

Class Size and the Erosion of Written Feedback Quality: A Dual-Perspective Qualitative Study of Teachers and Students in Overcrowded Classrooms

by

Koffa Abdouraman

University of Garoua (Cameroon)

akoffaa@yahoo.co.uk

Abstract

Written feedback survives almost every reform agenda in contemporary schooling, yet remarkably little is known about what happens to it once classes become very large. This article examines how class size shapes the quality, frequency and pedagogical usefulness of written feedback, drawing on a dual corpus of 21 teacher interviews and 21 student questionnaires addressing the same problem. The analysis combines thematic coding with descriptive quantification. Teachers reported class sizes ranging from 42.5 to 80 students on midpoint estimates ($M = 59.05$); students, asked about the same classrooms, perceived still more congested rooms, from 60 to 85 ($M = 73.71$). Written feedback was reportedly provided on about 60% of student work across the teacher corpus, with a wide spread (35–85%) tied to workload and enrolment. What emerges is not a simple reduction in quantity but a compression effect: as enrolment rises, written feedback becomes shorter, less frequent, less individualized, and more strongly oriented toward error marking than toward guidance for revision. Students corroborate this pattern almost unanimously, describing comments that are legible but brief, repetitive, weakly personalized, and insufficiently explanatory. Both groups nonetheless maintain that feedback remains valuable, producing a paradox we describe as feedback that is useful but insufficient. Teachers compensate through peer work, standard comment banks, selective marking, collective correction and digital tools; none of these strategies removes the structural burden created by very large classes. The article argues that feedback quality should be theorized as a problem of sustainability, workload and institutional design as much as an instructional one, and draws implications for staffing policy, protected correction time, teacher professional development, and the redesign of feedback systems in high-enrolment contexts.

Keywords: written feedback; class size; formative assessment; teacher workload; qualitative research; student voice

ADVANCING KNOWLEDGE
• IMPACTING SOCIETY •

1. Introduction

Written feedback occupies a central place in contemporary conceptions of learning because it can close the distance between present performance and desired performance, provided it clarifies goals, interprets current work, and indicates what should happen next. In the classic model, effective feedback answers three questions (Where am I going? How am I going? Where to next?) and is expected to be specific, individualized and actionable rather than merely corrective (Hattie & Timperley, 2007).

A strong theoretical case for feedback, however, says nothing about how easily high-quality feedback can be sustained in ordinary classrooms. Research on school-based formative assessment repeatedly shows that meaningful feedback is mediated by workload, institutional expectations, time pressure and the organizational conditions of teaching. In secondary mathematics, teachers report continuing to value written feedback while struggling to sustain it under large class sizes, limited instructional time and heavy workload, which has led scholars to frame the issue in terms of sustainable feedback rather than effectiveness alone (Machisi, Selepe, & Manamela, 2026).

Feedback is also poorly understood when treated as an isolated pedagogical act. Work on classroom writing assessment in school contexts has argued that assessment and feedback are interdependent, that learner participation matters, and that tensions between recommended principles and actual teacher practices usually reflect contextual constraints rather than any deficit of teacher goodwill (TESL-EJ review of Lee, 2017). Studies in language education point the same way: large class size, overloaded curricula, limited class time and exam-oriented cultures reduce opportunities for formative interaction and redirect teacher effort toward class control rather than individualized support (Ma, 2021).

Structural constraints shape the sustainability of marking practices as well. Evidence from school improvement research suggests that redesigning feedback systems to reduce excessive written marking, in favour of more immediate pedagogical interaction, can cut teacher workload substantially, though such shifts depend on school-level coherence, leadership support and professional trust rather than isolated individual effort (Aquinas Teaching and Learning Trust, 2018). Broader formative assessment reports stress that quality feedback must be descriptive, aligned with learning goals and usable by students, while warning that teachers need systemic support if such practices are to become routine rather than aspirational (Education First, 2016).

Against this background, the present article analyzes two complementary corpora centred on the same question: how does class size affect the quality of written feedback? The first contains responses from 21 teachers across primary, lower secondary and upper secondary settings; the second contains 21 student responses from overcrowded classrooms. This dual-perspective design allows the analysis to move beyond teacher self-report and to ask whether the degradation of feedback under high enrolment is also visible from the learner side.

Three questions organize the study: (1) How do teachers describe the effects of class size on the frequency, depth and personalization of written feedback? (2) How do students experience the feedback they receive in large classes? (3) What institutional and pedagogical implications emerge when both perspectives are read together?

2. Method

This study is a qualitative secondary analysis of two bounded corpora. Corpus 1 consists of 21 completed semi-structured interview sheets from teachers; Corpus 2 consists of 21 completed student questionnaires in tabular format. Both instruments address the same focal topic: the impact of class size on the quality of written feedback.

The dataset originates from a convenience sample of teachers and students working in a francophone school system under conditions of chronic high enrolment, spanning urban and rural establishments. The exact national setting has been anonymized to protect participant identities, since the schools are few, locally identifiable, and the teachers' accounts of workload and institutional support could expose them to professional consequences. We treat this decision as an assumed methodological cost rather than a neutral precaution. Anonymizing the country removes contextual information that readers would legitimately want: national staffing norms, curricular pacing requirements, marking regulations and examination pressures all differ between systems, and each would sharpen the interpretation of what teachers describe here. Readers should therefore weigh the findings as an account of feedback under overcrowding in one identifiable type of system, not as a portable description of any national context. Because the available files do not document a full sampling protocol, the dataset is treated throughout as an exploratory convenience corpus rather than a statistically representative sample.

The teacher corpus spans primary, lower secondary and upper secondary levels, public and private institutions, and both urban and rural settings. Teacher experience ranges from 2 to 22 years ($M = 9.8$). Midpoint estimates of teacher-reported class size range from 42.5 to 80 students ($M = 59.05$). The student corpus includes learners in the 3e, 2nde, 1ère and Terminale tracks; student-perceived class size ranges from 60 to 85 students ($M = 73.71$), indicating a shared perception of severe overcrowding.

Analytically, the article combines thematic analysis with descriptive quantification. All responses were first read across both corpora to identify recurrent patterns concerning feedback form, frequency, timing, personalization, perceived usefulness, constraints and proposed solutions. Simple descriptive counts were then used to indicate salience without converting the inquiry into a quantitative design, a strategy that preserves interpretive depth while reporting recurring tendencies transparently.

Trustworthiness rests principally on perspective triangulation: themes identified in teacher discourse were checked against the student corpus, particularly on brevity, standardization and time pressure. Because the corpus consists of short written responses rather than long interview transcripts or classroom observations, the findings should be read as robust perception data rather than as direct behavioural evidence. The convergence between the two datasets is what gives the analysis its credibility; the absence of observational material is what bounds it.

Table 1. Overview of the two corpora

Dimensions	Teachers (n = 21)	Students (n = 21)
Data source	Completed semi-structured interview sheets	Completed questionnaire tables
Educational span	Primary to upper secondary	Lower secondary to upper secondary

Class size	42.5–80 midpoint estimate (M = 59.05)	60–85 perceived class size (M = 73.71)
Institutional profile	11 public, 10 private; 16 urban, 5 rural	Student-reported classroom perspective
Feedback pattern	Average feedback reported on ~60% of work	Feedback mainly after major assessments
Core issue	Workload constrains quality and personalization	Feedback experienced as short and repetitive

3. Results

Three macro-patterns structure the findings: feedback compression under enrolment pressure, feedback standardization at the expense of personalization, and the emergence of coping strategies that redistribute rather than eliminate the burden of written response.

3.1. Feedback compression under high enrolment

The most visible result is a compression effect. Teachers working with larger classes report lower feedback coverage, shorter comments and less time per script. Across the corpus, written feedback is reportedly given on about 60% of student work on average, but the range is wide (from 35% to 85%) which suggests that teachers in comparatively lighter class conditions preserve more formative capacity than those in highly congested settings. Correction time follows the same logic: in the more crowded settings teachers commonly report 2 to 5 minutes per copy, while more moderate class sizes allow substantially longer attention to each script.

What teachers say about the content of their comments makes the mechanism concrete. Several characterize the shift in a vocabulary of subtraction: comments become "shorter", "more general", "more standardized". The three terms are not interchangeable, and the distinction matters analytically. Shorter is a matter of volume; more general is a matter of address, marking the point at which a comment ceases to be about this script and becomes about a category of error; more standardized is a matter of provenance, the moment at which the comment stops being composed and starts being retrieved. Teachers at the top of the enrolment range, working with classes approaching eighty, describe a further stage in which a sizeable share of scripts must be processed rapidly or minimally when workload peaks. The mean percentage embedded in teachers' own examples of diminished treatment is 37.76%, but the spread matters more than the average: some teachers mention only modest reduction, whereas others report that most scripts receive minimal feedback in crowded periods. The analytical point is that class size does not simply add more copies to the pile; it changes the texture of what can be written on each one.

Students confirm this compression almost unanimously, and their accounts locate it in time as well as in length. Feedback arrives after major tests, marked assignments or end-of-unit tasks rather than after everyday exercises, and the comments themselves are widely described as short, brief or fairly short. Read together, the two corpora suggest that written feedback in large classes is not absent. It is rationed.

3.2. From individualized guidance to standardized error marking

A second finding concerns what feedback is about. Students repeatedly report that comments focus on errors, faults and weaknesses, with much less attention to strengths, successful strategies or explicit next steps, feedback that signals deficits without adequately supporting revision. The dominant student experience is

not unintelligible feedback; most comments are considered readable. It is underdeveloped feedback: too generic, too repetitive, too weakly explanatory to function as guidance.

Perceived personalization is particularly low. Students state that many learners receive "the same observations", "similar comments" or "almost identical remarks". On the teacher side this perception aligns with the frequent use of standard comments, global summaries and collective correction. Such practices are entirely understandable under workload pressure, but they move feedback away from the individualized diagnostic function that formative assessment theory places at its centre. What is pedagogically sensible as an emergency response may nonetheless erode the student's sense of being addressed as a particular learner.

Not all teachers are affected in the same way. A minority working with comparatively smaller or more academically specialized classes report maintaining more detailed or more targeted feedback, especially where orientation decisions or high-stakes progression are at issue. This internal variation matters: feedback quality is not determined by class size alone but also by discipline, assessment stakes, teacher experience and the availability of compensatory routines. The overall direction of the corpus is nonetheless unambiguous, higher enrolment pushes written feedback toward standardization.

3.3. Coping strategies: pedagogical adaptation under constraint

Teachers do not respond passively to this pressure. The corpus reveals a repertoire of adaptive strategies: peer work (6 cases), standard comment banks (6), selective or sample-based marking (5), and digital tools (4), alongside collective correction, triage by difficulty and guided self-correction. These show pedagogical ingenuity and a determination to preserve some feedback function even when individualized written response to every script has become unrealistic.

Each strategy, however, carries a trade-off. Peer work and self-correction can increase learner participation, but only where students understand the criteria and teachers can supervise quality. Standard comment banks buy consistency and time, yet they may intensify the very impersonality students criticize. Sample-based marking reduces workload at the price of unequal formative attention. Digital tools are often imagined as the obvious remedy, but the corpus gives little support to that hope: they require infrastructure, training and implementation time that overcrowded settings rarely have to spare.

3.4. The paradox of useful but insufficient feedback

One of the more analytically interesting findings is that students continue to value the feedback they receive even while criticizing its brevity and impersonality. They say consistently that feedback helps them notice repeated errors, improve presentation, organize answers more effectively, or write better in subsequent tasks. Teachers likewise connect higher-quality feedback to better progression, stronger engagement or improved error correction, though they differ in how strongly they claim that effect.

Hence the paradox. Feedback under overcrowded conditions is widely judged insufficient, yet almost never judged useless. Minimal feedback evidently retains some diagnostic power; what it loses is the fuller formative function described in the literature. The problem, in other words, is not whether feedback matters. It is that structural conditions cap the amount of pedagogical value any feedback system can deliver.

3.5. Structural conditions and institutional demand

The corpus points decisively toward institutional rather than individual explanations. Thirteen of the 21 teachers report having received no specific training for managing written feedback in large classes. Asked

what support would make the greatest difference, they most often request smaller classes (9 cases), more specific training (5), digital tools (3), additional correction time (3) and stronger initial training (1). Teachers, in short, do not read the problem as one of motivation; they read it as one of conditions and professional support.

The student corpus reinforces that reading. Students overwhelmingly attribute limited feedback quality to the teacher's lack of time, caused by too many scripts and an overloaded classroom. The convergence is analytically important because it weakens the temptation to interpret weak feedback as a failure of teacher effort or as student dissatisfaction. In this dataset both groups identify the same bottleneck: excessive enrolment generates time scarcity, and time scarcity degrades the pedagogical richness of written response.

Table 2. Cross-perspective thematic convergence

Theme	Teacher corpus	Student corpus
Frequency of feedback	Reduced coverage as class size rises; mean ~60% of work	Mostly after major assessments, not routine exercises
Length/depth	Comments become shorter, quicker, more general	Comments experienced as brief or fairly short
Personalization	Standard comments and collective correction used to cope	Many students report receiving similar remarks
Usefulness	Feedback still seen as linked to progress and engagement	Feedback still helps correct errors and improve later work
Main constraint	Workload, time scarcity, overcrowding	Too many copies; teacher lacks time to detail comments
Preferred solutions	Smaller classes, training, tools, protected time	More detailed comments, more explanation, reduced class size

4. Discussion

A caveat about the frame of reference is needed before any implications are drawn. The feedback models mobilized in this article were, for the most part, developed and validated in well-resourced school systems, largely within the OECD, where a class of thirty counts as large and where marking loads, however heavy they feel, are bounded by staffing norms that the settings described here do not enjoy. Hattie and Timperley's three questions presuppose a teacher with enough time per learner to answer them; recommendations that feedback be specific, individualized and dialogic presuppose a ratio of scripts to hours that this corpus simply does not contain. The studies closest to our conditions are precisely the ones that reframe the question: Ma's (2021) account of Vietnamese EFL classrooms and Machisi and colleagues' (2026) work on secondary mathematics both start from feasibility rather than from effectiveness. We therefore read the international literature here as a repertoire of regulative ideals rather than as a description of attainable practice, and we take seriously the possibility that some of its prescriptions are not merely difficult but structurally unavailable in classes of eighty. Whether the compression effect documented below takes the same form in other high-enrolment systems is an empirical question this dataset cannot settle.

With that qualification, the findings support a strong interpretation: in overcrowded classrooms written feedback is not simply reduced in quantity, it is structurally reconfigured. The corpus shows movement from dialogic, criterion-based, guidance-oriented feedback toward compressed, error-focused, semi-standardized response. This is consistent with feedback theory, which holds that feedback works best when it clarifies goals, interprets present performance and points learners toward concrete next steps (Hattie & Timperley, 2007). When class size and workload reduce comments to brief correction, one or more of those functions disappears.

The study also aligns with research on sustainable feedback. Machisi et al. argue that feedback practices must be understood in relation to feasibility, learner participation and institutional support rather than as teacher acts detached from context (Machisi et al., 2026), and the present corpus reinforces the claim in some detail. Teachers continue to value written feedback; what they cannot do is scale individualized response indefinitely in environments where classes are exceptionally large. Sustainability is not a peripheral concern here. It is constitutive of feedback quality.

The student data add a layer that teacher self-report alone could not supply. Students do not merely complain that feedback is sparse. They identify a specific loss of explanatory detail and personal relevance. This resonates with literature emphasizing the integration of assessment and feedback and the need for learner involvement in making feedback productive (TESL-EJ review of Lee, 2017). Where comments are repetitive or generic, students may still extract some benefit, but the feedback relationship thins out and loses its dialogic character. Overcrowding, on this reading, affects not only teacher labour but student recognition and uptake.

The corpus also resists a purely linear account of class size effects. A minority of teachers in less crowded or more examination-sensitive contexts report a greater ability to preserve detailed feedback, often through selective routines such as sampling or guided self-correction. Policy interventions should not, therefore, rely on a single lever. Reducing class size matters, but so do assessment design, protected marking time, moderation structures and professional development. Ma's findings on the barriers created by large classes, limited time and overloaded curricula are echoed here, particularly in the way teachers describe reallocating effort from individualized explanation toward basic task management (Ma, 2021).

At the level of policy and school organization, the corpus supports a shift away from compliance-oriented marking cultures. Workload-reduction research shows that feedback redesign can save substantial teacher time once schools abandon the assumption that more written marking is always better (Aquinas Teaching and Learning Trust, 2018). Such redesign appears especially urgent in overcrowded systems, but it must be done carefully, because the answer is not simply to write less. It is to create high-leverage forms of feedback that students can understand and use, which is also the burden of formative assessment reports insisting that feedback be descriptive, actionable and systemically supported (Education First, 2016).

Several implications follow. Educational systems facing chronic overcrowding should treat feedback as a staffing and time-allocation issue rather than a matter of pedagogical exhortation. Teacher education and professional development should include explicit preparation for feedback design in large classes: peer-review routines, success criteria, comment banks with adaptation protocols, digital annotation strategies. And institutional policies should distinguish between evidence of marking and evidence of learning, because where teachers are pressured to produce visible traces of correction on every script, the predictable result is superficial compliance rather than formative depth.

5. Limitations

The article rests on a bounded secondary corpus and its limitations are correspondingly clear. Responses are relatively short; the sample is not statistically representative; the dataset includes no classroom observation, no document analysis of marked scripts, and no follow-up interviews. To these must be added the interpretive cost accepted in the Method section: because the national setting has been anonymized, readers cannot situate the findings within a specific regulatory and staffing environment, and comparison with other systems is correspondingly limited.

None of this nullifies the findings. The convergence between teacher and student accounts is strong enough to carry interpretive weight, and it does so precisely on the points where self-report is most vulnerable to bias. Future research could extend the work through classroom observation, collection of actual feedback artefacts, longitudinal tracking of revision uptake, and comparative analysis across schools with different enrolment patterns.

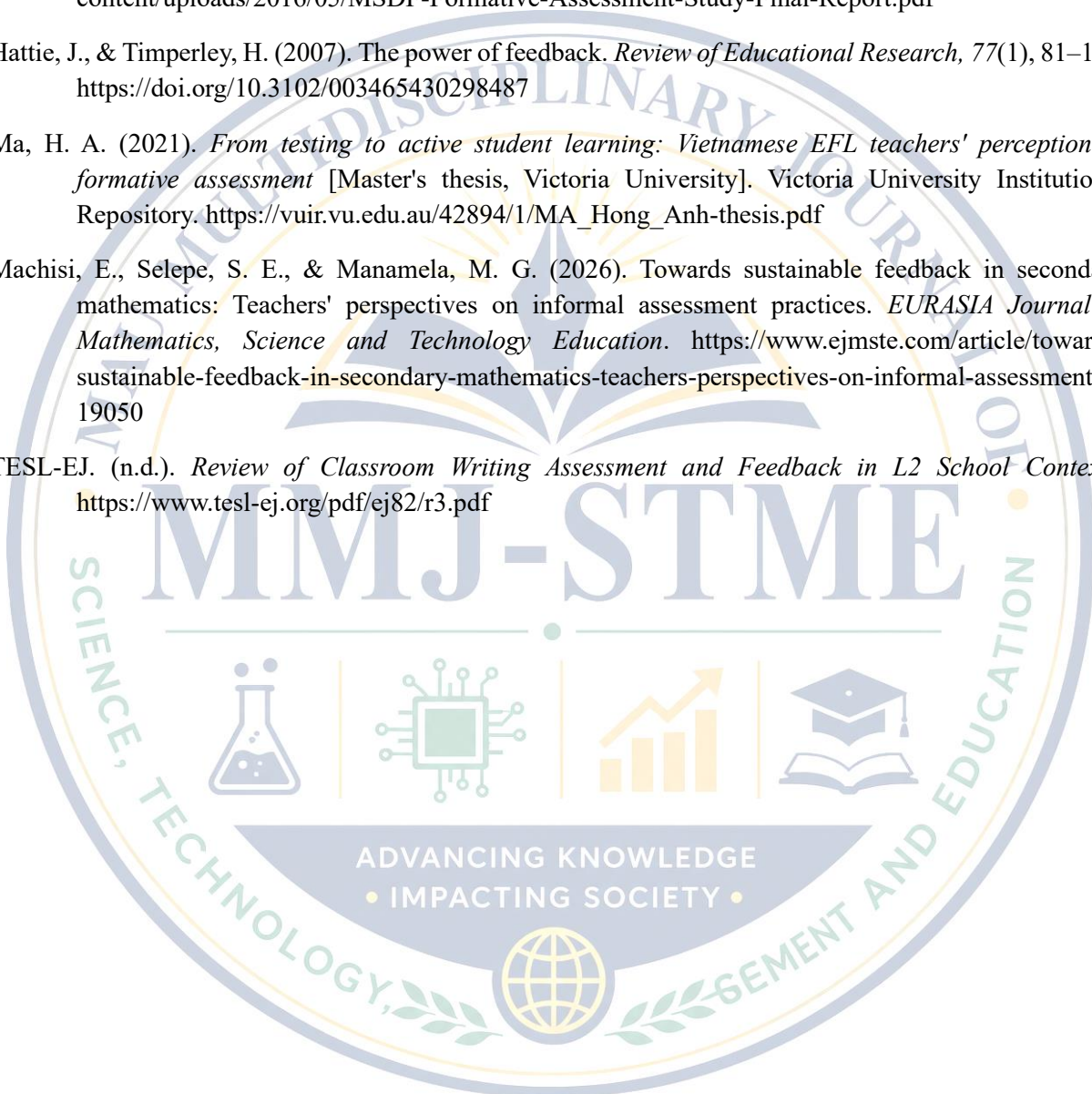
6. Conclusion

This dual-perspective study shows that overcrowded classrooms place written feedback under severe strain. As class size rises, feedback becomes less frequent, less detailed and less personalized, even though teachers and students alike continue to regard it as important for learning. The central contribution of the article is to show that the quality of written feedback cannot be separated from the material conditions of teaching. In large classes feedback is not only a pedagogical technique; it is a labour-intensive practice whose sustainability depends on class size, time, training and institutional design. Any serious effort to improve feedback quality in such settings must therefore move beyond exhorting teachers to do more, and redesign the structural conditions under which feedback is produced and used.



References

- Aquinas Teaching and Learning Trust. (2018). *Reducing teacher workload through "real-time" personalised feedback*. UCL Discovery. https://discovery.ucl.ac.uk/id/eprint/10045200/1/Aquinas_Trust_Reducing_teacher_workload.pdf
- Education First. (2016). *Formative assessment study final report*. <https://www.education-first.com/wp-content/uploads/2016/05/MSDF-Formative-Assessment-Study-Final-Report.pdf>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Ma, H. A. (2021). *From testing to active student learning: Vietnamese EFL teachers' perception of formative assessment* [Master's thesis, Victoria University]. Victoria University Institutional Repository. https://vuir.vu.edu.au/42894/1/MA_Hong_Anh-thesis.pdf
- Machisi, E., Selepe, S. E., & Manamela, M. G. (2026). Towards sustainable feedback in secondary mathematics: Teachers' perspectives on informal assessment practices. *EURASIA Journal of Mathematics, Science and Technology Education*. <https://www.ejmste.com/article/towards-sustainable-feedback-in-secondary-mathematics-teachers-perspectives-on-informal-assessment-19050>
- TESL-EJ. (n.d.). *Review of Classroom Writing Assessment and Feedback in L2 School Contexts*. <https://www.tesl-ej.org/pdf/ej82/r3.pdf>



Appendix. Selected descriptive indicators derived from the corpus

Teacher-reported feedback objective counts: improvement (9), grade emphasis (5), motivation (4), orientation (3).

Teacher-reported support priorities: smaller classes (9), specific training (5), digital tools (3), additional time (3), stronger initial training (1).

Teacher-reported coping strategies: peer work (6), comment banks/standard comments (6), sample-based marking (5), digital tools (4), plus collective correction and triage routines.

Across the student corpus, all respondents linked large class size to reduced feedback quality, typically through the mechanism of limited teacher time and excessive correction load.

