

Review of “Assessing English for Professional Purposes (Knoch & Macqueen, 2020)”

Anahita Bordbar*

Department of Foreign Languages Teaching, Ahv. C., Islamic Azad University, Ahvaz,
Iran

DOI: [10.71864/LCT-2025-1217683](https://doi.org/10.71864/LCT-2025-1217683)

Received: 06/04/2025

Revised: 12/07/2025

Accepted: 15/07/2025

1. Introduction

English for Specific Purposes (ESP) is an approach to language teaching that focuses on developing the linguistic and communicative skills required for a particular field, occupation or academic discipline. Unlike general English courses, ESP programs are designed around learners' specific goals such as business English for international trade, medical English for healthcare professionals or academic English for research scholars. ESP is characterized by a needs-based analysis, meaning that course content, vocabulary, and communicative activities are drawn directly from the contexts in which learners will use English. It emphasizes practicality and relevance often incorporating authentic materials such as workplace documents, technical reports or real-life scenarios. By tailoring instruction to the precise communicative tasks and professional realities of its learners, ESP enables students to acquire the targeted language skills that support their career advancement and academic success.

Since ESP is essential for equipping learners with the precise language skills required in their academic or professional fields, evaluating ESP is equally important to guarantee that these programs remain effective,

* Corresponding Author's E-mail address: Anahita.bordbar@iau.ac.ir



This work is licensed under a [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/).

relevant, and aligned with learners' specific needs. Evaluation of English for Specific Purposes, particularly in professional contexts, has attracted growing attention over the last two decades. This increased interest is largely driven by globalization and worldwide labor mobility which have heightened the need for valid, reliable, and contextually appropriate measures of English language proficiency. In response to this demand, *Assessing English for Professional Purposes* (2020) by Ute Knoch and Susy Macqueen provides a timely and valuable addition to the literature. The volume as a part of Routledge Research in English for Specific Purposes series revisits pioneering ideas in ESP assessment and provides new theoretical perspectives and methodological approaches. Knoch and Macqueen in this volume have presented a work which not only revises the discussion begun by innovators like Douglas (2000) and MacNamara and Roever (2006), but more socially engaged and ethically informed vision of language testing. The volume has seven chapters, thematically organized to help the readers explore the multifaceted nature of test development and validation in professional settings. It does so in due consideration of the sociopolitical and ethical aspects of language testing. A unifying thesis of the authors is that professional language testing is more than a technical task and that such testing is socially embedded and informed by power relations, institutional requirements, critical perspective on the consequences of tests, and context-sensitive validation, the authors strive to bring the field back into more comprehensive and humanistic purposes.

In the following pages, this review provides an overall critical examination of the book's contents, research methods, theoretical advances, and contribution to the field. It takes up central themes, places the book within the current literature, and provides an assessment of strengths and weaknesses. Lastly, it takes into consideration the book's contribution towards language assessment practice and future research.

2. Structure of the Book

This book includes 7 chapters that are relatively self-contained but interrelated. In another word, readers can read the chapters consecutively, or they can select individual chapters of interest which can leads into other chapters.

Chapter 1: Scope

The first chapter establishes the context through an operationalization of the field of English for Professional Purposes (EPP) testing. Knoch and Macqueen delineate EPP from more broadly defined ESP through an

emphasis on high-stakes, professional settings like aviation, health care, and the law. These scholars assert that conventional methodologies of ESP testing have all too frequently been very narrow-based, emphasizing linguistic correctness and communicative ability while being blind to the sociocultural and professional stakes inherent in these tests (cf. Douglas, 2000; Fulcher & Davidson, 2007). Another of the chapter's strengths is the redefinition of professional language tests as a process of "risk management." Adapting sociological theories (e.g., Giddens, 2023), the authors argue that EPP tests act not only as gatekeeping mechanisms but as devices for managing professional incompetence-related risk in society. This conceptualization provides a precursor to later debates on the political and moral aspects of testing.

Chapter 2: Construct

In the second chapter, Knoch and Macqueen explore construct theory and critique the classical psychometric conception of language constructs as fixed and de-contextualized. Rather, they argue for a socially constructed approach where constructs are co-constructed through stakeholders and dynamic and context-sensitive. Worth noting is their introductions of the "Codes of Relevance" concept. These codes consist of the professionally specific and culturally embedded norms of workplace language use. This chapter also discusses how constructs should be derived from authentic communicative practices rather than from abstract linguistic models. The authors emphasize the importance of construct representation and construct validity, referencing Messick (1996) and Kane's (2013) argument-based validation framework. They assert that a well-articulated construct should reflect not only the linguistic features of professional discourse but also its pragmatic, social, and institutional dimensions.

Chapter 3: Needs

Chapter 3 is dedicated to needs analysis as the empirical support for test design. Knoch and Macqueen outline a systematic process of undertaking needs analysis including ethnographic observation, corpus studies, interviews of stakeholders, and discourse analysis. Needs analysis has to go beyond the identification of language tasks, according to them, in order to unearth underlying competencies, communication-related roles and risk inherent in communication at the workplace. This chapter builds on Brown (2016) and Long (2005), and it emphasizes the multi-method, iterative character of needs analysis. These writers apply healthcare and legal careers as illustrations of how different types of data might be

integrated through the process of triangulation and then applied to informing test specifications. Their approach builds on more recent work on performance-based testing and domain analysis (Douglas, 2000; Purpura, 2016).

Chapter 4: Development

The authors here move from the theoretical to the practical and provide an extensive blueprint for test construction. They guide readers through item development, piloting, scoring rubric formation, and standard-setting methods. Special care is taken with authenticity, difficulty of task, and construct alignment. Empirical data from the development of the Occupational English Test (OET) informs the chapter and offers readers practical illustrations of decision-making in test design. Another of the strengths of this chapter is that it combines both qualitative and quantitative methods of testing development. Knoch and Macqueen present classical test theory (CTT), item response theory (IRT), and G-theory (Generalizability theory), and yet focus on the use of expert judgment and consultation with stakeholders as well. This balanced view is best practice in language testing research (Alderson, Clapham, & Wall, 1995; Bachman, 2010).

Chapter 5: Validity

Validation lies central to any measurement undertaking, and Chapter 5 offers an extensive framework founded on Kane's (2013) argument-based model of validation. The editors outline the inferences within test use and interpretation, such as extrapolation, generalization, scoring, and decision. They also insist on multifaceted sources of evidence such as test-taking perceptions, rater reliability, consequential validity, and internal consistency. Knoch and Macqueen concern themselves particularly with questions of the ethical nature of validity. Adapting Roever and McNamara (2006), they hold that validation must also include technical adequacy and social implications. They caution against tests being instrumentalized towards political ends and instead recommend openness, fairness, and responsibility.

Chapter 6: Policy

The second-to-last chapter investigates the policy landscape of EPP testing. In this one, the authors grapple with questions of regulation, accreditation, and professional responsibility. They examine how governments, licensing bodies, and employers use tests as a way of controlling entry into professions. This chapter also takes a look at

conflicts between national language policies and international standards in multilingual cultures.

A key insight is the recognition that language tests are often used as proxies for broader competencies, such as professional judgment or cultural adaptability. The authors critique this tendency and advocate for more nuanced assessments that reflect the complexity of professional practice. They also call for greater stakeholder involvement in test policy and design.

Chapter 7: Future Directions

The final chapter summarizes the book's arguments and outlines avenues for future research. Knoch and Macqueen stress interdisciplinary collaboration, especially with sociologists, ethicists, and experts in the field. They also note new areas like digital testing, workplace-based testing, and multilingual practice.

3. Critical Evaluation of Methodology and Argumentation

Knoch and Macqueen's methodological approach is theoretically sound and empirically informed. Adopting mixed-methods designs, triangulating ethnographic observation, corpus analysis, stakeholder interviews, and psychometric modeling, authors demonstrate a “multi-perspective” approach to language testing (what Fulcher, 2024, calls such an approach). Their methodological thoroughness is revealed through the in-depth case studies, and the development of the Occupational English Test (OET), in particular, shows consistency between theoretical construct and operationalization.

The emphasis on argument-based validation following Kane (2013) represents a significant strength. Rather than relying solely on statistical indicators of reliability or item-level discrimination indices, the authors construct a validity argument that is both comprehensive and contextually sensitive. This is further enriched by their critical stance on consequential validity, echoing Roever and McNamara's (2006) call for assessments that are socially responsible and ethically informed.

However, the volume does not seriously engage new language testing technology. Whereas Chapter 7 briefly discusses computer testing, the discussion is more suggestive than substantial. As AI-driven score engines and remote proctoring become more widespread in applied testing practice (Harding, Brunfaut, & Unger, 2020; Aryadoust & Fox, 2016), more mature consideration of these tools, and their prospective influence on validity and fairness, would have better placed the volume's discussion in the context of the times.

In addition, although the authors move effectively back and forth between theory and practice, the text sometimes presumes a great deal of background knowledge of sophisticated psychometric and sociolinguistic theory. For example, expositions of generalizability theory and critical discourse analysis might usefully have included more explanation addressed to less experienced readers.

4. Contribution to the Field

Knoch and Macqueen's work has multidimensional significance. Theoretically, their work goes beyond the conventional fields of grammar, vocabulary, and discourse in constructing language proficiency and instead proposes a context- and ethical-sensitive conceptualization of language proficiency. Their notion of "Codes of Relevance" augments prevailing frameworks of communicative competence by underscoring the institutionally and socially oriented norms governing professional discourse.

At a methodological level, the book provides a replicable blueprint for needs analysis, test development, and validation grounded in both qualitative inquiry and quantitative measurement. This aligns with calls in the literature for empirically grounded test design (e.g., Bachman, 2010; Brown & Abeywickrama, 2010).

In practice, the volume serves as an important reference guide to test developers, policymakers, and researchers working on high-stakes language assessments. Its critical analysis of the political applications of language tests, especially in migration and employment, provides a timely response to current debates regarding fairness, access, and inclusivity in assessment (Shohamy, 2020; Shohamy, Or & May, 2017).

One of the most impactful elements is the book's explicit engagement with ethics in assessment. The authors do not treat ethical considerations as ancillary but as integral to the validity of an assessment. This ethical framing aligns with contemporary movements in language testing that seek to humanize assessment practices and make them more accountable to diverse stakeholders.

5. Comparative Literature Analysis

Compared to foundational texts such as Douglas's (2000) *Assessing Languages for Specific Purposes*, Knoch and Macqueen offer a broader and more critical perspective. Whereas Douglas concentrated primarily on theoretical models, test specifications, and alignment of tasks to test constructs, Knoch and Macqueen foreground the social stakes and institutional consequences of assessment: they attend not only to what a

test measures but to who is advantaged or disadvantaged by those measures. Empirical work similarly shows that learners' perceptions of classroom fairness are deeply influenced by institutional practices. For instance, Sadeghi, Jalali and Fatehi Rad (2023) revealed that problems of transparency, consistency, and teacher bias are perceived as central by learners. Further, Baghoulizadeh, and Nosratinia, (2023) demonstrated that assessment practices that engage multiple stakeholders (teachers, peers, self) affect learners' agency and sense of fairness. In this way, Knoch and Macqueen extend beyond test design to address how test systems are embedded in institutional power and cultural expectations.

More proximally, Roever and McNamara (2006) might be considered the closest intellectual antecedent. As with McNamara, Knoch and Macqueen argue for a critical stance that interrogates the purposes and effects of tests. However, Knoch and Macqueen move this agenda further by presenting tangible methodologies for achieving *contextual* and *ethical validity*, not merely critique. Several studies underscore this shift. For instance, Gorji, Afraz, and Samimi (2023) built a model of language quality grounded in learners' lived experience, showing how assessment criteria must be responsive to context. Above and beyond, Kashanizadeh, Ketabi, and Shahrokhi (2024) demonstrated rigorous scale development, including validation procedures that consider cultural settings and specific pedagogical aims. Thus, the work of Knoch and Macqueen not only echoes other researchers' specifically Roever and McNamara's concern with social meaning but also advances the field by proposing tools and procedures for operationalizing ethical and contextual assessment.

Moreover, in relation to contemporary developments such as Harding *et al.* (2020) work on automated scoring or Aryadoust and Fox's (2016) research on multimodal assessment, Knoch and Macqueen's emphasis on human judgment and stakeholder engagement stands out. While they recognize the potential of technology, they remain cautious, advocating for human oversight particularly in design, interpretation, and decision-making. This is paralleled in several studies. For example, Badpa and Zare-Behtash (2023) examined how different modes of feedback (which could be mechanized or human) have diverging effects and suggested that human-mediated feedback remains crucial. Similarly, Sabzevari, Fatehi Rad, and Tajadini (2023) considered technology (LMS) as a mediating environment but reported that learners' perceptions of fairness, autonomy, and thoroughness depend heavily on how teachers and peers interact in and around the system. These findings reinforce Knoch and Macqueen's argument that while automated or technological tools can contribute, they

must be complemented by human judgment and meaningful stakeholder engagement to preserve interpretive validity and ethical legitimacy.

6. Concluding Remarks

Assessing English for Professional Purposes is an innovative text redrawing the scope, purposes, and methodology of ESP testing. Knoch and Macqueen offer an intellectually disciplined and ethically informed approach critiquing reductive views of language testing. In grounding assessment practice in social, institutional, and political contexts, they draw together a theoretically valid and pragmatically useful paradigm. Its most important strengths are in reconceptualizing constructs, methodologically comprehensive approach, and ethical assessment orientation. Its message is relevant to a vast majority of people across diverse roles such as researchers, practitioners, and policy-makers. Though some domains, especially digital testing, need more development, the volume's broader contribution remains important and enduring.

Researchers, also, can extend the contributions of this book in several interconnected directions. First, they can empirically validate the reconceptualized constructs proposed by Knoch and Macqueen across varied professional and cultural contexts, examining whether elements such as communicative competence and workplace discourse skills operate similarly across industries and countries. Second, there is ample scope to expand digital and AI-assisted assessment by exploring the implications of automated scoring, adaptive testing, and online platforms within ESP contexts, while ensuring alignment with the ethical and contextual principles emphasized by the authors. Third, policy-oriented studies can investigate the impact of ESP assessment policies on workforce readiness, equity, and inclusion, including how test outcomes influence hiring, promotion, and professional development. Fourth, the book's positioning of assessment within broader social and institutional frameworks invites cross-disciplinary research that connects language assessment with organizational studies, professional communication, and educational technology. Fifth, longitudinal and impact-focused research can track the long-term effects of ESP assessments on professional performance, career advancement, and educational trajectories, offering evidence for the effectiveness of ethically grounded assessment practices. Finally, inspired by the book's methodological breadth, researchers can pursue innovative mixed-methods designs that integrate qualitative insights from workplace stakeholders with quantitative performance measures, ensuring that assessments remain both technically reliable and relevant to real-world contexts. In short, *Assessing English for*

Professional Purposes is both a conceptual guide and a practical toolkit. It provides a foundation for rethinking ESP assessment as socially responsible, context-sensitive, and ethically sound, and offers multiple pathways for researchers to advance the field, particularly in areas like digital testing, policy impact, and cross-contextual validation.

In an area increasingly characterized by technological advancement and demands for policy direction, Knoch and Macqueen bring home the reality that assessment ends and purposes are not about simple measurement but significant, fair, and socially accountable evaluation. For this reason, *Assessing English for Professional Purposes* ought to be on the reading list of anyone who involves themselves in language assessment in professional life.

Funding: This research received no external funding from any agency.

Conflicts of Interest: The author declares no conflict of interest.

References

- Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge University Press.
- Aryadoust, V., & Fox, J. (Eds.). (2016). *Trends in language assessment research & practice: The view from the Middle East and Pacific Rim*. Cambridge Scholars Publishing.
- Bachman, L. F. (2010). *Language assessment in practice*. Oxford: Oxford University Press.
- Badpa, H., & Zare-Behtash, E. (2023). The impact of explicit and implicit corrective feedback on the writing proficiency of Iranian pre-intermediate EFL learners. *Journal of Language, Culture, and Translation*, 6(1), 144–162. <https://sanad.iau.ir/Journal/lct/Article/1122612/FullText>
- Baghoulizadeh, S., & Nosratinia, M. (2023). The comparative effect of self, peer, and teacher assessment in speaking on EFL learners' autonomy. *Journal of Language, Culture, and Translation*, 6(1), 66–88. <https://sanad.iau.ir/Journal/lct/Article/1012417/FullText>
- Brown, J. D. (2016). *Introducing needs analysis and English for specific purposes*. Routledge.
- Brown, H. D. (2010). *Language assessment: Principles and classroom practices* (3rd ed.). Pearson.
- Douglas, D. (2000). *Assessing languages for specific purposes*. Cambridge University Press.
- Fulcher, G. (2024). *Practical language testing*. London: Routledge.

- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Routledge.
- Giddens, A. (2023). *Modernity and self-identity: Self and society in the late modern age*. Stanford University Press.
- Gorji, F., Afraz, SH., & Samimi, F. (2023). Assessing English language quality of EFL students: Links to a grounded theory. *Journal of Language, Culture, and Translation*, 5(1), 133–149. <https://sanad.iau.ir/Journal/lct/Article/1012400/FullText>
- Harding, L., Brunfaut, T., & Unger, J. W. (2020). Language testing in the 'hostile environment': The discursive construction of 'secure English language testing' in the UK. *Applied Linguistics*, 41(5), 662–687. <https://doi.org/10.1093/applin/amz017>
- Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of educational measurement*, 50(1), 1-73.
- Kashanizadeh, I., Ketabi, S., & Shahrokhi, M. (2024). Designing and validating a technological innovation scale for the English as foreign language learning context. *Journal of Language, Culture, and Translation*, 7(1), 19–42. <https://sanad.iau.ir/Journal/lct/Article/1129883/FullText>
- Long, M. H. (2005). *Second language needs analysis*. Cambridge University Press.
- Roever, C., & McNamara, T. (2006). Language testing: The social dimension. *International Journal of Applied Linguistics*, 16(2), 242-258.
- Messick, S. (1996). Validity of performance assessments. In G. Wiggins (Ed.), *Assessing student performance* (pp. 1–19). Jossey-Bass.
- Purpura, J. E. (2016). Second and foreign language assessment. *The Modern Language Journal*, 100(1), 190-208.
- Sabzevari, M. Fatehi Rad, N. Tajadini M. (2023). A Qualitative Study of Reciprocal Scaffolding Techniques on Iranian EFL Learners' Views via Language Management System, *Journal of Language, Culture, and Translation*. 5 (2), 92-106 [10.30495/lct.2023.1988940.1090](https://doi.org/10.30495/lct.2023.1988940.1090)
- Sadeghi, M., Jalali, V., & Fatehi Rad, N. (2023). Appraising Iranian EFL learners' perceptions towards the teachers' implementation of classroom fairness: Focus on the main problems. *Journal of Language, Culture, and Translation*, 5(2), 175–189. <https://sanad.iau.ir/Journal/lct/Article/1012405/FullText>
- Shohamy, E. (2020). *The power of tests: A critical perspective on the uses of language tests*. London: Longman.

Shohamy, E., Or, I. G., & May, S. (Eds.). (2017). *Language testing and assessment* (3rd ed.). *Encyclopedia of language and education*. Cham: Springer. <https://doi.org/10.1007/978-3-319-02261-1>