



## Open Research Online

### Citation

Bektik, Duygu (2018). XIPIt: Updating the XIP Dashboard to Support Educators in Essay Marking at Higher Education. In: L@S '18: Proceedings of the Fifth Annual ACM Conference on Learning at Scale, ACM, New York, article no. 38.

### URL

<https://oro.open.ac.uk/55403/>

### License

(CC-BY-NC-ND 4.0) Creative Commons: Attribution-Noncommercial-No Derivative Works 4.0

<https://creativecommons.org/licenses/by-nc-nd/4.0/>

### Policy

This document has been downloaded from Open Research Online, The Open University's repository of research publications. This version is being made available in accordance with Open Research Online policies available from [Open Research Online \(ORO\) Policies](#)

### Versions

If this document is identified as the Author Accepted Manuscript it is the version after peer review but before type setting, copy editing or publisher branding

# XIPIt: Updating the XIP Dashboard to Support Educators in Essay Marking at Higher Education

Duygu Bektik

Institute of Educational Technology

The Open University

Milton Keynes, UK

duygu.bektik@open.ac.uk

## ABSTRACT

Effective written communication is an essential skill which promotes educational success for undergraduates. However, undergraduate students, especially those in their first year at university, are unused to this form of writing. After their long experience with the schoolroom essay, for most undergraduates academic writing development is painstakingly slow. Thus, especially those with poor writing abilities, should write more to be better writers. Yet, the biggest impediment to more writing is that overburdened tutors would ask limited number of drafts from their students. Today, there exist powerful computational language technologies that could evaluate student writing, saving time and providing timely, speedy, reliable feedback which can support educators marking process. This paper motivates an updated visual analytics dashboard, XIPIt, to introduce a set of visual and writing analytics features embedded in a marking environment built on XIP output.

## Author Keywords

Writing learning analytics; visual dashboards; learning analytics; marking interface; undergraduate writing.

## ACM Classification Keywords

K.3.1 [Computers and Education]: Computer Uses in Education

## INTRODUCTION

“The best way to improve one’s [academic] writing skills is to write, receive feedback from an instructor, revise based on the feedback, and then repeat the whole process as often as possible” [5]. This cycle requires tutors to read and provide feedback on student essays, which can create an enormous workload. Assessing written texts is a labour-intensive process for academic tutors. Marking and giving detailed feedback, and commenting on essays can be time-

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [Permissions@acm.org](mailto:Permissions@acm.org).

L@S 2018, June 26–28, 2018, London, United Kingdom

© 2018 Association for Computing Machinery.

ACM ISBN 978-1-4503-5886-6/18/06...\$15.00

<https://doi.org/10.1145/3231644.3231696>

consuming. This problem led researchers to study ways of developing applications that can automatically analyse and evaluate essays for assessment purposes. When assessing their students’ writing, educators will, among other features, be looking for scholarly meta-discourse as an indicator of argumentation. Meta-discourse refers to the features of text that provide linguistic cues which engage the readers, and explicitly convey the authors’ intended meaning, expressing their viewpoint, argument and claim [4].

One approach to automatic identification of meta-discourse, is ‘Xerox Incremental Parser’ (XIP) [1]. The XIP detects and labels rhetorically salient sentences in scholarly writing based on the identification of meta-discourse conveying the author’s rhetorical strategy. The labels are the following:

- SUMMARY (summarising the goals or results of the article),
- EMPHASIS (emphasising the importance of ideas),
- BACKGROUND (describing background knowledge necessary for understanding the article’s contribution),
- CONTRAST (describing tensions, contrasts between ideas, models or research directions),
- NOVELTY (conveying that an idea is new),
- TENDENCY (describing emerging research directions),
- and OPEN QUESTIONS (describing problems that have not been solved).

This paper updates the previous study on visualising the XIP output, and investigates the ways in which the output of the XIP analysis of student essays should be delivered to educators so that they make use of this output to provide feedback.

## BACKGROUND

In order to help readers to make sense of the published literature and to improve reading skills, previously the XIP tool was used to analyse the literature within the LAK/EDM dataset [8] automatically and to identify the important sentences within documents together with the rhetorical type of these sentences. The earlier XIP Dashboard [5] was designed to provide visual analytics from this XIP output to help its end-users, potentially researchers reviewing literature, to assess the current state of the art in terms of trends, patterns, gaps and connections within the learning analytics and educational data-mining literature. This previous study showed that a visual form of

the XIP output can be used to provide a scaffold for making sense of the published literature.

The primary aim for the XIP Dashboard at that time was broadly to explore ways in which the XIP tool's raw output could be visualised in a more user-friendly way. Following that study, the focus turned to possible ways of visualising the XIP analysis of student essays for supporting educators' assessment practices, which is resulted in the working prototype called XIPIt. The aim of the XIPIt study explained next is to support the educators' essay assessment practices and investigate this research question: "In what ways should the output of the XIP analysis of student essays be delivered to educators so that they can make use of this output to give feedback on student essays?"

### STUDY

The design of visualisations starts with a deep understanding of what stakeholders want [3]. As the research question explores the ways in which XIP output can be presented to educators so that they make use of this output to give feedback on student essays, it was necessary to collect user requirements from educators. XIP connects with education to the degree that there is an overlap in the hallmarks of research articles and the kinds of writing that academics seek to nurture in undergraduate students, and reward through grading criteria [2]. The XIP parser, however, was originally designed to work on peer-reviewed academic research writing by people who were not specifically focused on education or on pedagogic value. Therefore, the user and system requirements of a potential XIP visualisation specifically targeting educators and aiming to support their assessment practices needed to be elicited.

### THE USER AND SYSTEM REQUIREMENTS

The aim of the XIP's educator dashboard, XIPIt, is to provide educators with visual analytics reflecting student activity and progress regarding their essays. The visualisation proposed in the previous study, XIP Dashboard, was intended to be adapted for markers whose user and system requirements are discussed below.

#### User requirements

The user requirements for the ultimate educator dashboard were elicited based on the seven one-to-one interviews with markers and one the focus group discussion with six writing teachers and researchers at The Open University, UK. During these sessions, the problems markers had been experiencing regarding assessment and feedback were investigated. The problems of subjectivity in human marking, time limitations, and the need to provide better feedback and examples to ensure students understand their reasoning, became evident in the one-to-one interviews with tutors, the focus group and the literature. Automated support of the XIP embedded in a visual analytics dashboard could potentially solve these problems as follows:

- **the subjectivity of human assessment and marking:** Human markers can disagree with each other and therefore they do not necessarily come to the same conclusion as their peers, which is a reliability concern. The automated output on the other hand is always the same, stays the same and is not subjective. The XIP's educator dashboard, XIPIt, could therefore potentially be useful to help educators overcome this problem by offering self-diagnosis to reflect on their assessment and marking.
- **the limited time that markers have to assess an essay:** Benefiting from the XIP output could also potentially help markers to overcome their time problems. Considering it took around 15 minutes for the focus group participants to highlight 13 sentences and an hour for academic tutors to mark four pages of a student assignment, it could be time-efficient to use the automated support, as the XIP analysis took less than a minute for the same piece. Additionally, automated support could potentially help to overcome the subjectivity of human marking.
- **the possibility that markers do not necessarily notice that their students are actually making an analytical point, since most of the markers are not linguists:** Since the XIP has been developed by linguists whose aim was to analyse the rhetorical elements of salient sentences from articles, it could potentially point markers to the analytical points their students are making with its output.

#### System requirements

The system requirements for the potential educator dashboard were also elicited based on the problems markers experience in the one-to-one interviews and the focus group discussion. During the focus group session, the participants suggested that the XIP should be embedded in a well-designed and developed area, such as course websites or online learning environments. The participants suggested two different dashboards: one for markers and educators, who could input student writing and see what categories the tool assigns, which could then be used to base their feedback on and as a self-help tool to support their communications with their students, i.e. to show examples of good argumentative sentences, and generate discussions; and another one for students who could input their work to revise, edit and critique their own work. Their suggestion has been taken forward for the XIPIt prototype. Considering the scope of this study and decisions made based on focus group participants' suggestions, it was decided to design the educator dashboard. Two other points raised by the focus group participants were taken into account as well: the provision of a manual, and the ability to customise the dashboard. The focus group participants emphasised that there should be an explanation of what the tool's output and results do/do not mean. The meaning should be transparent so that what it does and what its limits are should be clearly stated. Therefore, any

visualisations for the educators should be implemented clearly, explaining what the XIP output means in respect of the students' writing. Since the participants emphasised that what makes good quality student writing is discipline and level specific, there should be an option to customise the tool to adapt it to different levels, disciplines, and modules. This suggested that the ultimate educator's dashboard requires a customisation feature.

### SPECIFICATIONS

For the implementation of the working prototype, a dataset consisting of student writing was required. The student writing used for generating the XIP analysis to be used in the dashboard came from an education and arts module at The Open University, UK (OU). The reason for selecting this module was because of the previous work [6], which demonstrated the relation of students' essay marks with the XIP analysis, and produced promising results about the XIP's performance students' essays that were in line with tutors' marking rubrics.

The detailed specifications for the working prototype of the XIPIt to be built based on the XIP analysis of this module's student writing, were identified through an unstructured interview with the module chair. During the interview, the process of how do educators mark and assess the written assignments was discussed as well as existing problems associated with it.

Since the late 1990s, the OU educators use a special software called e-TMA to download all the submitted assignments/essays which are zipped and downloaded into their own computers. Educators then open these documents with any word processor they have and then mark it based on the rubric. Although all markers use a pre-defined marking rubric, they are allowed to choose the way they give feedback. For instance, they can use track changes, or they can give in-line feedback, and they can choose whatever colour scheme, fonts or caps they want to use. It had been mentioned during the interview that there are inconsistencies between the styles of giving feedback to writing which need to be addressed because all students need to receive consistent and effective feedback. Considering there are more than 60 educators for this module, each choosing their own way of providing feedback, consistency would not be easily achieved. The disadvantage of this could be that if for instance a tutor gives feedback all in capital letters, which is usually perceived as someone shouting in online environments, and using a red font, which is usually used for pointing to an error, a student might interpret this feedback in a negative way.

Once the markers have finished giving feedback to all the student essays that they are supposed to, they then need to save these and give the final essay grade on another online system called PT3 where educators also have an infinite scrollable text box for their overall comments and feedback which is described as time-consuming by the module chair.

As also pointed out in the literature, the one-to-one interviews with markers and the focus group session, markers are pushed for time, so the XIPIt needs to save them time rather than adding to their workload. Finally, markers zip all the essays they have assessed and upload them back to the system for students to see it. From the unstructured interview with the module team chair, the following points have been taken forward for the prototype:

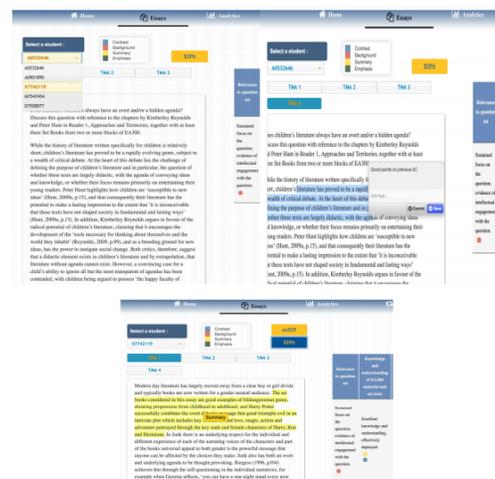
- Markers are inconsistent with the feedback styles therefore the XIPIt needs to provide consistency to address this issue.
- Markers are pushed for time, so the XIPIt needs to save them time rather than adding to their workload. Therefore, it needs to be connected with the systems markers already need to use like e-TMA and PT3.

### PROTOTYPE

The design of the proposed working prototype was informed by the literature, user and system requirements and specifications, as discussed in the previous sections. Based on these, the following design decisions were made.

#### All steps in one

The literature review, one-to-one interviews and focus group study with educators suggested that markers have limited time and they are burdened with the overload of marking they are required to do. Therefore, proposing an additional system, a dashboard that educators need to visit besides their usual marking steps, would take even more time, thus is not plausible. Instead, it was decided that the XIP analysis and output of their students' writing and visual analytics of those need be part of their usual marking steps. That is why the proposed educator's dashboard prototype for markers would allow them to: reach all the essays they need to mark through the dashboard; assess; check the XIP's analysis for these essays; give feedback; and mark them through the same system.



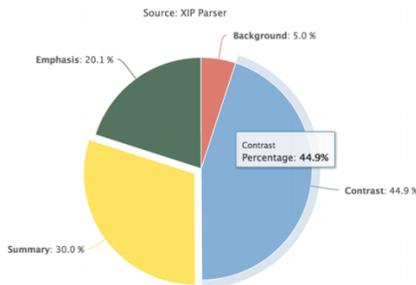
**Figure 1 The XIPIt educator dashboard prototype: selection of student essays, annotating and viewing the XIP output**

As shown in figure 1, when educators logged in to the system, they can see all their students' essays they need to mark. They can select the student whose essay they want to assess. They no longer need to go to e-TMA system to download the essays that they are required to assess as they used to do. They can annotate the students' essay, write comments, give feedback, mark it and save it to the system without going to the PT3 system anymore. Annotating in this way also overcomes the consistency problem that module chair raised in the interview, regarding the different feedback styles markers use.

While assessing and giving feedback, markers can also benefit from the XIP's output to reflect on the rhetorical elements just by clicking the 'XIPit' button, shown in amber in the figure 1. Using the tool in this way as a self-diagnosis to reflect on their assessment and marking, markers, can potentially check their students' analytical points and rhetorical acts and reflect on the XIP's output.

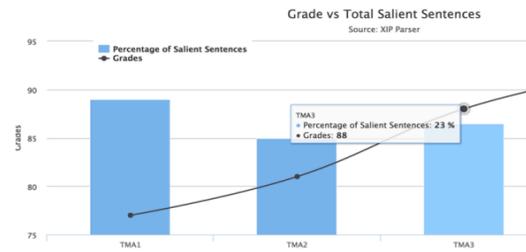
### Analytics

Considering the time problems that markers face, it was decided not to overload them with graphs that are difficult to interpret. Therefore, based on the reflections from the previous XIP Dashboard study and the literature, it was decided to use the popular, widely used and easy to interpret graphs: pie charts, bar charts and line graphs.



**Figure 2 Distribution of rhetorical sentence types of one student's assignment**

As shown in figure 2, in order to provide educators with a quick overview of the distribution of rhetorical elements in a student's assignment, a pie chart was used, which was motivated from the pilot study. Additionally, as shown in the figure 3, tutors were provided with a bar chart combined with a line graph, showing students' overall progress with regard to the XIP output. TMA refers to the tutor marked assignment and in this case the student had three assignments marked by the tutor. The graph shows how this student's grade changed per assignment and what the percentage of salient sentences identified by the XIP was. In this way, the tutor could reflect on how salient sentences affect the essay grade and follow the student's progress.



**Figure 3 Trend of the essay grade compared with the salient sentences**

### CONCLUSION

The ultimate aim of the XIP's educator dashboard, XIPIt, is to provide educators with visual analytics which reflect student activity and progress in respect of their essays. A potential application of such a system by academic tutors would be to use it to inform their judgement, to evidence their judgement and to explain their judgement. This study provided an exemplary prototype for this type of dashboard. The user and system requirements described in this paper can be used as a basis for a potential ultimate educator's dashboard that aims to support educators' essay assessment practices using any computational language technology like the XIP. The XIPIt study proposes an example prototype for such an educator dashboard, hoping that it can inspire further research to implement such a dashboard.

### REFERENCES

1. Ait-Mokhtar, S., Chanod, J.-P., & Roux, C. (2002). Robustness beyond shallowness: incremental deep parsing. *Natural Language Engineering*, 8(2-3), 121-144.
2. Bektik, D. (2017). Learning analytics for academic writing through automatic identification of metadiscourse (Doctoral Thesis). The Open University, UK
3. Börner, K., & Polley, D. E. (2014). *Visual insights: A practical guide to making sense of data*: MIT Press.
4. Hyland, K. (2005). *Metadiscourse*: Wiley Online Library.
5. Simsek, D., Buckingham Shum, S., Sandor, A., De Liddo, A., & Ferguson, R. (2013). XIP Dashboard: visual analytics from automated rhetorical parsing of scientific metadiscourse.
6. Simsek, D.; Sandor, A.; Buckingham, S.; Ferguson, R.; De Liddo, A. and Whitelock, D. (2015). Correlations between automated rhetorical analysis and tutors' grades on student essays. In: *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge*, ACM, pp. 355-359.
7. Taibi, D., & Dietze, S. (2013). Fostering analytics on learning analytics research: the LAK dataset. Paper presented at the *Proceedings of the LAK Data Challenge at The Third Conference on Learning Analytics and Knowledge*, Leuven, Belgium.