Evaluation of the Impact of Automated, Instant Feedback on Writing for Students with Language-based Learning Differences

RB-30716-001A_PAE

Vantage Learning805 Lower York Road

New Hope, PA 18938

www.vantagelearning.com

Paul Edelblut, Jim Lehman, Ph.D., Matt Daneker, MBA

Introduction

As IntelliMetric™ approaches its nineteenth birthday, research into its use has begun to take on a decidedly different focus. Whether or not open-ended assessments can be scored using computers is no longer the key question. Rather, inquiry into IntelliMetric™ centers on questions of how to leverage the highly reliable information provided by IntelliMetric™ into an adaptive learning environment that creates a customized learning pathway for every student. In short, the focus has turned to adaptive learning as a way to transform the educational process.

MY Tutor™ is a patented process of selective writing assessment with tutoring, specific to the individual needs of the user. MY Tutor™ feedback incorporates comments and prompts provided to the user which are generated in response to assessment of the written text. Skill level and developmental level comments, remarks and prompts on various genre specific domains areas are provided in real-time and on-demand modalities as selected by the user. The data presented within this paper provides a valid argument for using MY Tutor™ with IntelliMetric™ scoring for students who have a language based learning disability.

Writing instruction can be thought of as a symbiotic interaction between two equally important participants, the teacher and the student. To be successful in the standard paradigm of writing instruction both participants, the teacher and student, need to put forth substantial effort in a timely manner with a complex set of interdependent variables. In the traditional classroom the nexus of those two participants and the many variables intersecting is extremely low and the teachable moment is often lost before the teacher ever realizes. As part of a technologically driven adaptive learning environment, with MY Tutor™, the very moment a student begins to write the essential elements are brought together in a dynamic, on-demand, adaptive learning environment.

The complexity and difficulty of the writing process is something that many take for granted and most never consider. In total, the process of writing requires a larger amalgam of basic skills than nearly any other activity and exposes the subjects' ultimate understanding of language. This complexity is compounded for educators working in classrooms where approximately 20% of the student population has some form of language based learning disability, undiagnosed in nearly 12%. This research investigates the impact of providing instantaneous, machine generated feedback for this sub-population.

Beginning with the earliest handwriting exercises and continuing through keyboarding where children are expressing their understanding of the English language they must combine complex physical and cognitive processes to render or select letters precisely and fluidly. As writing tasks become more difficult, students must call on an increasingly wide range of skills to not only write legibly, logically, and in an organized way but also to invoke rules of grammar and syntax. This combination of requirements makes writing the most complex and difficult use of the English language.

Students as young as first graders validate their progress with English by writing nearly every day and are asked to do more with this skill than with any other except reading. Writing requirements — from homework assignments and class work to note taking and tests appear to be increasing across the curriculum. Admissions tests and high-stakes assessments for state or Common Core standards have added a variety of writing tasks. The result of the added quantity of writing is a greater emphasis for teachers to teach all students with a heavier emphasis on writing. Yet, many of these students, nearly 1 in 5, have an underlying disability interfering with their ability to maximize the instruction provided.

A language based learning disability can have devastating impacts on an individual child's progress and on a teacher's ability to support that child. In most other school subjects there are reasonably well accepted ways of both quantifying and supporting learning disabilities. School requirements demand a high level of writing

proficiency, and a child who struggles with an unrecognized writing disability will find it increasingly difficult to express his knowledge on many subjects, as the writing process itself will stand firmly in the way of learning.

Through the statistics presented below we demonstrate the positive impact of providing machine generated, instant feedback in a quantifiable manner for students in grades 3 through 8.

Before turning to the study, it may be useful to review what IntelliMetric™ and MY Tutor™ are and what each does. The next section provides an introduction to IntelliMetric™, the feature structure used in evaluation of responses and the links to the MY Tutor(TM) feedback, and a summary of the findings.

Overview of IntelliMetric™

IntelliMetric™ has been shown to be an effective tool for scoring constructed response questions across K-12, higher education and professional training environments as well as within a variety of content areas and assessment purposes.

IntelliMetric™ is an intelligent scoring system that emulates the process carried out by human scorers and is theoretically grounded in the traditions of Cognitive Processing, Computational Linguistics, Natural Language Understanding, and Classification. The system must be "trained" with a set of previously scored responses containing "known score" marker papers for each score point allowing for greater reliability in analytic features. These papers are used as a basis for the system to infer the rubric and the pooled judgments of the human scorers. Relying on Vantage Learning's proprietary CogniSearch™ and Quantum Reasoning™ Technologies, the IntelliMetric™ system internalizes the characteristics of the responses associated with each score point and applies this intelligence in subsequent scoring. The approach is consistent with the procedure underlying holistic scoring.

IntelliMetric™ creates a unique solution for each stimulus or prompt. This is conceptually similar to prompt-specific training for human scorers. For this reason, IntelliMetric™ is able to achieve both high correlations with the scores of human readers and matching percentages with scores awarded by humans.

IntelliMetric™ is based on a blend of artificial intelligence, natural language processing and statistical technologies. It is essentially a learning engine that internalizes the characteristics of the score scale through an iterative learning process. In essence, IntelliMetric™ internalizes the pooled wisdom of many expert scorers. It is important to note that AI is widely believed to better handle "noisy" data and develop a more sophisticated internalization of complex relationships among features. IntelliMetric™ was first commercially released in January 1998 and was the first AI based essay-scoring tool made available to and used successfully by educational agencies.

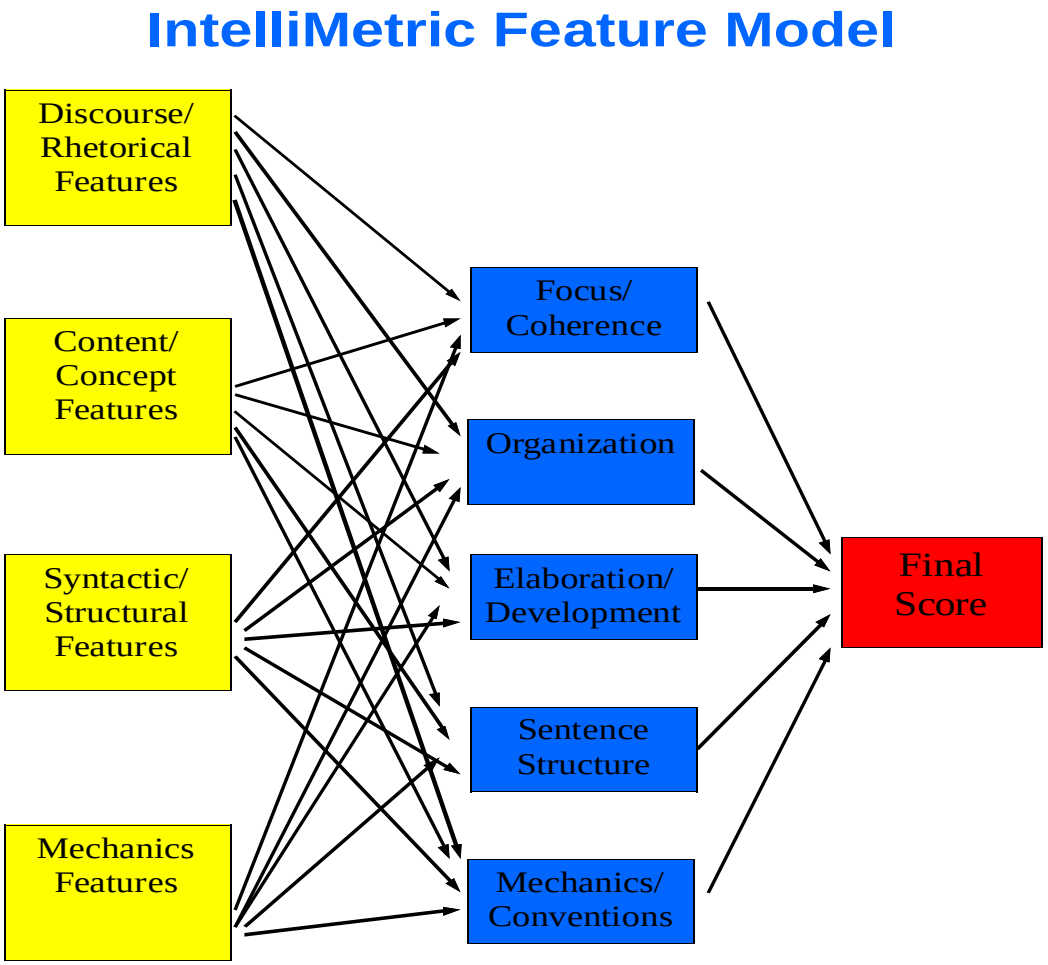
IntelliMetric™ uses a multi-stage process to evaluate responses. First, IntelliMetric™ is exposed to a subset of responses with known scores from which it derives knowledge of the characteristics of each score point. Second, the model reflecting the knowledge derived is tested against a smaller set of responses with known scores to validate the model developed and confirm IntelliMetric™'s ability to accurately and reliably score responses to that prompt. Third, once generalizability is confirmed, the model is applied to score novel responses with unknown scores. Using Vantage Learning's proprietary Legitimatch™ technology, responses that are anomalous either based on the expectations established by the set of essays used in initial training or with respect to expectations for edited American English are identified as part of the process.

Features

IntelliMetric™ analyzes more than ninety semantic, syntactic and discourse level features. These features fall into five major categories.

- Focus and Unity – Features pointing towards cohesiveness and consistency in purpose and main idea.
- Development and Elaboration – Features of text looking at the breadth of content and the
- support for concepts advanced.
- Organization and Structure – Features targeted at the logic of discourse including transitional fluidity and relationships among parts of the response.
- Sentence Structure – Features targeted at sentence complexity and variety.
- Mechanics and Conventions – Features examining conformance to conventions of edited American English.

This model is illustrated in the diagram below.



Overview of MY Tutor(TM)

When students submit their writing, IntelliMetric™ provides immediate assignment scores so that they can begin the process of revision at once. Prompts trained to be evaluated through our IntelliMetric™ scoring system ***with*** MY Tutor(TM), provide students with even more feedback than a comprehensive holistic score, breaking down their submission into each of the standard domains of writing and MY Tutor(TM)® provides goals and examples to help students focus their revision activities through immediate, detailed and developmentally appropriate prescriptive feedback. For each domain, Focus and Mearning, Content and Development, Organization, Language Use, Voice and Style, and Mechanics and Conventions students have a narrative learning pathway that adjusts to their changes.

Additionally the MY Editor® tool, a subset of the MY Tutor™ functionality, helps students in identifying potential errors and recommends corrections in grammar, style, mechanics, and usage.

These robust tools provide text and audio support in multi-lingual delivery during the process of composition, giving students with a learning disability a multi-sensory environment and ELL/ ESL students the opportunity to improve their English writing proficiency by receiving feedback in their native languages.

Students are guided through the recursive writing process based on the feedback they receive along with the various available resources within the system.

MY Tutor™ Research

For this study the data is comprised of nearly 18,400 student essay responses from 42 different states. Students responded to persuasive writing tasks appropriate for the proximal grade level. Students were identified by their school administration as having an **IEP**. Papers were scored on both 6 point holistic and 6 point analytic rubrics. IntelliMetric™ scores and MY Tutor™ narrative feedback were provided both summatively as well as during the drafting process. Narrative feedback was provided on Focus, Content, Organization, Language Use and Style and Mechanics. The essential research question was “Did Students who received MY Tutor(TM) feedback improve their writing from first to last submission?”

Case Processing Summary

	Cases					
	Included		Excluded		Total	
	N	Percent	N	Percent	N	Percent
Avg Holistic Score * Submission Number	18389	99.8%	42	.2%	18431	100.0%

Report of Average Holistic Score by Submission

Avg Holistic Score

Submission Number	Mean	N	Std. Deviation	Kurtosis	Skewness	Minimum	Maximum
1	3.344019	7633	1.1616373	-.358	.187	1.0000	6.0000
2	3.551508	4277	1.1449967	-.500	.074	1.0000	6.0000
3	3.650378	2833	1.1223732	-.444	.085	1.0000	6.0000
4	3.759976	2072	1.0649335	-.400	.010	1.0000	6.0000
5	3.882579	1574	1.0288661	-.353	-.023	1.0000	6.0000
Total	3.532442	18389	1.1447850	-.432	.079	1.0000	6.0000

Reviewing the mean holistic score across five submissions showed a steady and relatively consistent increase with each successive submission.

Furthermore, since the number of students who resubmitted declined as the resubmission number grew, this smaller group continued to see incremental improvement suggesting continued benefit despite the decreasing sample size

Submission Number

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	7661	41.6	41.6	41.6
	2	4286	23.3	23.3	64.8
	3	2836	15.4	15.4	80.2
	4	2073	11.2	11.2	91.5
	5	1575	8.5	8.5	100.0
	Total	18431	100.0	100.0	

.

In the final table we note that, with the exception of 6th grade, scores were clustered in a relatively tight grouping.

Report

Avg Holistic Score

Grade Level	Mean	N	Std. Deviation	Kurtosis	Skewness	Minimum	Maximum
Eighth Grade	3.705670	6206	1.1567836	-.466	.001	1.0000	6.0000
Eleventh Grade	3.665414	556	1.0301313	-.209	-.128	1.0000	6.0000
Fifth Grade	3.720877	342	1.1669140	-.433	.015	1.0000	6.0000
Ninth Grade	3.422683	1431	.9956568	.118	.026	1.0000	6.0000
Seventh Grade	3.557206	5397	1.1667127	-.512	.048	1.0000	6.0000
Sixth Grade	3.178566	3772	1.0849804	-.208	.282	1.0000	6.0000
Tenth Grade	3.758272	486	1.0291620	-.220	-.223	1.0500	6.0000
Third Grade	3.689091	110	1.1294657	-.951	.008	1.2100	6.0000
Twelfth Grade	3.698354	79	1.2751505	-.714	-.002	1.1100	6.0000
Total	3.532186	18379	1.1443993	-.431	.079	1.0000	6.0000

Conclusion

Rarely in educational research, particularly educational research dealing with a population of students who are on an IEP is the data so robust. The evidence of MY Tutor™'s positive impact on student scores for students with an IEP is overwhelming clear.