

Playing the Grading Game

Teachers, Performativity and the Dispute over Grade Inflation in Portugal

CO/V/STATED Project, sponsored by FCT. WWW.CONSTATED.PT

2023.14129.PEX; doi.org/10.54499/2023.14129.PEX

Tiago Neves, Cibelle Toledo, Gil Nata and Alexandra Doroftei. University of Porto and University of Trás-os-Montes e Alto Douro Correspondence to tiago@fpce.up.pt



From the visible face of statistics to their production

The CONSTATED project

- Asks how educational statistics are socially produced — not simply read off the world.
- Indicators (grade inflation, retention, dropout) are seen as artefacts of contested definitions, organisational routines and power relations.
- How were these numbers produced? Why these figures and not others?

Prior stage (indicator mapping)

- National & international dropout/retention indicators are conceptually fragmented: different populations, thresholds, infrastructures.
- Aggregation and discontinued series create analytical opacity (“ghost students”, one-rate reporting).
- Measurement choices have equity effects: they make some trajectories visible and others invisible.

This strand (the producers)

- Moves from the published number to the agents closest to its production: teachers.
- 18 semi-structured interviews on grade inflation, retention and dropout in Portuguese schools.
- Goal: the invisible side of the figure - “the icing under the cake”.

Statistics can enable transparency — and produce blind spots. This study looks at where the grade itself is made.



Numbers as conventions, grading as performance

Porter - Trust in Numbers

- Quantification as a technology of trust: it travels where personal judgement is distrusted.
- The grade substitutes for face-to-face trust; and, in travelling (rankings, access averages), it acts.

Desrosières - The Politics of Large Numbers

- Measures are conventions stabilised through work, not mirrors of nature.
- “Investment in form”: grids and formulas convert contested judgement into defensible numbers.

Espeland & Stevens - Commensuration

- Rendering different qualities comparable on one scale is a social process.
- Commensuration transforms what it measures: internal grades vs. exams, public vs. private.

Ball & Biesta - performativity and its costs

- The teacher between authentic pedagogical judgement and legible performance for inspection, rankings, parents.
- What is measured can quietly displace what is educationally valued.

These lenses interpret, rather than overwrite, what teachers say.



Research question — and the translate-and-constitute tension

Research question

- **How do teachers make sense of grade inflation, and what do their accounts reveal about grading as a situated practice within a performative system?**
- Retention and dropout as the backdrop: in teachers' talk the three phenomena are inseparable.

“artificial non-retention also amounts, in a sense, to grade inflation”

E17 — Portuguese/English, public TEIP

The analytic spine

- Teachers' accounts are subjective: situated, partial, sometimes self-justifying.
- Yet the talk does not float above the world: a voted grade, a re-routed student, a grid “built so no one fails” are real acts with real effects.
- Interviews are read neither as windows onto facts nor as pure discourse: they **translate** a reality and simultaneously help **constitute** it — grades, tracks, university places, certified futures.

The tension is not a caveat: it is the finding the analysis is built to show.



Corpus & participants

18

semi-structured
interviews with
teachers (E01–
E18)

2

sectors: public
(including TEIP)
and private

6

Regions (1st, 2nd
and 3rd cycles
+ secondary)

Snowball recruitment; Algarve, Alto Minho,
Norte, Alentejo, Grande Lisboa, Grande
Porto.

Interview topics

- Definitions and historical evolution of retention, dropout and grade inflation.
- Data collection and processing; the standards governing grading.
- Effects of metrics and performativity on pedagogical work and collegial relations.

Ethics & language

- Profiles reconstructed from self-presentation, kept non-identifying features.
- Verbatim extracts coded in Portuguese and translated for presentation: translation is itself interpretative; untranslatable terms are flagged (e.g. *empurrãozinho*, the “little push”).



Reflexive thematic analysis (Braun & Clarke)

Procedure

- Immersive reading of full transcripts.
- Semantic + latent coding, inductive + deductive, per interview (coding log E01–E18).
- Clustering around central organising concepts; review against the whole corpus; naming.
- **13 working codes → 7 themes**, organised as one argument.

Rigour apparatus

- Traceability: every extract anchored to its interview; per-paragraph coordinates (Pnn, ¶N) kept in the coding log.
- Saturation noted, not assumed: 17/18 interviews feed the three central axes (E05 reduced).
- Two independent codings of the same corpus integrated — convergence as analytic triangulation; strict non-duplication of evidence across themes.

Reflexivity

- Researcher subjectivity as resource: themes generated at the intersection of data and a quantification/performativity lens.
- Analyst's trajectory (quantitative → critical) sharpens attention to construction and power, and risks under-reading teachers' defence of legitimate judgement.
- Dissenting voices treated as findings, not noise (E16).



Seven themes, one argument

1

A field of interpretive dispute

what counts as “inflation” is itself contested

2

Comparing incommensurables

the internal–external gap (and the ranking) as artefact

3

Every grade carries the system

legal + institutional cascade converts demand into upward pressure

4

Two competing normativities

inclusion vs. rigour, under resource scarcity

5

The figure made (in)visible

classification work decides what even counts

6

Material force

grades allocate access and sort students

7

The teacher in the middle

negotiated, relational, moral practice

The translate-and-constitute tension is not an eighth theme — it runs through all seven.



Grade inflation is not one thing: a field of interpretive dispute

The pattern

- Not disagreement about whether inflation exists, but disagreement about what would count as inflation.
- Teachers draw boundaries: legitimate good grades / legal permissiveness / pedagogical or humane adjustment / manipulation as dishonesty (E06, E02, E05).
- **Asymmetric suspicion:** only national exams over internal grades are ever audited. Deflation is real but invisible — E06's student left 12th-grade Portuguese at 8 internally and scored 20 in the exam.
- Latently: a struggle over who gets to define the measure — professional definitions vs. definitions imposed from above.

In their words

“we can have very permissive criteria and not be inflating grades... it is the criteria that are permissive”

E06 — Portuguese, public non-TEIP

“at least 20% of the final Portuguese grade goes to oral assessment... that is not inflation, it is doing what the Ministry tells us”

E18 — Portuguese, elite private secondary

“only when the external grade is lower than the internal one is there concern”

E03 — on asymmetric scrutiny



Comparing incommensurables: the gap - and the ranking - as artefacts

The argument teachers make

- An internal grade aggregates many domains and a year-long trajectory; an exam is one written instrument on one day.
- The MAIA reform widened the internal construct: a gap without misconduct (E01: “different assessments”, not inflation).
- **The figure tracks the instrument, not the knowledge:** students sent to a 2nd-phase exam jumped 14 → 18–20 in a month (E18: “they didn’t spend a month learning that much more”).
- Counter-case: public schools may grade defensively below the exam - deflation, not inflation (E17).
- The public/private league table compares unequally resourced worlds on one scale.

In their words

“it is like comparing a horse race with a bicycle race”

E06 — on the internal/external gap

“they are not comparable things... it is unfair to public schools, creating a pressure they have no weapons to fight”

E17 — public TEIP; E18: “comparing a Fiat with a Ferrari”



Every grade carries the system on its back - the legal-institutional cascade

How the cascade works

- **DL 54/55 shift the focus from the student to the school:** retention becomes “exceptional”; it is now the school that must exhaustively document everything done to avoid it. Legislation makes retaining difficult, and thus retentions fall partly by design (DL 54).
- Pressure arrives as a staircase, not an order: EU/PISA → Ministry → school board → teacher (E11: “the pressure comes down like a staircase”).
- **TEIP schools are under sharper pressure than others:** contractualised targets, colour-coded dashboards, and funding at risk: missed targets mean fewer support teachers, psychologists, materials (E17).

In their words

“the pressure comes down like a staircase... it starts up at the top... ‘retaining a student costs so many thousands of euros’”

E11 — Physical Education, public (TEIP this year)

“a school that fails to meet certain targets... its funding can be cut”

E17 — public TEIP (still teaches with a chalkboard)



Pressure without orders: asymmetric bureaucracy, the “little push”, appraisal

Bureaucratic asymmetry

- The cost of resisting falls on the dissenting teacher; inflating is frictionless.

“if I change the grade and inflate, nothing happens and everyone is happy... if I don’t, my life is made hell — report after report”

E13 — Philosophy/Psychology, public

The little push, normalised

- Colleagues “helping” on test day; councils lifting near-misses; grades nudged to close a cycle (E07, E08).

“we give a little push to whoever is on the edge... it is no longer pressure, it is a given”

E07 — the push as routine

Teacher assessment: genuinely three-sided

- **Coercive:** a 0 on the appraisal descriptor for not ceding (E10); private-sector dismissals (E12).
- **Self-imposed:** peer competition under quotas — teachers “fought each other over grades” (E17), easing after the freeze.
- **Denied:** E16 recalls the results-indicator defeated by a 100,000-teacher movement.
- *The spread is itself a finding.*



Two competing normativities: inclusion and rigour

The tension

- Both duties are legitimate: include and not exclude; maintain standards.
- **The discomfort targets the asymmetry between declared inclusive goals and available resources, and** not the decline of retention itself (E08: the law is fine on paper, “we would need the resources... in practice we’re not sure we have them”).
- Inclusion $\neq$ facilitism: raising a fragile student to a pass can be the harder professional work (E02: “more work of merit for a student to reach a 3 than... a 5”).
- When means are missing, the grade - cheap, flexible, immediately available - absorbs the difference.

In their words

“retention... is a failure on the part of the student, the teacher and the education system itself”

E15 — public TEIP (shared responsibility; “there can be no facilitism” from either side)

“a public school, inclusive for all... but is it, really?”

E15 — interrogating the inclusive rhetoric from inside

The axiological face of translate-and-constitute: grading arbitrates between two duties under scarcity, and the arbitration is inscribed, invisibly, in the number.



The figure is made – and made (in)visible

Classification work makes the number

- Authority displaced upward: votes in the class council, two-value adjustments in the minutes, director overrides - “they take the pen, cross it out and put another grade on top” (E13).
- What is countable is decided too: a mid-year leaver becomes a “Class Transfer” (E14); Schools under 15 students vanish from Infoescolas (E03); enrolment kept alive until the obligation lapses at 18 (E15).

“the class council is sovereign... it goes into the minutes. Who reads the minutes? No one. What comes out is the grade sheet”

E17 — public TEIP

Ethical slippage ↔ exhibitionist performativity

- Front stage / backstage: the clean grade sheet is the display; the engineering lives in grids, votes and minutes nobody reads.
- **Fabrication (Ball):** results curated for exhibition - rankings as advertising, success rates as institutional self-presentation.
- The slippage is incremental: the little push → routine → a grid built so that no one fails (E17: undefined fields - “compliant, proactive, reactive student” - that always tilt upward).
- Ethical erosion and display feed each other: the more the number is exhibited, the more its production must be managed.



A contested object with material force

What the grade does

- Admission currency: exam + internal marks feed the higher-education access average. The pedagogical judgement becomes a gate.
- Sorting: vocational push-out “to influence without influencing directly” (E17); a “draining off” into crowded professional tracks (E15).
- **Polarisation of student quality:** “there is a huge gap between the good student and the weak one” (E15) - segregation between schools 500 m apart; elite-private non-abandonment as class composition, not pedagogy (E18).

In their words

“my son needs the grade because he is going into Aerospace Engineering, and his classmate going into IT no longer needs such a high grade”

E18 — parental logic, elite private secondary

“they made him understand, very insistently, that this was not for him”

E13 — shaping who belongs

The metric does not simply record trajectories: it produces them.



The teacher in the middle: negotiated, relational, moral

Where judgement meets the grid

- Assessment is “always a little subjective”, dressed up as something “exact and mechanised, with the grids and the Excel” (E01).
- Rigour as identity under inversion: “a demanding teacher... is perhaps incompetent” (E06); “victims of my cruel rigour” (E18); a heavy conscience for not inflating (E13).
- The non-discursive erupts through the statistical frame - school life as a “tangle” of relations the outsider reads only as a clean figure (E17).
- Neither victims nor cynics: agents exercising contested judgement in conditions not of their choosing (Ball’s performative subject).

In their words

“I am assessing young people in training, who are people, not numbers in front of me”

E14 — Physics & Chemistry, public

“a student was trying to kill another with a pair of scissors... this is far more pressing than a statistic”

E16 — the day before the interview



A real construction: two near-laboratory cases

E17 — the audit case

- The input is engineered: Excel grids “built so that students do not fail”; the pressure is real.
- The aggregation is faithful: inspections audit; the published statistic reports exactly what was decided.

“the way they got there is based on pressure. But from that point on... what is in there is what comes out”

E17 — constructed input, trustworthy count

E18 — the Covid case

- **Exams made optional (2020) changed the mindset of teachers and students alike:** teaching pivoted from exam-prep to projects and orality; criteria were openly adjusted.
- Remove the external anchor and the figure is visibly reconstituted, and this teacher can narrate exactly how the grades were arrived at.

“the moment the exam was no longer compulsory, everyone’s mindset changes completely”

E18 — elite private secondary

Constructed and real are not alternatives: the inputs are manufactured, the count is audited: the figure is a real construction.



What the indicator sees — and what it produces

A selective instrument

- The gap indicator commensurates what its producers regard as incommensurable.
- Sensitive where inflation is most defensible (exam subjects); blind where teachers say it is most real (non-exam subjects: E03, E13).

A feedback loop

- Rankings, funding and the legal status of retention feed the number back into the practice it measures.
- The indicator helps generate its own object (E09: “a little snake biting its own tail”).
- Once success rates are monitored, success becomes something to be produced - and the grade is the cheapest instrument.

The SES-data mystery

- SES-adjusted ranking figures for private schools appeared with no traceable production chain (Ministry: “zero response” to journalists).
- Even the researchers cannot see how the input is assembled.
- *The project's thesis enacted inside the data.*



For assessment, equity and policy

Reading the numbers

- Read grade-based indicators together with their conditions of production (and of non-production).
- Separate the value of the figure from its existence: transfers without categories, sub-threshold schools, deferred retentions never become statistics.
- Present professional disagreement as data: the three-sided appraisal evidence and the deflation cases belong in the picture.

Acting on the conditions

- Equity: inflation asymmetries feed directly into higher-education access. Who benefits from which gap is a distributional question.
- Support inclusion with proportional resources. Otherwise, the grade absorbs the difference between declared goals and available means.
- Acknowledge the grading pressures schools and teachers face, and work on the contexts that produce them, not only on the individuals.

Adequate accounts of grade inflation attend not only to indicators, but to how grading decisions are made legible — or kept invisible.



A real construction — with real beneficiaries and real casualties.

“Who reads the minutes? No one. What comes out is the grade sheet.”

E17 — Portuguese/English teacher, public TEIP school

Thank you. Tiago Neves, University of Porto — tiago@fpce.up.pt

CONSTATED • 2023.14129.PEX • doi.org/10.54499/2023.14129.PEX • WWW.CONSTATED.PT

