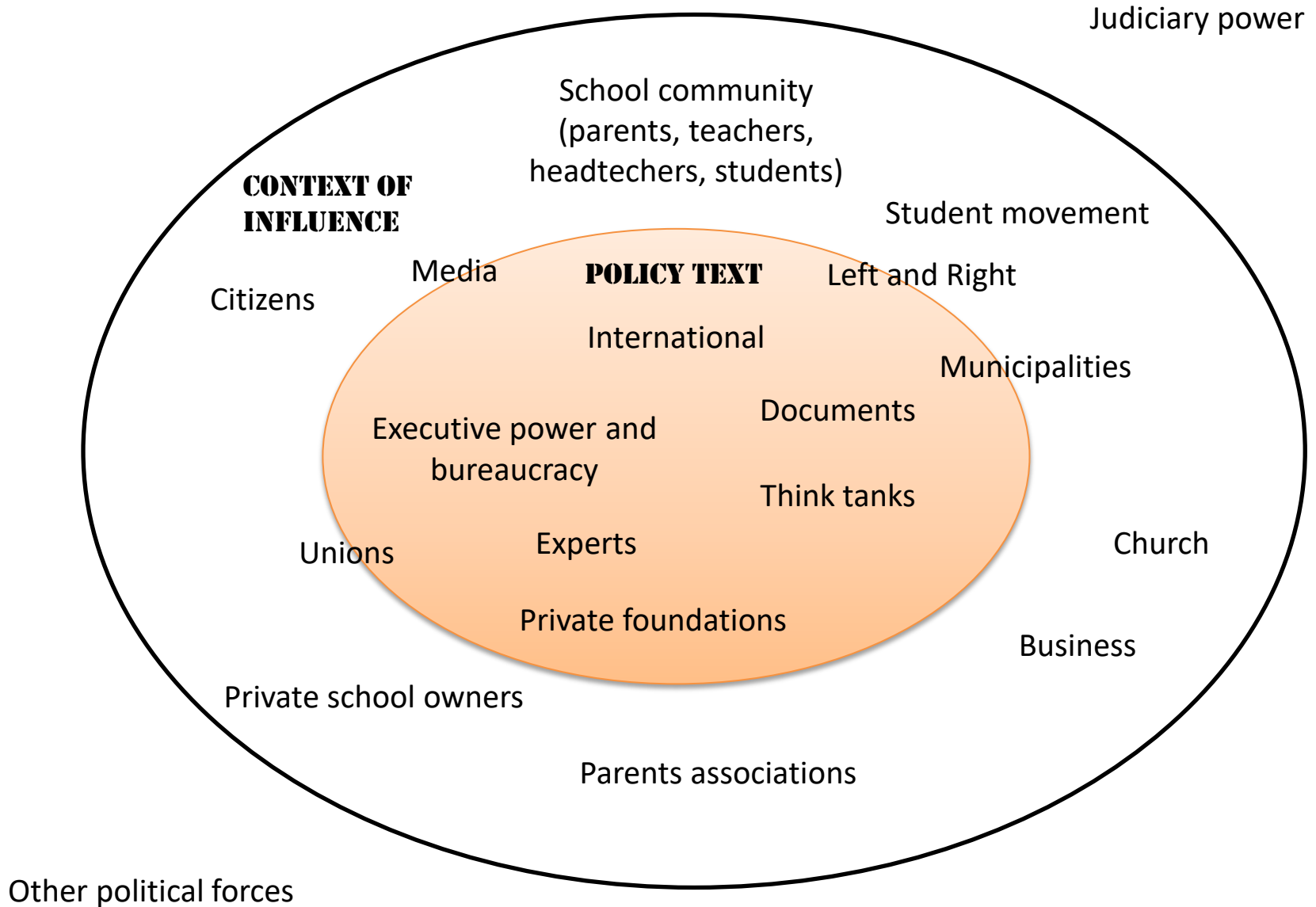
They focus efforts in teaching to the test and mechanistic teaching, hindering more reflective and creative approaches (CIDE, 2012)

Incentives and consequences are damaging the identity and motivation of schools and teachers (Falabella, 2014) and are impoverishing students' educational experience (Flórez, 2013)



WHOSE DATA AND FOR WHAT?: VOICES IN THE CONSTRUCTION OF ASSESSMENT POLICY

FROM THE CONTEXT OF INFLUENCE TO THE POLICY TEXT



Interactions and views between policy actors

«**Interviewer:** *What is the role that you think teachers should have in the construction of education policy?*

Interviewee: Well, the topic of teachers in Chile is quite complicated because teachers, especially primary teachers, don't know much, I mean, it is difficult that they can contribute in a very informed way to a well designed policy. I would spend all the money in the world to teach teachers well but I know is very difficult.»

(Interview with expert, policy-maker and member of private foundation)

«I think there is a weight of the technical above life, then, with all the respect I have for [name of expert in charge of the initial design of SIMCE], she only had a technical perspective, real life was not there and in real life this technique didn't work and she realised this later. But then in the Committee she had a bit, like this guided her to coordinate the Committee.»

(Interview with former member of the studies area of Chile's Teachers Union)



«I mean, I think that what one has to look at there is consequences, not the test. If the test is well constructed and what is the aim of taking it, I mean, consequences are political decisions. Then if political deviations, how much do these political consequences impact, right? But, I mean, large-scale policies have a framework that goes beyond SIMCE anyway, right?»

(Interview with technical expert in charge of developing SIMCE)

«Well, the Committee, in these things you normally end up summoning people who know about the topic and those are usually the ones that are convinced, so obviously the Committee didn't have a very strong position with regards to SIMCE and there were some minority perspectives where we thought that we had reached such an extreme situation that our advice was to stop SIMCE for 3 or 4 years, rethink, redesign (...). This perspective didn't succeed (...) There I realised there was another influence in maintaining SIMCE, which is education researchers (...) For people who works from more quantitative perspectives, these thing of interrupting measurements, of not taking them every year, makes them quite upset.»

(Interview with academic expert)



The construction of a false consensus



STUDENT ORGANISATIONS



TEACHERS

Inclusion/exclusion:
confirmationst bias

EXPERTS COMMITTEE 2014



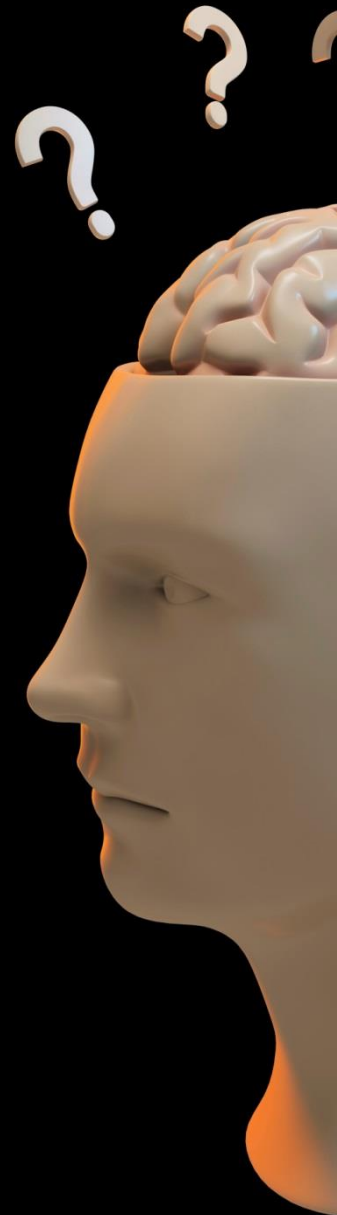
Among declared motivations:
students, social movements and some researchers and academics being critical about SIMCE



SOCIAL AND CITIZEN ORGANISATIONS

CONCLUDING REMARKS

- If validity refers to interpretations and actions derived from results, can validity studies in relation to the design of assessment systems ignore actions and uses derived from policy like they were external to the test?
- Is it ethically and technically valid to use the results from a test to make decisions like the ones that characterise current intended consequences of assessment systems? (salaries, closing schools, funding allocation, school choice)
- Who is eventually using the information derived from large-scale assessments in a beneficial and meaningful way?
- What changes should be made in order to promote pedagogically consistent and meaningful large-scale assessment systems?
- To what extent the dimension of participation in policy design affects the scope and type of use that is made of the information generated by assessment systems?
- Can we ignore that the intended purposes, uses and actions of an assessment system are functional to specific ideologies and to particular interests?



A word cloud of the word "thank you" in various languages and scripts. The central and largest text is "thank you". Other prominent words include "danke", "teşekkür ederim", "gracias", "obrigado", "merci", "dank je", "gracias", "dziękuję", "sukriya", "kop khun krap", "arigatō", "dank je", "gracias", "dziękuję", "sukriya", "kop khun krap", "arigatō", "dank je", "gracias", "dziękuję", "sukriya", "kop khun krap", "arigatō".

Other visible words include: **danke**, **teşekkür ederim**, **gracias**, **obrigado**, **merci**, **dank je**, **gracias**, **dziękuję**, **sukriya**, **kop khun krap**, **arigatō**, **dank je**, **gracias**, **dziękuję**, **sukriya**, **kop khun krap**, **arigatō**.

Other visible words include: **danke**, **teşekkür ederim**, **gracias**, **obrigado**, **merci**, **dank je**, **gracias**, **dziękuję**, **sukriya**, **kop khun krap**, **arigatō**, **dank je**, **gracias**, **dziękuję**, **sukriya**, **kop khun krap**, **arigatō**.