

Book Review of Rachael E. Gabriel & Richard L. Allington. (2016). *Evaluating Literacy Instruction: Principles and Promising Practices*. New York, NY: Routledge

Shawnee Hardware, York University, Canada

Introduction

Evaluating Literacy Instruction: Principles and Promising Practices, edited by Rachael E. Gabriel and Richard L. Allington, offers great insight for English Language Arts/Literacy (ELA) educational researchers, ELA teachers, and educational administrators. The book emphasizes the need for both valid and reliable ELA observational evaluation that captures students' differentiated literacy needs while still meeting the stipulations of the Common Core States Standards (CCSS)¹, the Race to the Top² grant, and the Elementary and Secondary Education Act³. The book's contributors noted that through the CCSS, the government of the United States (U.S.) had legislated a series of standardized ELA evaluations at the expense of more culturally responsive and differentiated ELA evaluation. Continuing with this line of thinking, Gabriel et al. alert us to the established practice of using the scores obtained on these ELA observational evaluations to determine teachers' performance. Speaking specifically to this, throughout the book, its contributors caution against the increasingly popular habit of using ELA evaluation scores as a correlation of teachers' performance by noting that standardized evaluations alone "cannot be used to support better instruction" (Gabriel & Wouflin, Chapter 2, p. 17) because "they have not necessarily considered the particular objective data about teacher performance in different settings" (Lopez, Proctor & Scanlan, Chapter 9, p. 146).

The practice of using classroom evaluations to draw a conclusion on teachers' performance is classified as Value Added Methods (VAMs). *Evaluating Literacy Instruction* shares Darling-Hammond's (2015) concerns about the use of VAMs that estimate teachers' contribution to students' classroom learning. For example, in Chapter 1, Gabriel and Allington (2016) warn against the use of most commercially available instruments for ELA teachers' evaluation, as while they attempt "to be utilitarian in most grade level, content area and curricular connect, they ignore the facts that matter most for literacy development" (p. 7). Some of these "facts that matter" are the students' own reading background, the choice of texts used in the classroom, the resources available to the teacher, and class size. In acknowledgement of these "facts," Gabriel and Allington recommend that ELA teachers' evaluations capture good teaching, students' meaning-making, students' English language development as well as promote culturally and linguistically responsive learning. It is worth noting that in today's neoliberal educational environment, these facts that matter have been largely neglected from all the core subject areas and VAMs have been more widely used to assess teachers' performance in these subjects. Unfortunately, two of the consequences that resulted in the increase of VAMs are the proletarianization of teachers' tasks and "deprofessionalization" of teachers' roles (Apple, 1986; Buyuk, 2014), as these teachers' contributions in the classroom are reduced to a single score rather than an assessment which captures different variables in the teaching and learning process.

The strength of *Evaluating Literacy Instruction* lies in the strong body of scholars (Brady, Heiser, McCormick & Forgan, 2016; Grossman, Cohen, Ronfeldt & Brown; Rothstein, 2016) who substantiate the main themes espoused in the book: a critique of VAMs as the sole evaluative tool for ELA teachers. For example, Grossman and colleagues (2016) recognize that standardized evaluations can be used to initiate the conversation about what students are learning in the classroom; however, like Gabriel and Allington (2016), they advise that this should not be used to paint the whole picture of the teaching and learning process in any classroom. This recommendation is a

¹ Started in 2008, the CCSS is a set of academic standards that guide mathematics and ELA instructions. The CCSS guides what students should know and learn per grade (Common Core State Standards Initiative, 2019).

² Race to the Top is a competitive grant that rewards states which have made significant gains in student achievement and that have a comprehensive plan to sustain those gains in the future (US Department of Education, 2009).

³ The Elementary and Secondary Act replaced the No Child Left Behind policy and attempted to encourage educational equity by offering grants to districts serving predominately low-income students and special education centers (US Department of Education, 2017).

timely, given the growing popularity in the use of VAMs in the United States (Campbell & Ronfeldt, 2018) and the increasing calls for teacher accountability in Canada (Crundwell, 2005; Clifton, 2013). Hence, the main arguments advanced in *Evaluating Literacy Instruction* buttress the academic rigidity of educational researchers and practitioners who advocate for more nuanced ways of evaluating teachers' performance.

Evaluating Literacy Instruction contributors' expertise in reading instruction, ELA evaluation, and education policy research was evident throughout the book's 11 chapters. To increase readability, the editors, Gabriel and Allington, divided the book into two sections. Section one, titled "Crafting Systems and Policies for Evaluating Literacy Instruction" comprises chapters two to six. These five chapters discuss the construction and components of popular observation schemes such as the Danielson's Framework for Teaching (FFT)⁴ and The Classroom Assessment Scoring System (CLASS)⁵ to enlighten education practitioners about the most useful aspects of these observational tools. Section two, named "Examples of Alternative Systems," comprises chapters seven to 11, which address the paucity in the literature about valid and reliable substitutes for VAMs and the nexus between reading instruction and teacher evaluation. This section provides useful considerations about the various ways to improve the utility and validity of any observational schemes. For example, in Chapter 7, Peterson explains the School for Change Project, which schools can follow to create their reading programs, curricula, materials, and models. Moreover, Hoffman and Sailor's Chapter 8 outlines TEX-IN3⁶ that allows for teachers to capture both the students' individualized and collective emic perspectives on reading and their classroom books. Each chapter is easy to understand as they follow a straightforward format: the problem, a literature review of present research in the area, an alternative way of addressing the issue and a "Take Action" section which presents a recommendation. Although the chapters of the book are short with 15 pages on average, the Take Action section provides a chapter summary; hence, it is useful for readers who have limited time to complete an entire chapter.

Apart from offering concise and practical recommendations, *Evaluating Literacy Instruction* is novel as it attempts to fill the gap in the literature on how to link reading instruction research with teacher evaluation policy. From the onset, the book's editors explained that there appears to be a scarcity in the literature on how to evaluate ELA teachers and their impact in the classroom. Again, given the rising dominance of VAMs and standardized ELA evaluation tools, the book has utility as a starting point in helping us to critically analyze the "value" that these forms of evaluation add to the classroom. I concur with the authors that there seems to be a paucity of qualitative evaluation tools that can capture rich descriptions of not only the teaching and learning process but also students' microgenetic development in the classroom. Qualitative evaluation tools (e.g., see Estacion, McMahon, Quint, Melamud & Stephens, 2004) represent a strength-based depiction of the students and teachers' development, highlighting how they evolve as learners (including how they navigate their various identities) and demonstrating the various ways they are engaged in the class. I posit that the use of more qualitative teacher assessment tools will improve the nexus between reading instruction and teacher evaluation policy as these tools can provide essential information on how the ELA teachers' pedagogy contributes to students' learning and development.

Despite the strengths of this book, I have identified three structural issues with it. The first issue is the widespread use of quantitative data. A book on ELA education and evaluation should include more narratives of teachers' perceptions of the strategies they employ in the classroom and their analysis of how this impacts their students. English Language Arts is a mainly qualitative subject; therefore, it stands to reason that qualitative evaluative tools be used in ELA teachers' assessment. For example, in Chapter 6, instead of analyzing samples of the students' work and interviewing the teacher for their chapter on the benefits of an observational system in a first grade writing class, Coker and colleagues used quantitative tools (table and graphs) to convey the observational data. Readers of the book would benefit from hearing the ELA teachers' accounts of the effectiveness of these observational systems, as these narratives could inform administrators about the advantages and disadvantages of using this observation system, hence, offering a more insightful review of the observational system.

⁴ This is a research based standardized protocol which means four learning components. These components are: 1. Planning and preparations, 2. classroom environment, 3. instruction, 4. teaching responsibility (Danielson Group, 2017).

⁵ CLASS is a multi-grade, general classroom observational tool which captures social and emotional assistance, organizational and management supports and instructional scaffolding in the classroom (Stuhman, Hamre, Downer & Pianta, n.d.).

⁶ TEX-IN3 contains three components and analyses the text environment, text use and text understanding in an English language classroom (Center on Great Teachers & Leaders, 2013).

Another issue that I found with the book is that it is very text heavy. Although the authors used infographics throughout (see pages 22, 23, 84, 90, 107-112, and 154), they could have included more info-graphics, potentially making it more accessible to a broad readership. Skillfully blended graphics and images would improve the books' comprehensibility for all readers, regardless their understanding of ELA evaluation language. Moreover, making the book more multimodal would only serve to increase different readers' understanding, as the information therein would be conveyed through multiple forms.

Finally, I recommend that the authors include a glossary in the next edition of the book. A glossary would make the book easier to understand for readers who are unfamiliar with but interested in U.S. English Language Arts/Literacy evaluations. Some chapters in the book are laden with pertinent but technical ELA evaluation jargon and from research in the field. In portraying this technical information, the book is laced with many acronyms. Therefore, a glossary would make *Evaluating Literacy Instruction* easier to understand and perhaps more widely read.

In closing, despite the cited limitations, *Evaluating Literacy Instruction* is a valuable read for education practitioners and researchers who want a more nuanced understanding of U.S. English Language Arts/Literacy evaluations. It reminds the aforementioned educational stakeholders that a single research tool, used at a specific instance in the classroom, is not sufficient to capture what teachers do in their classrooms and how this impacts students' learning. This reminder echoes Larsen's (2005) position that although standardized assessments "may provide a framework to start thinking about quality teaching, it is important that mechanism is not put in place (e.g., checklist and short-answers tests) that ignore the complexities and highly contextualized nature of teaching" (p. 302). *Evaluating Literacy Instruction* reinforces Larsen's (2005) message that valid and reliable evaluations should be conducted at multiple instances in the classroom using various data collection methods to capture both the in-school complexities and students' lived experiences that impact their learning.

REFERENCES

- Apple, M. (1986). *Teachers and texts: A political economy of class and gender*. London, England: Routledge and Kegan Paul.
- Brady, M. P., Heister, L. A., McCormick, J. K., & Forgan, J. (2016). Value added models for teacher preparation programs: Validity and reliability threats, and a manageable alternative. *The Educational Forum*, 80(3), 339-352.
- Buyurk, H. (2014). Professionalization or proletarianisation: Which concept defines the changes in teachers' work? *Social and Behavioural Sciences*, 116, 1709-1714.
- Campbell, S. L., & Ronfeldt, M. (2018). Observational evaluation of teachers: Measuring more than we bargained for? *American Educational Research Journal*, 55(6), 1233-1267.
- Center on Great Teachers and Leaders. (2013). *Guide to evaluation products- TEX-IN3 Observation System*. American Institutes for Research. Retrieved January 09, 2019 from <http://resource.tqsource.org/gep/GEPTool.aspx?gid=33>
- Clifton, R. A. (2013). *Obtaining better teachers for Canadian public schools: A review of the "Teacher Effectiveness" research literature*. Retrieved January 10, 2019 from <https://www.fraserinstitute.org/sites/default/files/obtaining-better-teachers-for-canadian-public-schools.pdf>
- Common Core: State Standards Initiative. (2019). Retrieved January 09, 2019 from <http://www.corestandards.org/other-resources/>
- Crundwell, R.M. (2005). Alternative strategies for large scale student assessment in Canada: Is value-added assessment one possible answer. *Canadian Journal of Educational Administration and Policy*, 41, 1-21.
- Danielson Group. (2017). The Framework. Retrieved January 11, 2019 from <https://www.danielsongroup.org/framework/>
- Darling-Hammond, L. (2015). Can value added add value to teacher evaluation? *Educational Researcher*, 44(2), 132-137.
- Estacion, A., McMahan, T., Quint, J., Melamud, B., & Stephens, L. (2004). *Conducting classroom observation in first things first schools*. Retrieved from https://www.mdrc.org/sites/default/files/full_64.pdf
- Grossman, P., Cohen, J., Ronfeldt, M., & Brown, L. (2016). The test matters: The relationship between classroom observation scores and teacher value added on multiple types of assessment. *Educational Researcher*, 20(10), 1-11.
- Larsen, M. A. (2005). A critical analysis of teacher evaluation policy trends. *Australian Journal of Education*, 49(3), 292-305.
- Rothstein, J. (2016). *Can value-added models identify teachers' impacts?* Retrieved January 09, 2019 from <http://irle.berkeley.edu/can-value-added-models-identify-teachers-impacts/>
- Stuhlman, M. W., Hamre, B. K., Downer, J. T., & Pianta, R. C. (n.d.). What should classroom observation measure: Part 2. Retrieved January 10, 2019 from https://curry.virginia.edu/uploads/resourceLibrary/CASTL_practioner_Part2_single.pdf
- The US Department of Education. (2009). *Race to the Top Fund*. Retrieved January 26, 2017 from <https://www2.ed.gov/programs/racetothetop/index.html>
- The US Department of Education. (2017). Every student succeeds act (ESSA). Retrieved January 26, 2017 from <https://www.ed.gov/essa?src=rn>
- Willms, J. D. (2009). *Value-added models in education*. Retrieved January 08, 2019 from https://www.unb.ca/research/institutes/crisp/resources/pdf/policybriefs/pb15_value-added_models_in_education.pdf

ABOUT THE AUTHOR

Shawnee Hardware, PhD is a Researcher with the Child Development Institute, Toronto where she is developing resources for the Stop Now And Plan program www.stopnowandplan.ca. Shawnee holds a PhD in Language, Culture and Teaching from York University, Toronto and a M.Ed. in Curriculum, Teaching and Learning Studies from Memorial University, Newfoundland. Her interests include student engagement, youth civic engagement, English language teaching and learning, and sociopsycholinguistics.