Writing Analytics: Methodological and Conceptual Developments

Joe Moxley, Founding Editor, *University of South Florida*

Norbert Elliot, Editor-in-Chief, *University of South Florida*

David Eubanks, Executive Editor, *Furman University*

Meg Vezzu, Managing Editor, *The Journal of Writing Analytics*

Matthew J. Osborn, Designer, *The University of Pennsylvania*

Welcome to Volume 2 of *The Journal of Writing Analytics*. As the scholars in this issue demonstrate, writing analytics is emerging as a vibrant field of study. As editors, we are encouraged to see that important methodological and conceptual developments are becoming apparent that serve to deepen and strengthen the field in significant ways. This issue contains seven research articles, two research notes, and a special section featuring research presented at a US educational measurement conference. As was the case with Volume 1 in 2017, our 2018 authors advance a remarkable range of research. And, as was the case last year, this year's authors continue to come from diverse fields advancing focused interest.

We begin by introducing the research of our colleagues and then turn to a reflection on the developments we see in their work.

1.0 Research Articles

Of the seven research articles in Volume 2, the first five are grouped according to studies of undergraduate student writing in the US. The two that conclude this section are devoted to the use of technology to assure fairness in the ways we gather and interpret information.

In "Evolution of Instructor Response? Analysis of Five Years of Feedback to Students," Susan Lang contributes to the new field of analytics by focusing on the use of text mining techniques to support the work of writing program administrators. Undertaken over five years (10 semesters) from August 2012 through May 2017, her study presents findings from 17,534 samples of undergraduate student writing and 141,659 discrete comments by instructors on that writing in order to address a fundamental pedagogical issue: how best to support graduate teaching assistants in their ongoing professional development. Using ProSuite, an integrated collection of text analytics tools, Lang found that instructors over time incorporated a principled, consistent response vocabulary to student writing. This study is especially significant in

DOI: 10.37514/JWA-J.2018.2.1.01

demonstrating that previous work in writing program administration—in this case, the wordlist created by Anson and Anson (2017) to identify high quality and novice responses—can be used across institutional sites to determine whether or not a gap exists in understanding what constitutes valuable instructional feedback.

Hannah Ringler, Beata Beigman Klebanov, and David Kaufer also take a programmatic approach in extending previous corpus-based research in "Placing Writing Tasks in Local and Global Contexts: The Case of Argumentative Writing." Using sociocognitive profiles of writing assignments created by Aull (2017), the researchers examine 438 first-year writing assignments using Docuscope, a platform targeting rhetorical dimensions of texts. Their research is innovative in that it is intended to allow both local and global inferences. While the local analysis revealed key rhetorical strategies that students exhibited between the two first-year writing courses, the global analysis demonstrated that the examined first-year writing primarily differs in academic language, personal register, assertive language, and reasoning from the larger reference corpus. As is the case with the instructor response study by Lang, the generalization study by Ringler and her colleagues explores analyses that are primarily descriptive and diagnostic, rather than prescriptive or evaluative. In both studies, administrators are invited to use such analyses in an iterative process, using the targeted profiles of successful students as a lens to evaluate their writing course sequences. In both studies, we see an expanding granular sense of analysis accompanied by an ability to offer both local and generalized inferences.

Analysis of first-year writing continues with Thomas Peele's "Is this Too Polite? The Limited Use of Rhetorical Moves in a First-Year Corpus." Using 548 research-based argument essays from the first semester composition sequence at City College of New York (CCNY), Peele examines the range and distribution of argument-based rhetorical moves made by students. As is the case with the first two studies in Volume 2, Peele centers his work in a programmatic way—here advancing research by Lancaster (2016) on concession and argument. As Peele finds, students in the CCNY sample make fewer concessions and more counterarguments than their peers, but they do so with a diminished range of linguistic resources. Peele positions the study as a new way to make visible the practices of our writing programs. Taking the findings to the classroom and using them in professional development activities, as he observes, constitutes an important new direction for writing program administration.

Studies of first-year writing continue in "Contemporary Peer Review: Construct Modeling, Measurement Foundations, and the Future of Digital Learning." In their study of 837 first-year student responses to questions on the quality of peer reviews given and received, Ashley N. Reese, Rajeev R. Rachamalla, Alex Rudniy, Laura L. Aull, and David Eubanks demonstrate the importance of construct articulation, the usefulness of fairness as an integrative measurement framework, and the advantages of research on student learning as it occurs in digital ecologies. Focusing on surveys asking questions about self-reflection (how students evaluate a review they gave) and transaction (how students evaluate a review written by their peers), Reese and her colleagues find statistically significant differences among sub-groups. Women students, for instance, awarded themselves higher responses for helpfulness than men. In similar fashion, Hispanic students believed their reviews were more helpful than non-Hispanic students, and

students who claimed dual language proficiency felt their own reviews were more helpful than English-only speakers. In the case of transaction, men were perceived as equally helpful as women in their feedback, and there were no statistically significant differences between non-Hispanic and Hispanic students' reviews in terms of helpfulness. White students, however, perceived the reviews they received as more helpful than did Asian students. As the authors demonstrate, an expanded concept of the writing construct—here, with attention to reflective and transactional impressions of the interpersonal domain of writing—is important to our understanding of how different groups of student writers respond to digital peer-review environments.

Because first-year writing is the largest single site for the collection and analysis of undergraduate writing in the US, there are few corpus studies that expand that sampling plan to investigate how students do over time. We are therefore pleased to present the first longitudinal corpus analysis of undergraduate student writing from the Stanford Study of Writing (SSW). Collected under Andrea A. Lunsford, the Louise Hewlett Nixon Professor of English, Emerita, and the Principal Investigator of the project, the SSW followed a random sample of Stanford University undergraduate students from 2001 to 2006 as they submitted all the writing they did for all of their classes, as well as extracurricular writing. In "Structural Features of Undergraduate Writing: A Computational Approach," Noah Arthurs provides an analysis of the SSW using topic modeling, stance analysis, and parse trees. Latent Dirichlet Allocation was used to identify topic features; metadiscourse markers from Aull and Lancaster (2014) were used to identify stance markers; and Natural Language Processing (NLP) parse trees were used to examine the ways sentence structures of groups of students change over the course of their undergraduate writing careers. As Arthurs concludes, the triangulated technique yields insight into undergraduate writing at Stanford: Students develop most as writers during their first two years, and students of different majors develop as writers in different ways. His hope—that educators will be able to use this kind of longitudinal analysis to understand how their students develop as writers—is at one with all who envision writing analytics as a way to improve student learning.

Our last two research articles focus on tools that ensure fairness in writing analytics. The first is a study of confidentiality, and the second is a study of construct-irrelevant response strategies.

In the protection of human subjects, de-identification, anonymization, and pseudonymization of textual data is an actively studied research area. To contribute to techniques ensuring that writing samples are analyzed without the names of their authors, Alex Rudniy offers a comparative study of three techniques to remove student names from undergraduate chemistry laboratory reports in "De-Identification of Laboratory Reports in STEM." His study is based on a corpus of student work submitted to MyReviewers, a web application purposed for peer review in written communication. Using brute force search with user lists, named entity recognition with OpenNLP, and NeuroNER, he finds that automating de-identification of STEM lab reports is not possible in the case at hand; that is, machine learning toolkits applied out of the box and artificial neural network techniques did not enhance performance of the brute forth approach based on user list matching. His study is exemplary in demonstrating that the

anonymization premise of the General Data Protection Regulation, implemented on May 25, 2018, may not be accurate in all cases. While Rudniy believes that an artificial neural network given an appropriate training set and valid parameters may be capable of learning the difference between names of student authors, peer reviewers, and referenced scientists, that work is still before us.

In "Developing an e-rater Advisory to Detect Babel-generated Essays," Aoife Cahill, Martin Chodorow, and Michael Flor focus on BABEL Generator (2014), a web-based tool which automatically generates semantically incohesive essays. Because such systems generate writing samples intended to fool automated scoring engines (often associated with automated writing evaluation [AWE]), they pose a threat to information interpretation and use. Using a corpus of Babel-generated essays and a corresponding corpus of good-faith essays, Cahill and her colleagues built a classifier to distinguish between the two and integrated it into an existing AWE system. The researchers found that the classifier built on Babel-generated essays and good-faith essays can distinguish the Babel-generated essays from the good-faith ones with 100% accuracy. When the classifier was integrated into the automated scoring engine (Educational Testing Service's e-rater®), it flagged very few responses that were submitted as part of operational submissions (76 of 434,656)—responses that, in fact, were previously flagged as either null (non-scorable) or 1 (lowest score) by human experts. This new measure of lexical-semantic cohesion is an important step in our ability to provide evidence of fairness in AWE.

2.0 Research Notes

In our two research note studies, attention is given to the use of corpus analysis in improving writing center consultation and the value of measuring motivation in writing instruction.

Writing centers—a key contact zone (Pratt, 1991) in writing program administration—function like large data repositories. Use of that information to support writing program administration is examined at four institutions in "It's All in the Notes: What Session Notes Can Tell Us About the Work of Writing Centers." Genie N. Giaimo (The Ohio State University), Joseph J. Cheatle (Michigan State University), Candace K. Hastings (Texas A&M University), and Christine Modey (University of Michigan) analyze over 44,000 writing center session notes (transactions completed with a tutor after instructional transaction with a student) using Voyant, a web-based application for performing text analysis. Appropriately, the analysis takes different forms at different campuses. At The Ohio State University, Giaimo's analysis provides insight about the ways in which writing center consultants consciously articulate their use of specific strategies learned in their training. At Michigan State University, Cheatle investigates the ways in which a writing center philosophy focuses on global composing issues as embodied in session notes. At Texas A&M University, Hastings focuses on single terms, such as *flow*, to demonstrate semiotic shifts in use and meaning. And, at the University of Michigan, Modey also focuses on specific word choices to propose a longitudinal approach to writing consultation in which attention to key terms yields a programmatic approach to feedback effectiveness.

In "Going Beyond 'That was fun': Measuring Writing Motivation," Tamara Powell Tate and Mark Warschauer focus on motivation, a key malleable factor involved in improving student writing. With special emphasis on students in grades 4 to 12, the researchers present the results

of a literature search undertaken to identify measures (in this case, instruments often consisting of scales) of writing motivation used within the last 10 years to evaluate school-age students. While discrete measures of motivation were identified, attendant properties for the instruments (in terms of fairness, validity, and reliability) were often absent. Additionally, important theories of motivation were not captured by the identified measures. This review provides researchers with the current landscape of instruments capturing writing motivation—and reminds us that additional, comprehensively validated measures are needed to capture this facet of the intrapersonal domain of writing. For researchers in writing analytics, identification of measures related to motivation—indeed, to all facets of the intrapersonal domain of writing—is significant in expanding the field. Such expansion is taken up by Burstein and Beigman Klebanov and her colleagues in the special section to which we now turn.

3.0 Special Section

As a new journal offered annually in a web-based, open application format, *Analytics* is in a unique position to advance research. As editors, we are alert to conferences featuring sessions on writing analytics so that state-of-the-art work can be published within a very short time of its presentation. Such is the case in the special section of Volume 2 featuring research presented on April 14, 2018, at a Coordinated Symposium hosted by the National Council on Measurement in Education (NCME) in New York City. Entitled *What Writing Analytics Can Tell Us About Broader Success Outcomes*, the session was chaired by Jill Burstein and Daniel McCaffrey, with Mya Poe as discussant. While two of the papers resulting from the presentations were promised to others, we were fortunate to secure an introduction, two papers, and a response to the session. Reviewed by NCME and Educational Testing Service peers and our own editorial team, the papers from this session provide a useful way to record state-of-the-art directions in writing analytics.

In her introduction to the Coordinated Symposium, Jill Burstein calls attention to the important connection between computational methods for the study of texts (the technical side of writing analytics) and ways that these methods help us to understand naturally occurring task genres and composing processes (the interpretative side of writing analytics). Alignment of the technical and the interpretative—or, more accurately, ensuring connections between both—is featured in her introductory remarks to the NCME panel. As a specialist in AWE, Burstein reminds us that these systems do indeed yield scores. However, as she notes, AWE can also be used to study the relationship between socio-cognitive writing achievement frameworks (those based on a model of writing skill and knowledge, an intrapersonal domain, and an interpersonal domain) and broader success outcomes (such as post-secondary retention and completion).

Following her introduction, Burstein and her colleagues present an application of AWE use in "Writing Mentor™: Writing Progress Using Self-Regulated Writing Support." As a Google Docs add-on, the Writing Mentor™ (WM) application is designed to help students improve their writing and thus promote their success in postsecondary settings. Specifically, WM provides AWE feedback using NLP methods and related linguistic resources in four areas: use of sources, claims, and evidence; topic development; coherence; and knowledge of conventions. The case study includes descriptive evaluations from an Amazon Mechanical Turk (AMT) usability task

situated in WM (n = 108) and from users-in-the-wild data ($n \approx 2{,}693$). From the AMT study, Burstein and colleagues report self-efficacy scores (addressing the intrapersonal domain of writing), as well as survey perceptions of the tool and open-ended user responses. From the users-in-the-wild study, a snapshot of information is presented from the event logs analyzed approximately eight months after the release of the application in late November 2017. Among the findings from this study, especially interesting is that writers are making feedback-related changes to texts as they are revising; that is, the feedback provided by WM is being used, in real time, as writers shift words among text positions to enable readers to better follow the flow of ideas in a given piece of writing. The study concludes with contextualization of WM as a form of personalized learning designed to provide systematic instruction according to specified domain models of constructs. In its ability to provide an integrated system by which self-assessment and performance information relevant to writing construct may be obtained, WM extends our understanding of AWE.

In "Utility-Value Score: A Case Study in System Generalization for Writing Analytics," Beigman Klebanov and her co-authors extend AWE research to examine utility value: the perceived value of coursework and individual engagement in it (Harackiewicz et al., 2015). While utility-value intervention (UVIs)—aimed at promoting student motivation and performance by having students consciously reflect on the value of what they are learning—have proven successful in improving grades and narrowing achievement gaps, human scoring is labor intensive. In this context, AWE is especially promising in its ability to identify, in student UV written statements, connections between science content and the student's own life by focusing on past test verbs (an indicator of narrativity) and related features. Using undergraduate homework assignments in introductory science and psychology courses that required UV responses (e.g., explaining why specific information is relevant to the writer's life), the researchers found that the features identified do, in fact, capture linguistic regularities associated with written UV expression. Nevertheless, caution is exercised in the recognition that shifts to a new student population, a new subject matter course, or a new variant of the original task may result in systematic changes in the textual features that render the original system inapplicable to the new context. Some features generalize better than others, and so challenges remain. Variability associated with site-specific use is a key finding of the study. AWE systems, it appears, are far more nuanced than many critics imagined, and in this benefit, we also find a limit to generalization.

In her response to the panel, Mya Poe observes that we are presently in a phase of second-generation AWE in which research may be characterized by aspirational pedagogical attention to varied forms of evidence related to construct validity. Practically, this shift means that linguistic characteristics identified in the AWE must align within a larger universe of interpretation if the systems are to be deeply helpful to student writing development. As she notes, second-generation AWE must be developed with a clear understanding of the contexts in which it is to be used. In turning to the theory-building that informs AWE development, Poe notes that both Burstein and Beigman Klebanov use linguistic models of text and syntactic features as markers of writing development. She contrasts this perspective with those drawn from writing studies that focus on textual features and genre manifestation as they represent (and are mediated by) rhetorical and

social aims of writing. As Poe proposes, one way forward for multidisciplinary research between educational measurement and writing scholars at the present moment is to focus on precisely what is meant by a socio-cognitive, situated view of language. In her emphasis on the shift in AWE from summative (score-based) to formative (feedback-centered) assessment, Poe also implicitly re-centers the aim of assessment to learning. Often termed Assessment for Learning (A*f*L), this re-centering calls attention to the important *as*, of, and *for* formulation: assessment *as* a form of learning, assessment *of* student learning, and assessment *for* learning how to improve stakeholder opportunities (Heritage & Wiley, 2018). Aiming toward complexity, she appropriately notes, is something that well-trained writing teachers value. Provocatively, Poe ends her response by encouraging both the educational measurement community and the writing studies community to advance evidence standards for writing assessment in general—and for automated models in particular—that make fairness a central consideration.

4.0 Reflection

In Volume 1, the editors provided a taxonomy of four interrelated programs of research in writing analytics: educational measurement, massive data analysis, digital learning ecologies, and ethical philosophy (Figure 1, p. x). In Volume 2, we see that this taxonomy remains useful as a way to understand how the field is being deepened, and thereby strengthened, in significant ways.

Methodologically, we are beginning to see increased attention to generalization inferences—those comparative claims made across a range of assessment conditions. As Kane (2016) writes in his example of such inferences, "[W]e generalize over test forms, occasions, test administrations and contexts, and we assume that the score would not vary much had these aspects of the testing been somewhat different" (p. 72). To lend precision to our generalizations, we identify different conditions of observation such as tasks and contexts. We specify our sampling plans, estimate sampling errors, and provide standard errors. To justify the generalizability of these facets, we provide our validity augment. As Kane observes, generalizability assumptions "should specify how widely the interpretation is to be generalized, and to the extent that generalizability over any facet is doubtful, it would need to be investigated empirically" (p. 73). Of course, as Kane notes, absence of generalizability does not mean that the assessment lacks validity; that is, if the interpretation does not include generalizability inferences, the assessment may remain valid for the specific site where the assessment was conducted.

Emphasis on generalizability inferences provides a useful way to understand methodological developments in writing analytics. Because of large sample sizes, substantial expenditure of fiscal and human resources, and the need for studies to be useful, writing analytics is desirous of generalization inferences. This desire is especially evident in Volume 2 in the research of Arthurs and Ringler et al. As Arthurs notes, the SSW had local goals, such as the ability to provide an overview of student writing at Stanford. Broader goals, such as the ability of the study to contribute to longitudinal development of undergraduate writers, were also part of the study. Following the interpretation framework provided by Kane, institutions wishing to compare the results presented by Arthurs would begin by studying the institutional profile of Stanford listed

in the federal Integrated Postsecondary Education Data System. If the student profiles were comparative, then the sampling of the SSW, in general, and the one used by Arthurs, in particular, would be examined for their representativeness of the undergraduate student population at Stanford. If both comparisons held, then the institutional comparative process could proceed. Ringler et al. provide a different approach. Using the ARGREF, an argument reference corpus, the researchers are able to provide an explicit comparative basis for the writing under examination. Out of the 16 use categories that constitute their stated rhetorical profile of an essay, the researchers compare the targeted writing samples to the ARGREF corpus to determine statistically significantly lower or higher categories of rhetorical use by students.

Both examples illustrate the desire for generalization inferences—in this case, the ability to generalize findings from a sample to a target population. In the case of the Arthurs study, comparison is dependent on institutional profile similarity. In the case of the Ringler et al. study, comparison is dependent on a comparative corpus. There are costs and benefits to both approaches, and depending on the level of specificity determined as prerequisite to comparison, neither approach may be sufficient. Dissimilar institutions would yield little useful information. If, for instance, distinct genre comparisons are being made, only institutions and corpora using similar pedagogical approaches would be candidates for comparison; that is, a source-based business proposal would have to be compared with a source-based business proposal. On the other hand, the benefits of appropriate comparative studies are equally related to specificity. If institutions have similar students and similar writing pedagogies, then comparison provides a useful point for targeting curricular development. The benefits of comparative corpora are equally clear in allowing the diagnosis of differences and similarities that can also be used in writing program design. Yet, even here, Beigman Klebanov et al. are explicit in urging caution when applying automated methods developed on smaller-scale of substantially homogeneous data to larger, heterogeneous datasets.

Conceptually, in Volume 2 we are beginning to see increased attention to fairness inferences—those assurances of protection anonymity and qualifications made to the generalization claims. We have classified above the study by Rudniy as one of confidentiality (protection of human subjects under 45 CFR 46 [Federal Policy for the Protection of Human Subjects, 2018]) and the study by Cahill et al. as one of construct-irrelevant response (required minimization under Standard 3.1 [American Educational Research Association et al., 2014]). Legal and standards-based guidelines are important ways of ensuring fairness, and these studies illustrate how they may be adapted for writing analytics research.

A close reading of Volume 2 also reveals concern for fairness in terms of sub-group analysis and construct representation. Reese et al. are explicit in valuing sub-group analysis: Studying fairness in peer review includes investigating whether any given demographic group disproportionately benefits or loses from the practice. Their study of attitudes toward self-reflection and transaction, elements of the intrapersonal and interpersonal domains of writing, also reveals attention to broad construct representation. While attention to sub-group analysis is a traditional prerequisite to score interpretation and use characterized as evidence of fairness, recent calls for explicit domain models link validity concerns to fairness. Because diminished representation of the writing construct is associated with forms of disparate impact (Kelly-Riley

& Whithaus, 2016), recent work in conceptualizing fairness has called attention to the need for broad construct representation as a way to increase equity (Poe, Inoue, & Elliot, 2018). In the studies by Tate and Warschauer, Burstein et al., and Beigman Klebanov et al., we therefore see the presence of a defined construct model, categorized according to domains, associated with evidence of fairness.

In Volume 2, we are therefore witnessing ways that researchers in writing analytics are using, adapting, and creating methodological and conceptual frameworks. Significant here is that calls for caution associated with massive (big) data are being heeded. Mangrum (2018) has provided a historical account of the value dualism between those who advocate aggregate-scale analysis and those who advocate close readings. While interpretative disjuncture no doubt exists, it is equally true that resonances are also present. Readers will note in Volume 2 that there is no advocacy of score use. Instead, each study is devoted to using information to help students improve their writing ability. There is none of the leveling of the objects of inquiry about which Mangrum expresses reservations. The objects of inquiry—information derived from corpus analysis—are never leveled, and caution about such leveling appears to be one hallmark of writing analytics. Attention to the contact zone of information interpretation and use is a hallmark of writing analytics. While fields such as writing studies are said to have experienced a linguistic turn (Zalewski, 2013), writing analytics seems to have been born linguistic. Hence, we note a second hallmark: The situated nature of language is uniformly acknowledged by researchers.

As we prepare Volume 2 for release, we are also preparing for The 7th International Conference on Writing Analytics. The conference will explore innovation in writing analytics, with special attention to broadening our annual conference community. The call for proposals is an extension of research we hope to see in future issues of the journal:

- Who are the stakeholders of writing analytics? Traditionally, we think of key educational stakeholders as the following: advisory boards, administration, faculty, parents, professional organizations, students, and the public. These groups, however, may need to be expanded when we think of information use and research impact.
- How may the stakeholder community of writing analytics be better understood? We must know more about the kinds of disciplinary knowledge and multidisciplinary collaborations that are needed to expand the body of knowledge associated with our community.
- How might writing analytics be more widely communicated to stakeholders? Conference presentations, peer-reviewed articles, and book chapters are common ways to distribute research findings. Yet these vehicles are limited to academic stakeholders. Needed is a new communication taxonomy for writing analytics.
- How may we better understand the consequences of our research for student learning? Many web-based platforms are now using archival and real-time analytics for feedback. Little is known however, regarding the impact of these platforms and the information they are capable of providing regarding diverse student groups.

We hope that our journal continues to deepen our understanding of written communication as we broaden our stakeholders.

5.0 Acknowledgements

Volume 2 would not be possible without the active support of colleagues who have agreed to serve on the Board of Reviewers, presented in Table 1 with the area of manuscripts each specialist reviewed. Our reviewers are, in fact, developmental editors. Because the journal uses a policy of desk rejection of manuscripts not fully developed or not directly related to the mission of the journal, reviewers know that the editors are committed to publishing the studies under review. As manuscripts are reviewed twice—and often three times by Board members—detailed advice is provided on how to strengthen work that is already well considered. The high quality of research presented in Volume 2 is due to their commitment to the journal.

Table 1

Journal of Writing Analytics Board of Reviewers and Review Specialization, Volume 2, 2018

Reviewer	Affiliation	Review Specialization
Chris M. Anson	North Carolina State University	Rhetoric and Composition
Ian G. Anson	University of Maryland, Baltimore County	Electronic Forecasting
Laura Aull	Wake Forest University	Corpus Linguistics
Ryan Baker	University of Pennsylvania	Learning Analytics
Duncan Buell	University of South Carolina	Computer Science
Hugh Burns	Texas Woman's University and US Air Force Academy	Computational Rhetoric
Scott Crossley	Georgia State University	Applied Linguistics
Irvin R. Katz	Educational Testing Service	Cognitive Psychology
David Kaufer	Carnegie Mellon University	Digital Textual Analysis
Andrew Klobucar	New Jersey Institute of Technology	Digital Humanities
Suzanne Lane	Massachusetts Institute of Technology	Writing in the Disciplines
Djuddah A.J. Leijen	University of Tartu, Estonia	English Language Learning
Collin F. Lynch	North Carolina State University	Intelligent Tutoring Systems
Mya Poe	Northeastern University	Writing Assessment
Valerie Ross	University of Pennsylvania	Critical Writing
Alex Rudniy	University of Scranton	Educational Data Mining
David Slomp	University of Lethbridge	Qualitative Research
Erica Snow	Imbellus	Artificial Intelligence
Swapna Somasundaran	Educational Testing Service	Sentiment and Discourse Analysis
Jennifer Pei-Ling Tan	National Institute of Education, Singapore	Creativity & 21st Century Competencies

In the process for this year, we want to give special thanks to Hugh Burns and Erica Snow for their additional reviews this year and the special care they took with each. We also want to thank Laura Runge, Chair of the Department of English at the University of South Florida, for

her support. We also wish to thank Mike Palmquist, founding director of the Colorado State University Open Press, for his continued support as our publisher.

6.0. Award

Please join us in congratulating David Eubanks, Assistant Vice President in the Office of Institutional Assessment and Research at Furman University and Executive Editor at *Analytics*, for winning the 2016 Charles F. Elton Best Paper Award. Presented by the Association for Institutional Research (AIR), the Charles F. Elton Best Paper Award celebrates scholarly papers presented at the AIR annual conference that exemplify the standards of excellence established by the award's namesake and that make significant contributions to the field of IR. The purpose of the award is to promote scholarship and to acknowledge that AIR members make a wide variety of scholarly contributions to the field, ranging from theory to practice. Offering a new quantitative technique for visualizing and assessing inter-rater agreement in discrete ordinal or categorical data such as rubric ratings, Eubanks' AIR paper was published the following year in Volume 1 of *Analytics* as "(Re) Visualizing Rater Agreement: Beyond Single-Parameter Measures" (pp. 276-310).

7.0 Design

With Volume 2, we say farewell to Sophie Elliot, our first designer, who has now left to join the design team of Target as a Visual Merchandiser. Ms. Elliot designed the cover for Volume 1, as well as our logo and publication template. We are in her debt and wish her well in her design career.

We now welcome our new designer, Matthew J. Osborn, Assistant Director of The Critical Writing Program at University of Pennsylvania. Inspired by research interests in the relations among aesthetics, rhetorics, and information, his design for Volume 2's cover juxtaposes the emergent complexity of analytics with flat and minimal sensibilities in vogue at the time of this writing. This dialectic reflects a range of relationships with information, scale, and insight to which the present volume responds. In addition to his own site at matthewjosborn.com, he maintains design for the Philadelphia Writing Program Administrators at phillywpa.com.

8.0 Advertisement

In this issue, we begin a collaborative advertisement by welcoming *The Journal of Teaching Writing* (*JTW*). Now in its 36th year, *JTW* is devoted to the teaching of writing at all academic levels and in any subject area. The journal holds the mission to publish refereed articles and reviews that address the practices and theories that bear on our knowledge of how people learn and communicate through writing. We are delighted to join *JTW* in promoting journals in the field of rhetoric and composition.

Analytics is part of a group of Writing Across the Curriculum and Writing Journals. For more on these journals, visit the WAC Clearinghouse Journal Listings:

https://wac.colostate.edu/resources/wac/journals. These listings include journals that are available online and/or in print. Whenever possible, a link has been provided to a website associated with a journal.

References

American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Aull, L. L. (2017). Corpus analysis of argumentative versus explanatory discourse and its implications for writing task design. *Journal of Writing Analytics*, *1*(1), 1–47. Retrieved from https://journals.colostate.edu/analytics/article/view/106

BABEL Generator (2014). Retrieved from http://babel-generator.herokuapp.com

Federal Policy for the Protection of Human Subjects (2018, July 19). Retrieved from https://www.ecfr.gov/cgi-bin/retrieveECFR?gp=&SID=83cd09e1c0f5c6937cd9d7513160fc3f&pitd=20180719&n=pt45.1.46&r=PART&ty=HTML

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2015). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *Journal of Personality and Social Psychology*, *111*, 745–765.

Heritage, M., & Wiley, C. (2018). Reaping benefits of assessment for learning: Achievement, identity, and equity. *ZDM: Mathematics Education*, *50*, 729–741.

Kane. M. (2016). Validation strategies: Delineating and validating proposed interpretation and use of test scores. In S Lane, M. R. Raymond, and T. M Haladyuna (Eds.), *Handbook of test development* (2nd ed., pp. 64–80). New York: Routledge.

Kelly Riley, D., & Whithaus, C. (2016). A theory of ethics for writing assessment. [Special issue]. *Journal of Writing Assessment*, *9*(1). Retrieved from http://journalofwritingassessment.org/article.php?article=99

Lancaster, Z. (2016). Do academics really write this way? A corpus investigation of moves and templates in *They Say/I Say. College Composition and Communication*, *67*, 437–464.

Mangrum, B. (2018). Aggregation, public criticism, and the history of big data. *Publications of the Modern Language Association*, *133*, 1207–1224.

Poe, M., Inoue, A. B., & Elliot, N. (2018). *Writing assessment, social justice, and the advancement of opportunity*. Fort Collins, Colorado: The WAC Clearinghouse and University Press of Colorado.

Pratt, M. L. (1991). Arts of the contact zone. *Profession*, 33–40.

SSW. (2018, 30 April). Retrieved from Stanford Study of Writing: https://ssw.stanford.edu/

Zalewski, J. (2013). The post-process era in composition studies and the linguistic turn of the 20th century. In D. Gabrys´-Barker, E. Piechurska-Kuciel, and J. Zybert (Eds.), *Investigations in teaching and learning languages: Studies in honour of Hanna Komorowska* (pp. 51–61). Switzerland: Springer.