

## Collaborative Development and Testing of a Textbook Evaluation Instrument

Thomas Baker, Kathy Montoya, Paddy Odu, and Carola Villegas

---

*Summary:* This article describes the collaborative development and testing of a pre-use, textbook evaluation instrument aimed at consensus-building to facilitate an acceptance or rejection decision.

*Keywords:* textbook evaluation instrument, collaboration, consensus building

---

Textbook evaluation is a complex activity (Tomlinson, 2003; Chambers, 1997). This is because teachers must consider a wide variety of factors, such as the teacher, the learners, the textbook, the context, the purpose, etc. A further layer of complexity comes when teachers are not included in the evaluation and selection process. As a result, teachers could become frustrated with this situation. This is likely when teachers feel their opinion does not matter in a fundamental issue that directly affects them. Therefore, it is advisable to involve all members of the teaching team in the textbook evaluation and selection process.

The purpose of this article is to describe the collaborative development and testing of a pre-use, textbook evaluation instrument. The evaluation is leveraged to promote consensus building in order to facilitate an accept or reject decision about the textbook. For this project, our course instructor gave a group of four, MA/TEFL students the following open-ended guidance: “Individually develop your own practical, user-friendly, evaluation instrument, using the Tomlinson (2003) and Rubdy (2003) articles as a resource. Afterwards, meet as a group and consolidate your efforts into a single instrument.”

### Instrument Development

Firstly, the development of our evaluation instrument was a collaborative process. To begin, each member of our team independently designed an assessment instrument. This included the pragmatic and pedagogic concerns each member of the group believed the textbook should address, based upon our training and experience as teachers. Next, we met to consolidate our efforts into one final instrument. At this meeting, rather than select the best model of the four, we decided to respect the knowledge

and experience of each group member by clustering all of our questions into five major categories. These categories were

- the curriculum,
- topics,
- activities,
- layout, and
- cost.

Under these categories, we placed subcategories. We added a cluster of questions for each subcategory to facilitate a holistic, subjective, qualitative evaluation. To make the instrument more robust, we used a 5-point Likert scale to provide a quantitative reference.

Here it is important to note that our instrument is not norm-referenced. We also did no training to increase the rater reliability. Had we done this, it is likely our evaluations would not have had the large gaps between the evaluation scores that we did. Having said that, we deliberately set out to develop an instrument that would be capable of actively involving teachers in the evaluation and textbook selection process. Ultimately, our willingness to reach a consensus and support the textbook selected by the group was positively affected.

Furthermore, we believe our instrument has value. First, it reflects our collective experience as educators. On a team of teachers, we feel this is no small matter. It means decisions of the group will likely meet less resistance and a higher degree of cooperation, since the instrument represents the teaching team collectively. Second, by clustering our questions in qualitative groupings for holistic assessment, we have the added benefit that a potentially time-consuming task (answering over a 100 questions), now requires only 15 – 20 minutes. This makes the best possible use of the limited time teachers have available.

A further benefit is consensus-building, which is a Chilean

cultural characteristic. This is leveraged to have the support of the group for the final decision. It is important because the perfect textbook does not exist. According to Tomlinson and Rubdy (2003), modifying and supplementing a coursebook is to be expected. The support of the group, therefore, is crucial to successful implementation and use of this textbook.

Overall, the instrument we developed (see Figure 1) is similar to Rubdy's dynamic cluster approach (2003, p. 46). Rubdy puts the learner, the teacher, and the materials at the center of textbook evaluation. Rubdy's model allows evaluators to use their subjective judgment as they consider questions holistically on multiple evaluation aspects rather than in isolation.

Our group expanded Rubdy's model further by creating two specific categories of our own: curriculum and cost. The curriculum is important because public school textbooks must follow the national curriculum set by the Chilean Ministry of Education (Mineduc). Likewise, cost is important because Mineduc does not have unlimited financial resources. Due to their importance, we added these two categories to our evaluation instrument.

## Instrument Testing

The book we chose to evaluate was the Chilean 6th grade English coursebook, *Get ready with English 6!* (2021). Our evaluation is important because recent studies in Chile (Guernica consultores, 2016; Estudios y consultoría Focus, 2017; EDECSA, 2017) have found that many teachers have an unfavorable opinion of the English textbooks provided by Mineduc. Therefore, there is a need for this kind of teacher-designed, coursebook evaluation.

With our instrument now in place, we could proceed to the actual evaluation of the coursebook. Each member of the group followed the same procedure outlined by Tomlinson (2003, p. 23) to test our evaluation instrument. Each member of the group evaluated the coursebook in isolation, with no input, knowledge, or guidance from any of the other members. That said, collectively our group has accumulated over 50 years of experience.

This is an asset for us, because we bring a wide range of teaching and learning experiences to this task. For example, one member of the group currently works in the private sector, with a binational organization. Previously, this member was with the English Opens Doors Program, a joint effort by the United Nations and Mineduc to promote the teaching and learning of English in Chile. Another works in the tertiary system, teaching English to future lawyers at one of the leading universities in the country. Another teacher is employed in the private sector, working with elementary students at the target level for this evaluation, sixth grade. Finally, one member of the group has over twenty years of experience working in both the public and private system, from tertiary level to

elementary level, even serving one term as the President of TESOL Chile in 2010.

As might be expected, given the wide range of experience within the group, our individual results were not convergent. One member tallied 70 points out of a possible total of 85. Another teacher tallied 60 points and a third teacher tallied only 42 points. We had the foresight, however, to hold one teacher's evaluation in reserve to facilitate an odd number of evaluators to ensure there would be no deadlocked final decision (2 votes for vs 2 votes against). In an actual English department with an even number of teachers, a tie-breaking mechanism of some sort should be devised.

However, we felt that three evaluators were sufficient to establish a pattern. If this did not occur, the fourth evaluator's assessment could as a minimum be used to establish the average score of the group as a whole. At best, this fourth score might closely mirror one of the other three evaluators. While seemingly inconsequential, it could aid in group consensus-building by adding another perspective for consideration.

## Scoring

The evaluation of the teacher who teaches future lawyers at the university level was not among the evaluations mentioned already. The average score the three evaluators arrived at was 57 points. We believe another group of evaluators will likely have a different average score than we did. This was also the sentiment expressed by Tomlinson (2003, p. 23).

Nonetheless, we had a significant difference (28 points) between the highest and lowest scores. We therefore asked our classmate who had not originally evaluated the textbook to do so now, without being told any of the original scores. The result was 70 points. Both teachers (who scored 70 points) stated that it was very important for them that the book covered all of the learning objectives stated in the curriculum.

The teacher scoring the coursebook at the lowest level stated that the teacher's guidebook used the PPP paradigm to teach grammar throughout the book. This behavioristic approach, based on the belief that

*Rubdy puts the learner, the teacher, and the materials at the center of textbook evaluation. Rubdy's model allows evaluators to use their subjective judgement as they consider questions holistically on multiple evaluation aspects rather than in isolation.*

learning happens in a linear fashion, has been widely discredited in ELT over the past 20 years. Moreover, it is teacher-centered in the presentation and practice phase. In the production phase, it prescribes that the grammar presented and practiced now be produced. This is not how real-world communication works. Overall, it is inconsistent with the communicative methodology the national curriculum calls for.

The more moderate teacher (who tallied 60 points) mentioned neither of these points. It was felt that the book had much to recommend itself to teachers. The major observation against the book, however, was that it often failed to account for students' prior knowledge. This teacher identified many instances where activities began by pre-teaching vocabulary, but did not account

for what students already knew about the topic.

We highlight our differences here for a reason. Our aim with this instrument is not to condemn the book. On the contrary, our goal is to facilitate a group consensus. The team's ultimate decision, to accept or reject, will have a solid pragmatic and pedagogic foundation underpinning it. If a team feels that the book, despite its known shortcomings, will benefit their students' learning, they will accept it for implementation. If, on the other hand, the book will not benefit their students' learning, the team will reject it, even if it has known strengths. This group consensus, we believe, is critical to the successful implementation and use of the textbook. For the record, had we been members of the same English department, our group would have accepted this book for implementation.

This is because issues that caused us concern (students' prior knowledge not addressed, PPP grammar presentations) are remediable with a minimum of effort by the teacher.

### Textbook Evaluation Instrument

Scale: 1 = Totally disagree 2 = Disagree 3 = Neutral 4 = Agree 5 = Totally agree

CATEGORIES & SUBCATEGORIES	1	2	3	4	5
<b>1. Curriculum</b>					
A. <b>Methodology:</b> Do the methodology and objectives match the curriculum?					
B. <b>Speaking:</b> Contextualised? Challenging? Interesting? Right level? Motivating? Promotes meaningful communication?					
C. <b>Listening:</b> Pre-, During, Post Activities? Challenging? Contextualized? Right level? Progression in complexity?					
D. <b>Reading:</b> Pre-, During, Post Activities? Challenging? Right level? Diversity of topics? Adequate quantity of text?					
E. <b>Writing:</b> Purpose? Audience? Genre? Interesting? Right level? Challenging? Motivating?					
F. <b>Vocabulary:</b> Contextualised? Cognates? Lexical sets? Right level? Progression? Concrete and abstract vocabulary? Recycling?					
G. <b>Grammar:</b> Accuracy? Fluency? Collocations? Functional? Socio-culturally situated? Progression from simple to complex? Recycling?					
H. <b>Pronunciation:</b> Phonology practice adequate, contextualized, and recycled? Connected speech? Appropriate?					
I. <b>Higher Order Thinking Skills (HOTS):</b> Analysis? Synthesis? Inference? Evaluation? Creativity? Critical thinking? Problem solving? Debate?					
J. <b>Metacognition:</b> Attention to processes used to plan, monitor, and assess one's understanding and performance?					
K. <b>Learning Strategies:</b> Helps students learn how to learn? Attention to ways of organizing & using a particular set of skills to learn content or accomplish other tasks?					
<b>2. Topics</b>					
A. <b>Authentic:</b> Current? Real? Interesting? Local and target culture inclusion? Motivating? Dis/abilities? Progression of knowledge? Culturally relevant?					
B. <b>Sensitivity:</b> Gender? Unbiased? Tolerance for immigrants? Caters to diversity? Cultural sensitivities?					
<b>3. Activities</b>					
A. <b>Classroom Activities &amp; Projects:</b> Fun? Challenging? Right level? Caters to different learning styles? Pairwork? Groupwork? Teamwork? Games address form and meaning? Variety (eg Posters, presentations, roleplays)?					
B. <b>Social skills:</b> Holistic? Citizenship? Democratic?					
<b>4. Layout</b>					
A. <b>Layout:</b> Attractive? Color? High-quality images? Print? Graphics? Contextualised? Informative?					
<b>5. Cost</b>					
A. <b>Cost:</b> Good value? Teacher's Guide matches the Student Book/useful for lesson planning and yearly planning? Workbook? Multimedia? Published within the last 2 years?					
<b>TOTAL POINTS:</b>					/85

### Conclusion

In conclusion, we recommend this evaluation instrument without reservation for teachers working in contexts similar to our Chilean context. This is because the aim of the instrument is to facilitate a consensus between the members of the teaching team.

As a final reflection, we would like to share three reasons for our recommendation. One, evaluating your coursebook is good professional development (Tomlinson, 2003). Two, it acquaints you with the strengths and weaknesses of your coursebook. Three, it puts you in a position to recommend a coursebook to your school leadership team that will better serve your students' needs. These potential benefits are well worth the time you invest in evaluating your coursebook.

Figure 1: Textbook Evaluation Instrument. Click to expand the image.

## REFERENCES

- Chambers, F. (1997). Seeking consensus in coursebook evaluation. *ELT journal*, 51(1), 29-35.
- Dunne, B. & Newton, R. (2021). *Get ready with English 6*. Richmond Publishing. [https://www.curriculumnacional.cl/614/articles-145464\\_textoescolar\\_muestra.pdf](https://www.curriculumnacional.cl/614/articles-145464_textoescolar_muestra.pdf)
- EDECSA, (2017). Proyecto de usabilidad de textos escolares en enseñanza básica: Informe final. [Usability project for elementary school books: Final report.] Encargado por [Commissioned by] MINEDUC y PNUD, Chile.
- Estudios y consultorías Focus, (2017). Recursos digitales y su impacto en el proceso de enseñanza y aprendizaje: Informe final. [Digital resources and their impact in the teaching and learning process: Final report.] Encargado por [Commissioned by] MINEDUC y PNUD, Chile.
- Guernica Consultores S.A. (2016). Estudio de Uso y Valoración de Textos Escolares: Informe final. [Study of the use and valuation of textbooks: Final report.] Encargado por [Commissioned by] MINEDUC y Oficina Regional de Educación para América Latina y el Caribe (UNESCO). Santiago, Chile: María Pía Olivera Vidal.
- Rudby, R. (2003). "Selection of materials." In Tomlinson, B. (ed.) *Developing materials for language teaching*. London: Continuum. pp. 37-57.
- Tomlinson, B. (2003). *Developing materials for language teaching*. London, UK: Continuum.



**Thomas Baker** has worked with EFL students from elementary to university level. He is the Past-president of TESOL Chile (2010 – 2011). Currently, he is doing an MA in TEFL at Universidad Andres Bello in Santiago, Chile. You may contact him at [thomas-baker1999@gmail.com](mailto:thomas-baker1999@gmail.com).



**Kathy Montoya** is currently doing an MA in TEFL and works in the elementary and preschool levels in a school in Chile. She has teaching experience in Australia and periodically teaches EFL for the British Council of Chile. You may contact her at [Kath.montoyar@gmail.com](mailto:Kath.montoyar@gmail.com).



**Paddy Odu** hails from Nigeria and is currently an Ambassador to English Opens Doors Program Chile. He has 6 years international teaching experience. He is doing an MA in TEFL at UNAB. You may contact him at [paddyworldsix@gmail.com](mailto:paddyworldsix@gmail.com).



**Carola Villegas** is an English-Spanish Interpreter with over 13 years of EFL teaching experience. Currently, she is teaching English at University of Chile, and doing an MA in TEFL at UNAB in Santiago, Chile. You may contact her at [villegas.carola@gmail.com](mailto:villegas.carola@gmail.com).

## WAESOL EDUCATOR

*a professional journal for English language educators*



**SUMMER 2022: Call for Submissions**

## Consider an article submission

**DEADLINE: APRIL 15, 2022**

- share teaching tips
- describe research that impacts your teaching
- reflect on a teaching practice that might benefit others
- report what you have learned at a conference
- explain the benefits of new texts, technology, and resources
- highlight the work of a colleague, student, or program

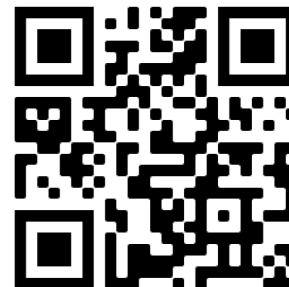
**SUBMISSION  
GUIDELINES**



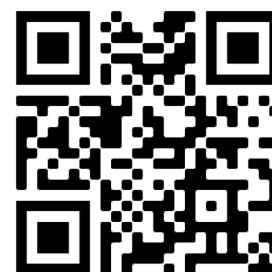
**READ THE  
JOURNAL**



We're so excited to announce the 33rd annual Spokane Regional ESL Conference happening on **Saturday, February 26, 2022!** Just like in 2021, it will be online via Zoom and SCHED, the popular event scheduling system. This conference has been a great place for teachers to get inspired, meet lifelong friends, find new opportunities, get OSPI clock hours, and fall in love with their craft all over again—our goal is to continue this magic in 2022! Please join us.



This year we are expanding our grant program to include more grants of higher amounts. Check out the details at this QR code.



**Spokane Regional** **E S L** **Conference**  
Inspiring Teachers

[spokaneesl.com](http://spokaneesl.com)