In the previous chapters, we introduced language program evaluation and an argument-based approach to blended language program evaluation. Further, we argued that blended language programs were best understood at three levels – micro, meso, and macro – with a view that the boundaries of each level could be tailored to fit the needs of a specific evaluation project. The aim of this chapter is to discuss program evaluation at the micro level. We define the micro level as the place where interactions with technology are made evident through the selection and use of activities, groupings, and tools that reflect both teachers and learners’ teaching and learning philosophies and capacities.

After discussing the reasons for evaluation at the micro level, we exemplify two of the four stages of argument development that were described in Chapter 2: planning an argument and gathering the evidence. In planning an argument, we discuss the importance of working with stakeholders to identify focal questions and to narrow down the scope of the evaluation project by selecting the consideration that best fits the level of inquiry. We also discuss how work with stakeholders is paramount to set up the structure of an argument and provide a sample argument to guide its development. In the gathering the evidence stage, we brainstorm potential measures that evaluators must take to comply with ethical considerations. Our argument that blended language program evaluations are best served by qualitative inquiry and particularly from action research approaches is made evident in this chapter. We then present instruments for data collection focusing on questionnaires, interviews, and classroom observations. We finish this chapter with an overview of data analysis procedures.
The purpose of a micro-level evaluation

Arguably, the micro level has received the most attention in the field of language program evaluation. In the work of Kiely and Rea-Dickins (2005), for example, the classroom is understood as the place where pedagogy is instantiated and becomes real. The primary purpose at the micro level is to see the extent, variety, and effectiveness of technology integration into language learning activities in the classroom and the impact of such integration on language learning outcomes (i.e., development of L2 proficiency). Students who are directed to search and read websites, for example, would presumably be less likely to advance further in language development than those who were directed to use the websites to build a presentation. In the latter example, the single technology – websites – can be used, not only for reading practice, but also as a basis for oral presentation skills and research experience. In later years of the program, for example, students could build websites themselves in the target language in ways that incorporate digital literacy with small group work in language proficiency (Laborda, 2009).

As discussed in Chapter 2, blended approaches differ from CALL projects and activities in that they are the result of an overall design philosophy that seeks to be purposeful, multimodal, appropriate, and sustainable (Gruba & Hinkelman, 2012). A CALL project may not have these attributes, but rather be judged only on the merits of its ability to promote second language acquisition. Ideally, the strength of blended approaches in the classroom – in this case, at the micro level – is that they allow for a variety of actions, text types, activities, times, and groupings to be configured within a single face-to-face setting (Gruba & Hinkelman, 2012). Nonetheless, it is this very ability to present a range of options that makes blended programs so challenging to evaluate. What particular element, or configuration, should be the focus of evaluation at micro level?

In her revision of thinking about approaches to instruction in tertiary institutions, Laurillard (2002) noted that lesson planning must incorporate a larger variety of teaching and learning activities. Rather than maintaining a reliance on narrative style of instruction that occurs, for example, when teachers present and students listen, Laurillard (2002) set out a framework that encouraged more interaction, adaptation, communication, and production amongst members of a learning community. In this way, the use of technology was motivated by a pedagogical action that encouraged the use of a greater range of activities as a result of the combination of different texts and tools.

Gruba and Hinkelman (2012) built on such work to propose five dimensions of technology that not only included a variety of actions,