



Review of *Blended language program evaluation*

Senta Goertler, Michigan State University

Blended language program evaluation

Gruba, P., Cárdenas-Claros, M., Suvorov, R., & Rick, K.

2016

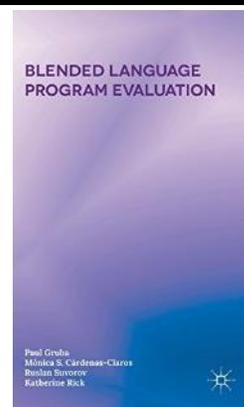
ISBN: 978-1-137-51436-3

US \$100.00

248 pp.

Palgrave MacMillan

Basingstoke, UK



As computer-assisted language learning (CALL) implementations and blended instruction become more prominent, it is—as many CALL researchers have argued—crucial that such pedagogical innovations be founded in second language acquisition (SLA) theory, building on SLA and CALL SLA research results, and evaluated using best practices and newest methodologies (see Chapelle, 2009). *Blended language program evaluation* by Gruba, Cárdenas-Claros, Suvorov, and Rick is an excellent example of combining program evaluation research and theory from a variety of disciplines with CALL, specifically blended language instruction.

Chapelle (2009) has outlined how different theoretical approaches can be implemented in CALL curricula. Furthermore, I have argued that blended and online education can be justified based on SLA research and theory (Goertler, 2011). Even though blended language learning already has a long tradition in comparison to newer formats such as the flipped classroom (see Kim, Park, Jang, & Nam, 2017), many blended program developments are still driven by logistical, and not pedagogical, considerations and top-down rather than bottom-up processes (Goertler, 2011). Early reports on blended learning mostly summarize the pedagogical building blocks of curricula and report on enrollment trends, logistical benefits, or student satisfaction. The second wave of blended research focuses on quasi-experimental comparison studies comparing blended delivery formats with face-to-face or solely online formats (for summary see Grgurović, 2011). As I and others have argued, quasi-experimental studies may be insufficient for evaluating the learning outcomes of blended learning, since the approach is built on the assumption that learning in a blended environment has to lead to the same learning outcomes as a face-to-face environment. This assumption closes off the possibility that blended delivery formats may lead to unique learning outcomes that go beyond the outcomes typically seen in face-to-face formats (e.g., improved digital literacy). It also assumes that pure face-to-face delivery formats will always be an alternative option, which, in reality, might not be true. To fully capture the challenges and benefits of blended learning, an evaluation framework is necessary that involves more stakeholders, data points, and analyses and that can accommodate greater complexity than simple comparison of test results and perception survey reports between face-to-face and blended course participants. To impact language teaching practice positively, such research needs to be conducted ethically and with maximal buy-in from a range of stakeholders.

Gruba et al. present a comprehensive and contemporary evaluation framework expanding Kane's (2006)

argument-based approach. Their proposed evaluation approach overcomes several shortcomings of previous research on blended learning and is a good guide for language program evaluation in general—and blended language programs in particular. The book is divided into three parts: (1) an overview of previous research and theories, (2) an introduction of the three different levels of program evaluation (i.e., micro, meso, and macro), and (3) sample case studies for the different levels across four contexts (i.e., Chile, USA, Vietnam, and Australia). Each part includes several dense chapters with a plethora of invaluable summarizing tables and graphs. The book is framed by a foreword by Chapelle and a concluding reflective chapter.

In Chapter 1, the author team introduces blended learning from a program evaluation perspective. This chapter builds on Gruba's previous co-authored book on blended learning (Gruba & Hinkelman, 2012). This chapter poignantly summarizes the shortcomings and challenges of language program evaluation, especially blended language learning. First, they point to the common issue of technology not being fully integrated, which leads to technological concerns being put above pedagogical and learner needs. Secondly, they correctly point to issues in teacher training, attitudes, and skills for CALL integration (see also Arnold & Ducate, 2015). Next, they discuss issues in blended and CALL evaluation such as the dynamic nature of technological advancement; the need for involving a multitude of stakeholders in evaluation and looking at outcomes over different time spans; and the lack of training and expertise of team members in program evaluation, research methodology, and data analysis. These shortcomings pose great difficulties for iterative program design and evidence-based decision making. In the final section of the chapter, the authors present their arguments for why program evaluation is the appropriate theoretical and methodological basis for improving blended language instruction. It is notable that this is the only chapter in which the work is contextualized in blended learning research, since discussion of blended learning is missing in much of the research typically presented in North American contexts.

If one were to only read one chapter in this book, it should be Chapter 2. This chapter is an incredible (and incredibly dense) synthesis of program evaluation theories and methodologies drawing from multiple disciplines and spanning both traditional and the most recent approaches. The chapter begins by presenting six challenges in language program evaluation: social and political context, stakeholder involvement and diversity, assessment criteria, data types, use of project outcomes, and technology integration. Their comprehensive review of program evaluation approaches particularly focuses on action research and those evaluations that are intended to positively impact the program being evaluated. In their proposed model of evaluation, they identify three evaluation levels (i.e., micro, meso, and macro). They further differentiate what these terms mean depending on the scope of the evaluation project, and neatly and compactly summarize information in tables. Building on Gruba and Hinkelman (2012), they further identify four primary considerations for blended language teaching: purpose, appropriateness, multimodality, and sustainability. In the section following, they present argument-based approaches to testing and evaluation as defined by Kane (2006). This approach to evaluation includes four phases: (1) planning an argument, (2) gathering the evidence, (3) presenting the argument, and (4) appraising the argument. What is particularly striking about the proposed model is the focus on complexity of context, ethics, and the impact of the evaluation approach. While an argument-based approach might be new to readers in North American foreign language education, it is an approach that has found a following in English language teaching circles around the world.

Chapters 3, 4, and 5 are the most applied chapters in this book and can serve as a guidebook for practitioners. Each chapter is dedicated to a different level of evaluation (i.e., micro, meso, and macro). In each chapter, the authors summarize concerns, considerations, questions, and data points for each of the four research project phases. The authors go to great lengths to exemplify how rich data can be collected in ethical and methodologically sound ways with the intention of positively impacting the program being evaluated, carefully weighing the benefits and appropriate uses of qualitative and quantitative data. The authors mention several useful concrete instruments and offer tables and questions that can be used as road maps in all phases of a language program evaluation project. Once again, it should be noted that these chapters are relevant not only for blended language program evaluation, but also for language program evaluation

in general.

The practitioner reading this book may want to skip Chapters 6, 7, 8, and 9, while those interested in research on blended language learning may want to skip ahead to these case study chapters after reading Chapter 2. In these four chapters, the authors present four case studies implementing the evaluation approach illustrated in Chapter 2. What is particularly impressive about these case studies is that they include all three evaluation levels and they take place on four continents in different program types. They are each written by four research teams led by one of the co-authors of the book. They share the same approach, are primarily conducted on English learners, and are all in college settings. Hence, the differences and commonalities provide an excellent overview of issues with regards to the implementation of the research framework and the implementation of a blended curriculum. Normally, switching between co-authored chapters and team-written chapters might lead to stylistic inconsistency and be challenging for readers. However, breaking up the co-authored book with these essentially single-authored book chapters is an effective way of illustrating the diversity and complexity of blended learning. The details in which the contexts and procedures are described in each of these four case study chapters serve as excellent examples of qualitative research and of the importance that contextual factors may play in program outcomes. In reviewing these four case studies, one finds great value in Kumaravadivelu's post-method argument for a pedagogy of particularity, practicality, and possibility (Kumaravadivelu, 2006).

The honesty of the authors' reflection in Chapter 10 is a refreshing reminder that both curriculum design and research is an iterative process where each iteration builds on the findings from the previous one. After brief individual reflections, the chapter revisits and redefines the evaluation approach presented in Chapter 2 by first walking the reader through the elements that were successful and then pointing out challenges. Notable strengths of this approach include its connection to existing applied linguistics research and its promise to make research results more transferable due to the rigor in methodology, as well as the salient focus on transparency and program improvement. A clear roadmap for the evaluation process, which makes all elements transparent for all stakeholders, is key to ethical research and to promoting buy-in from stakeholders. The most valuable aspect of the approach is the focus on providing programs with evidence for decision-making and program improvement. The authors also acknowledge that implementing the approach was not without challenges. Most notable was the complexity of the framework, which was not learned easily (indeed, this excellent overview chapter should help future researchers overcome some of the shortcomings of the case studies). In the final section of the chapter, the authors provide directions for future research and modifications to the approach. They propose adding a fourth consideration to the considerations of Gruba and Hinkelman (2012), namely alignment, which demands that macro, meso, and micro level initiatives and evaluations are linked to each other for more cohesion. Furthermore, they point out the value in learner analytics and their potential in blended and online learning. As might be expected, they also encourage a stronger connection and possible foundation in complexity theory, specifically the theory of complex adaptive systems. While the connection to complexity theory was just a small note in the book overall, given the approach presented and the context (blended learning), a stronger connection to and integration of complexity theory would have made the book even stronger.

For those working in language program administration or in language resource centers, Gruba et al. have put together an excellent resource to assist in evaluation and program development. Rarely does one see such a diverse presentation of contexts and resources in a single volume, as each author clearly provides an integral component. The publication is also a good reminder of the different discourses held in North America and elsewhere and in English language teaching in comparison to foreign language education, which is both a strength and a weakness of the book. On one hand, it is an important contribution to the discussion of language program evaluation and blended language learning in US venues for CALL research. On the other hand, in various places, the connection to these discussions is missing, especially in regard to blended learning. Finally, the book title is somewhat misleading. As illustrated in this review, the book is an excellent practical and theoretical guide for program evaluation. However, the first two words in the title, *blended language*, do not receive as much attention, which might have been intentional, since the first author has already published a book on blended language learning. A reader interested in the topic may

want to consider reading both books in sequence to truly gain expertise in both fields. Similarly, those who are new to program administration and evaluation might want to also consult Lord's (2013) guide for language program coordinators and the language program evaluation guide by Norris, Davis, Sinicrope, and Watanabe (2009). Those already experienced and knowledgeable in program administration and research on language programs will find in this book an invaluable resource for comprehensive, contemporary, inclusive, and innovative language program evaluation that takes into account the complexities and particularities of specific contexts.

References

- Arnold, N., & Ducate, L. (2015). Contextualized views of practices and competencies in CALL teacher education research. *Language Learning & Technology*, 19(1), 1–9. Retrieved from <http://ilt.msu.edu/issues/february2015/commentary.pdf>
- Chapelle, C. (2009). The relationship between second language acquisition theory and computer-assisted language learning. *Modern Language Journal*, 93, 741–753.
- Goertler, S. (2011). Blended and open/online learning: Adapting to a changing world of language teaching. In N. Arnold & L. Ducate (Eds.), *Present and future promises of CALL: From theory and research to new directions in language teaching* (pp. 471–501). San Marcos, TX: CALICO.
- Grgurović, M. (2011). Blended learning in an ESL class: A case study. *CALICO Journal*, 29(1), 100–117.
- Gruba, P. & Hinkelman, D. (2012). *Blending technologies in second language classrooms*. Basingstoke, UK: Palgrave Macmillan.
- Kane, M. (2006). Validation. In R. Brennen (Ed.), *Educational measurement* (pp. 17–64). Westport, CT: Praeger.
- Kim, J., Park, H., Jang, M., & Nam, H. (2017). Exploring flipped classroom effects on second language learners' cognitive processing. *Foreign Language Annals*, 50(2), 260–284. doi: 10.1111/flan.12260
- Kumaravadivelu, B. (2006). TESOL methods: Changing tracks, challenging trends. *TESOL Quarterly*, 40(1), 59–81.
- Lord, G. (2013). *Language program direction: Theory and practice*. New York, NY: Pearson.
- Norris, J. M., Davis, J. M., Sinicrope, C., & Watanabe, Y. (Eds.). (2009). *Toward useful program evaluation in college foreign language education*. Honolulu, HI: University of Hawai'i National Foreign Language Resource Center.

About the Author

Senta Goertler (PhD, University of Arizona) is Associate Professor of Second Language Studies and German at Michigan State University and currently serving as Resident Director for the Academic Year in Freiburg program in Germany. Her research focuses on blended and online learning, computer mediated communication, and education abroad.

E-mail: goertler@msu.edu