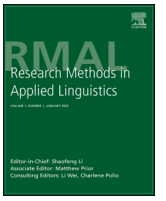
Introducing and validating the teaching-in-English-competencies-scale (TECS) for English-medium instruction in Sweden and Spain

Research Methods in Applied Linguistics 5 (2026) 100326

Contents lists available at ScienceDirect

Research Methods in Applied Linguistics

journal homepage: www.elsevier.com/locate/rmal



Introducing and validating the teaching-in-English-competencies-scale (TECS) for English-medium instruction in Sweden and Spain

Hans Malmström [a,*], Heath Rose [b], Christopher Shepard [b], Emma Dafouz [c]

[a] Department of Communication and Learning in Science, Chalmers University of Technology, Hörsalsvägen 2 412 96, Göteborg, Sweden
[b] Department of Education, University of Oxford, 15 Norham Gardens, Oxford OX2 6PY, UK
[c] Department of English Studies, Universidad Complutense de Madrid, Plaza de Menéndez Pelayo s/n, Ciudad Universitaria 28040, Madrid, Spain

ARTICLE INFO

Keywords:
EMI
Lecturers
Self-efficacy
English proficiency
Language skills
Scale
Teaching challenges

ABSTRACT

This paper presents the development and validation of the Teaching-in-English-competencies-scale (TECS), a novel instrument designed to assess lecturers' perceived language competence, confidence, and preparedness to deliver subject content in English-medium instruction (EMI). After establishing the psychometric properties of the scale, we illustrate its potential analytical applications using data from two contrasting EMI contexts, Sweden and Spain. TECS addresses a critical gap in EMI research, as existing instruments have predominantly focused on lecturers' general English proficiency challenges and on students' linguistic needs. Moreover, to date, no empirically validated instrument has been identified that matches TECS in scope or application. Validation procedures included exploratory factor analysis (EFA) and confirmatory factor analysis (CFA), which yielded a robust and consistent four-factor structure encompassing both productive and receptive competencies, in addition to a single factor for items loaded onto a self-efficacy scale. The statistical performance of the indicators confirmed the suitability of TECS for group comparisons and institutional and individual diagnostics. Its application in two contrasting EMI contexts further illustrates its capacity to capture variation across settings, providing initial evidence of its applicability beyond a single context, while acknowledging that structures found may be context dependent due to the dominance of our sample residing in Sweden. The results indicate that TECS can function as a systematic tool for a wide range of EMI stakeholders, including researchers, teacher educators, and pre- and in-service EMI lecturers.

1. Introduction

In line with the growing use of English as a medium of instruction (EMI) (Agnew & Neghina, 2021; Wingrove et al., 2025), several reports have highlighted challenges experienced by a key stakeholder group: EMI lecturers (e.g., Dang et al., 2023; Shao & Rose, 2024). Many of these challenges relate – directly or indirectly – to lecturers' English skills and their confidence in teaching through English, both of which have been questioned by higher education (HE) stakeholders (e.g., students), but also by EMI lecturers themselves (Dafouz, 2018; Guarda & Helm, 2017; Lasagabaster, 2022; Macaro, 2018). Despite these concerns, few studies have examined this

* Corresponding author at: Department of Communication and Learning in Science, Chalmers University of Technology, Hörsalsvägen 2 412 96, Göteborg, Sweden.
E-mail address: mahans@chalmers.se (H. Malmström).

https://doi.org/10.1016/j.rmal.2026.100326
Received 4 December 2025; Received in revised form 20 May 2026; Accepted 20 May 2026
Available online 1 June 2026

issue systematically across diverse EMI contexts, and none have done so using a single, standardized instrument. The present study introduces such an instrument, enabling an updated account of perceived English-language skills and support requirements.

The Teaching-in-English-competencies-scale (TECS) is an instrument designed to elicit information about EMI lecturers' confidence in their English and their perceived competencies in using English *specifically for teaching in EMI contexts*. TECS enables systematic assessment of lecturers' perceived English communication competencies, including their teaching-in-English self-efficacy, and identification of areas where support and professional development might be needed. In addition, by drawing on comparable data across EMI settings – exemplified here through the administration of the instrument in Sweden and Spain – TECS provides a basis for cross-contextual mapping of EMI lecturers' competencies that can inform both local and international EMI policy and practice.

2. Literature review

2.1. EMI lecturers' classroom English proficiency, language-related challenges and confidence

The centrality of language – English specifically – in EMI teaching is widely acknowledged (Lasagabaster & Doiz, 2021; Macaro, 2020; Pecorari & Malmström, 2018), but research on EMI lecturers' English competence often presents a mixed and context-dependent picture. A growing body of research has examined EMI lecturers' English proficiency and classroom interaction.

Regarding proficiency, some studies offer a positive account of EMI lecturers' English competence, reporting that lecturers demonstrate satisfactory levels of English (sometimes confirmed by institutional certification) as well as strong confidence in their language ability (e.g., Kling & Dimova, 2015; Vander Borght et al., 2025; van Splunder et al., 2022). However, alongside these positive accounts, other studies have highlighted perceived or alleged shortcomings and the challenges these may pose for teaching and learning. Dang et al. (2023, p. 844) describe EMI lecturers' English proficiency as "the most common EMI challenge." Lasagabaster (2022, p.14) similarly observes that lecturers' "English language competence appears as the main stumbling block, to the extent that it may negatively affect the teaching approach and teaching strategies." Kling (2015, p. 218) further notes that lecturers "underplay their limitations in general linguistic proficiency," including difficulties retrieving vocabulary, producing accurate grammar, and managing pronunciation. Recent research by Malmström et al. (2025) indicates that the challenges some EMI lecturers experience might not be confined to general English proficiency but may be even more pronounced in their academic language use.

Concerning classroom interaction, several studies (e.g., Airey, 2011; Lasagabaster, 2022) have highlighted the consequences of EMI lecturers' limited English proficiency for interactional practices. These include reducing the amount or complexity of content, showing less flexibility in teaching, relying on more monologic/lecturer-centered instruction, displaying reduced fluency, struggling to manage spontaneous interaction, explaining complex disciplinary concepts, and responding ineffectively to students' questions. However, research also cautions against attributing potentially "ineffective" classroom performance solely to lower levels of English proficiency. Vander Borght et al. (2025) examined the relationship between EMI lecturers' English proficiency, pedagogical training, self-efficacy, and classroom interactivity, finding that pedagogical training was the strongest predictor of lecturers' adopting a more interactive teaching approach, while English proficiency had only a relatively modest effect.

Research has also examined EMI lecturers' teaching self-efficacy (Bandura, 1977; Tschannen-Moran & Hoy, 2001) and its relationship to their English proficiency. In a case study of EMI lecturers in Taiwan, Tsui (2018) found that they initially reported low self-efficacy, linked to language anxiety and pedagogical uncertainty, but an EMI training program significantly enhanced their confidence, sense of empowerment, and orientation toward student-centered teaching. Similar concerns, particularly when lecturers are new to EMI, are evident in the systematic review by Dang et al. (2023), who found that many EMI lecturers experience reduced teaching self-efficacy due to linguistic insecurity, limited pedagogical preparation, and unclear role expectations. This suggests that language-related challenges are central to lecturers' feelings of uncertainty; as Lasagabaster (2022, p. 14) observes, "lecturers on EMI courses tend to feel insecure due to their inability to tackle language problems."

The relationship between English competence and teaching self-efficacy also seems to vary depending on how proficiency is conceptualized and measured. Two studies may serve as examples. In the first, drawing on 188 EMI lecturers at a Chinese university, Wang (2021) found a strong positive correlation between classroom English proficiency and teaching self-efficacy. Among the dimensions assessed, instructional language use/classroom instructional discourse (lecturers' ability to explain concepts, give instructions, and signal transitions) was the strongest predictor of higher teaching self-efficacy, whereas grammar, pronunciation, and language of interaction appeared to be less important. In the second, a large-scale study of 978 EMI lecturers in Sweden (Vander Borght et al., 2025), a more limited effect of English proficiency (self-reported CEFR) was found: within a highly proficient sample (88.8% self-rated at CEFR C2 level), only speaking proficiency significantly predicted teaching self-efficacy, while reading, writing, and listening appeared to be unrelated. These studies suggest that (certain) productive language skills – particularly classroom instructional discourse and spoken classroom interaction – are associated with greater lecturer teaching confidence; however, at the same time, it is possible that the relative importance of proficiency may diminish once a high baseline level is reached or institutional proficiency requirements are met.

Cross-country research also indicates that lecturers' English proficiency and preparedness to deliver subject content are central concerns among students. In Spain, for example, Aguilar and Rodríguez (2012) found that students' perceptions of lecturers' English proficiency strongly influenced their overall attitudes toward EMI, with "lecturers' insufficient level of English" identified as the most prominent negative aspect (p. 183). Similarly, Galloway and Ruegg (2022), drawing on data from nine universities in Japan and China, found that lecturers' English proficiency and willingness to teach through English were students' main concerns, often leading to comprehension difficulties. In Tunisia, research by Abdeljaoued (2023) showed that students rated their lecturers' English proficiency as lower than their own, and about two-thirds reported difficulty understanding them.

2.2. Comparative perspectives on EMI teaching

Despite the rapid growth of EMI worldwide, there remains a notable lack of comparative research for examining teaching across disciplines, institutions, and countries (Shao & Rose, 2024), limiting the field's ability to identify broader patterns in EMI practice. Reflecting this gap, Macaro (2018, p. 94) calls for "comparative methodologies…in different contexts," and Dimova et al. (2015, p.319) similarly argue that "the field [of EMI research] could be enhanced by cross-national, contrastive studies," specifically highlighting language proficiency as an area where comparative perspectives are particularly needed.

Although still few in number, recent comparative studies provide illustrative examples of how EMI lecturers experience their roles and teaching challenges across different institutional and national contexts. In a study of 28 lecturers from 10 European universities, Orduna-Nocito and Sánchez-García (2022) identified both convergences and divergences between policy and practice: while policies emphasized minimum proficiency thresholds and international visibility, lecturers drew attention to pedagogical strategies, assessment challenges, and the need for stronger educational support. Similar tensions between institutional ambitions and teaching realities were evident in Austria, Italy, and Poland, where Dearden and Macaro's (2016) interviews with 25 lecturers revealed generally positive views of EMI but also concerns about linguistic insecurity, the pedagogical demands of teaching through an additional language, and reduced confidence in explaining complex concepts, with some cross-country variation linked to institutional support and EMI implementation models. Extending these insights beyond Europe, Shao and Rose's (2024) cross-case study of 19 business and management lecturers in China, Japan, and the Netherlands likewise highlighted contextual variation: while Dutch lecturers reported relatively few difficulties, their Chinese and Japanese counterparts expressed greater insecurity and teaching challenges.

These studies show the value of comparative research in EMI settings. Yet, as Rose et al. (2023) note in their call for the community to strive for comparative equivalence in EMI research, not all comparative designs allow for findings to be interpreted on the same terms: while Shao and Rose (2024) ensured disciplinary equivalence by focusing on business and management, other studies do not offer the same level of control over contextual variation making findings less directly comparable. Building on previous work in this cross-contextual space, we believe there is value in the development of research instruments that can accommodate disciplinary and institutional diversity while still producing systematically comparable data, thereby capturing both the context-specific realities of EMI lecturers' experiences and enabling broader, cross-context generalizations about their confidence, teaching language proficiency, and perceived challenges.

3. Instrument development: Designing and trialing TECS

To develop an instrument that elicits information about EMI lecturers' confidence in teaching through English and their perceived competencies in performing classroom-communicative tasks in English, we drew on two main complementary sources: (i) an EMI challenges scale (ECS), mainly used in EMI research in Asia, and (ii) the skills descriptors contained in the Common European Framework of Reference for Languages (CEFR).

3.1. Sources of inspiration and conceptual foundations

The first source of inspiration was the instrument commonly referred to in EMI research as the "EMI Challenges Scale" or "EMI Challenges Survey" (Gaffas, 2025; Soruc et al., 2021), hereafter ECS. The 45 items in this scale were originally developed by Evans and Green (2007) to investigate students' difficulties when studying disciplinary content through English and were later employed by Evans and Morrison (2011) in their study of undergraduate students' language-related challenges at an EMI university in Hong Kong. The ECS has since been used in a number of EMI studies internationally (e.g., Gaffas, 2025; Shepard & Morrison, 2021; Soruc et al., 2021). Organized around the four macro-skills (reading, writing, speaking, listening), its multidimensional self-report structure informed the design of TECS. However, because TECS focuses on lecturers rather than students, ECS items were adapted and reconceptualized to reflect the pedagogical communicative activities involved in teaching disciplinary content through English.

Our second source of inspiration was the CEFR Companion Volume (Council of Europe, 2023); the CEFR skills descriptors provided a foundation for conceptualizing what it means to be linguistically and communicatively competent in EMI teaching and learning settings. The "four skills" framework offers a widely recognized, transparent and comparable basis for organizing descriptors of communicative performance, and with some minor revisions to the "can-do" descriptors, it aligns well with the practice of teaching. The CEFR frames language proficiency not as static linguistic knowledge but as the ability to perform communicative language activities – to act effectively through language in real contexts. Its emphasis on interaction and mediation was particularly useful for describing EMI lecturer competence, yet in our design these dimensions were not treated as separate constructs. Rather, the notion of mediation, understood as the ability to facilitate understanding, build conceptual bridges, and manage interactional flow, was integrated throughout the prompts relating to speaking, writing, listening, and reading. For instance, several items tap into lecturers' ability to present disciplinary content clearly, interpret students' questions, provide feedback, and select or adapt written materials – all communicative acts that involve mediating knowledge and meaning. In this way, CEFR-based notions of mediation and interaction informed the overall construct and item formulation, while remaining embedded in the task-specific operationalization of the four language skills for teaching in EMI.

The ECS and the CEFR Companion Volume together guided the operationalization of the TECS constructs. We conceptualize EMI lecturers' English competence as a multifaceted construct comprising their self-perceived confidence and their perceived ease or difficulty in performing communicative and pedagogic tasks in English. The resulting instrument, TECS, is structured around two complementary dimensions. The first targets lecturers' self-efficacy beliefs, capturing how confident they feel using English for key

teaching functions, preparing materials, lecturing, interacting with students, and assessing learning. The second focuses on language-related competencies, asking lecturers to evaluate the ease or difficulty of specific speaking, writing, listening, and reading tasks as they occur in EMI teaching situations. These tasks, ranging from giving verbal instructions and presenting disciplinary content to understanding students' comments and reading their assignments, are intended to mirror the communicative demands of EMI teaching and align conceptually with CEFR descriptors for communicative language competence.

The instrument design is meant to link language proficiency frameworks and teaching practice, and it integrates CEFR-based conceptualizations of communicative ability with empirically grounded descriptions of EMI classroom language use. In other words, the overarching aim of TECS is not to measure linguistic proficiency per se, but to capture how EMI lecturers perceive their confidence and preparedness to use English as a tool for disciplinary teaching and interaction.

As a self-report instrument, TECS captures respondents' perceptions of their communicative competence and teaching self-efficacy rather than objective performance. Such responses reflect not only individuals' abilities, but also their interpretation of items, use of response scales, and self-evaluative norms. We acknowledge that responses may vary across demographic and contextual variables, including gender and cultural background; such variation should be considered when interpreting the results.

3.2. Instrument design and development

Initial design and development of the survey took approximately three months (December 2023–February 2024). The project team first met in early December 2023 to discuss the conceptualization and feasibility of the instrument design and validation study, after which a first draft of the TECS was generated by the first author later that month. This was followed by a team review in January 2024 and further discussions in February 2024, resulting in a revised TECS draft circulated in mid-February.

The mid-February version comprised three sections. Section 1 focused on lecturer efficacy and initially contained one item asking respondents to rate their overall confidence in using English for teaching. Section 2 targeted EMI lecturers' language-related challenges and included 35 items in total, which consisted of 31 Likert-type items across four skills – speaking (8), writing (8), listening (7), and reading (8), as well as four optional open-ended questions. Section 3 included nine items related to respondents' teaching profile, background, and English proficiency.

During a subsequent revision cycle, three additional confidence items were added to Section 1 (four items total). This was based on our review of the literature on self-efficacy and EMI, which called for multi-item measures of self-efficacy due to the limitations of using a single item in subsequent analyses (Thompson et al., 2022). The scale was therefore expanded to include four items representing key EMI teaching functions: delivering teaching (CONF1), interacting with students (CONF2), assessing learning (CONF3), and preparing for classes (CONF4) (see Appendix A). Section 2 remained unchanged, and Section 3 expanded to 12 items by incorporating questions about respondents' EMI training, university affiliation, and institutional EMI policies.

The questionnaire was designed to be administered in English for several reasons. First, as the intended targeted survey respondents are EMI subject-content teachers who have been using English to teach in English-taught courses or programs, the general levels of English proficiency of potential participants are expected to be sufficient to fully understand all of the questionnaire items. Next, the questionnaire was developed using common high-frequency vocabulary and simple Likert-type scales to ensure that participants can easily comprehend each item and respond without difficulty. Additionally, no language comprehension issues were either observed or reported during piloting of the questionnaire. Moreover, translating the questionnaire in different languages for specific EMI contexts would require developing several different versions, which would negate the practicality of the instrument (that it can be easily used in a wide variety of contexts) as well as decrease the generalizability of findings across different EMI contexts due to potential language differences. Previous studies demonstrate that using translated instruments can result in information or meaning being inadvertently altered, diminished, or simply lost in translation (Chang et al., 1999; Chapman & Carter, 1979; McGorry, 2000; Yu et al., 2004). Harkness et al. (2004) contend that "small differences in formulation across languages can affect understanding" and that translation "can result in respondents not being asked what the researchers intended to ask" (p.454).

In its current form (see Supplementary Material for the full instrument), TECS consists of two main sections and a final section containing items eliciting demographic and background information. Section 1 includes four items, while Section 2 comprises 28 items (three items were eliminated following CFA results); four additional open-ended questions in Section 2 are optional. Both sections use seven-point Likert-type scales, commonly employed in social science research (Croasmun & Ostrom, 2011) and increasingly used in EMI studies (Curle & Derakhshan, 2021). As scale format can affect response distribution (Weijters et al., 2010), and response categories and labels can influence levels of response bias (Greenleaf, 1992), particular consideration was given to the various options.

The first part of the survey focuses on lecturers' self-efficacy beliefs regarding their ability to use English in EMI teaching contexts. Respondents are asked to rate their confidence in using English to (i) teach (e.g., deliver lectures and seminars), (ii) interact with students (e.g., consultations and discussions), (iii) assess learning (e.g., set assignments, mark, and provide feedback), and (iv) prepare for teaching (e.g., select readings and create lecture materials). In terms of orientation and directionality, a vertical descending format is used with a seven-point scale ranging from *Highly confident* to *Highly unconfident,* which includes a mid-point, or 'neutral' point labelled *Neither confident nor unconfident.*

We chose a seven-point Likert-type scale for three main reasons. First, research suggests that seven-point scales are preferred for unmonitored online questionnaires and provide "a more accurate measure of a participant's true evaluation" (Finstad, 2010, p. 104). Additionally, seven-point scales can reveal stronger correlations between mean differences and significance levels (Lewis, 1993). Thirdly, data from Likert-type scales can be considerably less accurate when scale points are fewer than five or more than seven (Johns, 2010). In terms of response-order effect, we opted for a decremental order (versus incremental) which begins with Highly confident and ends with Highly unconfident. This reflects a slightly more conservative approach to reduce potential cases of inflated or

exaggerated lack of confidence due to primacy effect, the tendency by respondents to select options which are presented earlier in a list, which is more common in online surveys.

The second section of the survey focuses on language-related competencies and asks respondents to indicate how easy or difficult they find performing specific communicative tasks in English across speaking, writing, listening, and reading. Each of the four skill sections begins with a brief instructional prompt ("Think of yourself in a situation where you are speaking /writing /listening /reading English for some teaching purpose"). Lecturers then rate 6–8 tasks per skill, using a seven-point Likert scale ranging from Very easy to Very difficult. A vertical format with a decremental order was used again in the second part in an effort to mitigate any potential primacy effect which might overemphasize respondents' teaching challenges. Hence, any primacy effect would tend to lean towards 'easy' rather than 'difficult' ensuring a more conservative approach to measuring actual challenges.

Section 3 contains 12 demographic and background items, including disciplinary field, university affiliation, student levels taught, years of (general and EMI) teaching experience, CEFR-aligned English proficiency, language background, gender, prior EMI training, and institutional EMI policies. These items support subgroup analyses and contextualize lecturers' self-assessments.

The survey is fully anonymous and takes approximately ten minutes to complete. To date, TECS has been administered online, using Microsoft Forms, though the instrument is platform-agnostic (if desirable, a paper-and-pencil administration is also possible).

3.3. Instrument piloting

In March and April 2024, versions of TECS were piloted twice with EMI lecturers at the first author's university to evaluate its clarity, coverage, and relevance to EMI teaching contexts. The average completion time for the pilot survey was approximately 11 min. Each pilot administration was followed by informal (whole) group interviews with the participants in order to obtain qualitative feedback on the design, terminology, and overall usability of the instrument, leading to several rounds of refinement prior to large-scale deployment.

Feedback from the first pilot ($n = 5$) indicated that nine items required clarification, particularly in relation to ambiguous or otherwise unclear phrasing and the consistency of response options. Accordingly, several linguistic and structural revisions were made between the two administrations. These included minor rewordings to improve the clarity of prompts, removal of potentially ambiguous wording, and refinement of the introductory guidance to make the purpose of the instrument more explicit. Before the second pilot administration, the introductory section of the instrument was also expanded to include necessary information regarding research ethics and informed consent, ensuring full transparency about the voluntary nature of participation and the treatment of survey data.

To strengthen the instrument's face validity, we also removed or replaced items that either pilot respondents – or we ourselves – judged not to be central or specific to EMI classroom interaction. This process resulted in a more focused and coherent set of prompts, better aligned with the communicative and pedagogic realities of EMI teaching.

A second round of lecturer piloting ($n = 7$) confirmed the overall clarity, coherence, and contextual fit of the revised instrument: participants reported that the survey items were well formulated, appropriately framed, and reflective of their everyday communicative practices in EMI classrooms. The average response time for completing the survey was now just under nine minutes, with some respondents reporting completion times of between 5 and 6 min. Minor layout and formatting issues were also identified and resolved prior to large-scale deployment. From the outset, the instrument also included a set of demographic items to enable subgroup analyses and interpretation of responses; these demographic items were only minimally revised during the two pilot stages.

3.4. Selecting study contexts for instrument trial

To trial TECS we selected two contrasting EMI contexts: Sweden and Spain. Both have seen substantial expansion of EMI over time (Wingrove et al., 2025), but they differ in national English proficiency, language policy traditions, and the degree to which English is embedded in society and academia, as briefly described below.

3.4.1. Sweden

Sweden represents a high-proficiency and highly internationalized EMI context. English is widely used in Swedish society, including in workplaces, media, and popular culture, and students are generally exposed to the language from an early age. By the end of upper secondary school, they are expected to reach at least CEFR B2, and many exceed this benchmark (Mežek, 2024). Moreover, Sweden consistently ranks among the top-performing countries in global English proficiency comparisons (EF Education First, 2024). This strong societal and educational foundation is reflected in higher education, where internationalization has made English the default academic lingua franca. Thus, while Swedish remains the legally mandated main language of virtually all universities, English plays a central role in teaching and research, particularly at master's level: around 18% of undergraduate courses and more than half of master's are taught in English, and approximately two-thirds of master's programs are entirely English-medium (Malmström & Pecorari, 2022). However, stakeholders' English proficiency is far from uniform. Research shows considerable variation in EMI lecturers' productive and receptive English competencies. Some lecturers report limitations that likely affect classroom interaction (Malmström et al., 2025), whereas others, especially in science, technology, engineering and mathematics, self-assess at CEFR C1–C2 and report high confidence in their English ability (Vander Borght et al., 2025).

3.4.2. Spain

Compared to Sweden, Spain represents a developing EMI context, where English has gained prominence more recently but is less

embedded in society and academia. Spain typically ranks below the European average in English proficiency (EF Education First, 2024), although English is increasingly becoming more visible in the media, tourism, professional sectors, and youth culture (Llurda & Mocanu, 2024). In education, major policy efforts have supported the expansion of bilingual (English/Spanish) schools across the country and in 1996, the Ministry of Education and the British Council launched the Bilingual Education Program (BEP) to integrate English into public schooling. From the early 2000s, regional initiatives have also extended bilingual instruction across primary and secondary education, particularly in subjects such as science, history, and the arts (Lasagabaster & Ruiz de Zarobe, 2010).

Higher education in Spain has followed suit, with universities increasingly adopting EMI to attract international students, strengthen research collaboration, and improve home students' graduate competitiveness. This agenda was recently formalized in the university act Ley Orgánica 2/2023 (LOSU), which explicitly promotes EMI as part of national internationalization strategies while ensuring the continued presence of Spanish and co-official languages. Nonetheless, implementation challenges persist as many lecturers report uncertainty about teaching through English and highlight the need for further linguistic and pedagogical support. Spain therefore offers a context where EMI is expanding rapidly (Fernández-Costales & Lasagabaster-Herrarte, 2024; Wingrove et al., 2025) yet lecturer preparedness remains a central concern, making it a particularly relevant setting for validating an instrument focused on classroom English use.

3.5. Participant recruitment

A total of 533 EMI lecturers participated in the first TECS administration: 347 from Sweden (in May 2024) and 186 from Spain (February-May 2025). Participants were recruited through a combination of purposeful and snowball sampling. Purposeful sampling ensured that invitations reached lecturers with experience of EMI, while snowball sampling encouraged participants to forward the invitation link to EMI colleagues in similar roles. An open invitation was distributed via e-mail to EMI lecturers through a dedicated list server established specifically for this purpose. The invitation was sent to lecturers at five universities in Sweden and three universities in Spain. The message briefly outlined the study's aims, provided a link to the online instrument, and explained that participation was voluntary, anonymous, and conditional on informed consent, consistent with the ethical procedures established during piloting. Recipients were encouraged to forward the invitation to other EMI lecturers in Sweden or Spain. Because the link could be shared, the relative contribution of purposeful and snowball recruitment cannot be determined. Participants received no compensation.

Overall, the recruitment strategy was designed to reach a cross-section of EMI practitioners across disciplinary fields, institutional types, and levels of experience. Respondents thus represented a diverse group of higher-education lecturers engaged in English-medium programs in both countries, spanning engineering, natural sciences, social sciences, health sciences, and the humanities.

We acknowledge that the validation sample is unevenly distributed across contexts, with a larger proportion of respondents from Sweden than from Spain (see below). While this does not affect the estimation of the factor structure, which is based on covariance patterns among items, it limits the extent to which cross-context equivalence can be assumed. With this in mind, researchers who seek to use the instrument with different EMI teacher populations to ours, may benefit from replicating our validation procedures to verify factor structures in their own research context.

4. Validating TECS

For the EMI language competency items, the data were randomly split into two sets using the random case selection function in SPSS. This procedure was used to create independent subsamples for exploratory and confirmatory factor analyses. The split was conducted across the full dataset using simple random sampling rather than stratified by variables such as gender country, discipline, or prior experience to avoid introducing bias, and because the purpose of these analyses was to identify the underlying factor structure of the instrument rather than to compare national groups. This resulted in 265 cases for use for the exploratory factor analysis (EFA)

Table 1
Participant characteristics.

Variable	Category	Sweden	Spain	Total
Allocation	EFA	165	100	265
	CFA	182	86	268
Discipline	Engineering	34	36	70
	Health sciences	3	24	27
	Humanities	59	8	67
	Medicine	1	1	2
	Natural sciences	3	39	42
	Social sciences	242	78	320
	Not reported	5	0	5
Gender identity	Male	185	89	274
	Female	151	90	241
	In another way*	7	7	14
Total	All cases	347	186	533
	After case removal	307	182	489

Note: *The gender identity question asked "Which of the following describes how you think of yourself?", following guidance from the Equality and Human Rights Commission (Balarajan et al., 2011).

and 268 cases for use in the confirmatory factor analysis (CFA) (see Table 1 for participant characteristics). Next, data were inspected for normality and whether they met the assumptions of factor analysis. Subsequently, 44 participants were removed before factor analysis, because they recorded scores of 7 out of 7 for all items across the entire scale. These responses contributed no variance and thus would provide no information for the factor structure. In EFA and CFA, covariance and variance among items are what drive factor extraction and model estimation, so including invariant cases (all responses of 7 across all items) would distort correlations and lead to estimation problems (e.g., singular covariance matrices or underestimated factor loadings). Therefore, these participants were removed before analysis. This resulted in 246 cases for use for the EFA and 243 cases for use in the CFA.

Regarding our measure of EMI lecturer self-efficacy, this comprised four newly created items for this study. Because a four-item scale is too small for a meaningful exploratory factor analysis, we assessed its structure using a confirmatory approach. A single-factor CFA with four indicators was specified to evaluate the coherence of the scale. The sample used for this was the total 489 cases. As is typical for a one-factor model with four items, the model is just-identified and therefore global fit indices are not informative; validation is based on factor loadings and reliability. To ensure research transparency, these data are available on GitHub (https://github.com/drheathrose/TECS.git).

4.1. Exploratory factor analysis

A series of EFAs were conducted on the EFA dataset ($n = 246$) in SPSS to examine the underlying theoretical structure and dimensionality of the 31-item instrument. Because the items measured latent constructs and were moderately correlated, principal axis factoring with oblimin rotation was employed. Sampling adequacy for factor analysis was excellent: Kaiser–Meyer–Olkin (KMO) = 0.964, and Bartlett's Test of Sphericity was significant, $\chi^2(406) = 8359.42$, $p < .001$, supporting factorability. Two items were removed

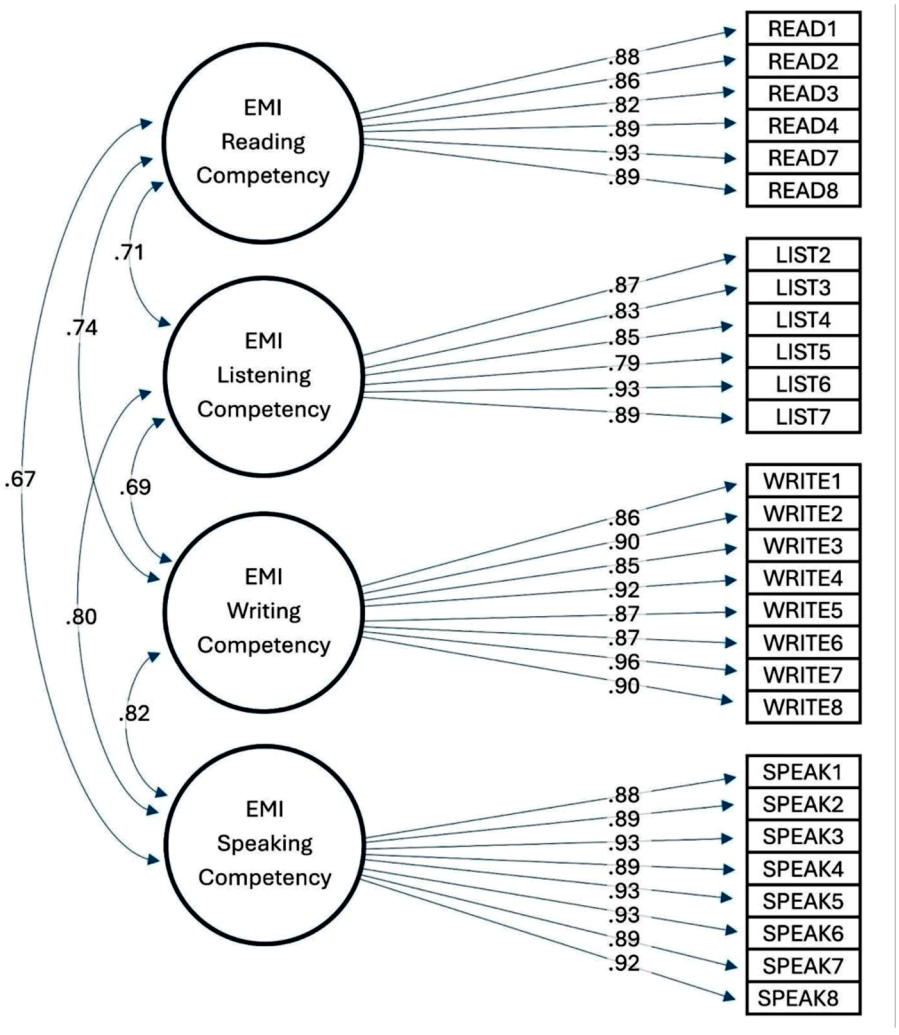


Fig. 1. Confirmatory factor analysis model of EMI competencies (standardized factor loadings for indicators and inter-factor correlations are presented).

after the initial EFA based on poor loading (< 0.40) and/or salient cross-loadings (> 0.30), which hindered theoretical clarity. These items were related to EMI reading competency (READ5 and READ6).

A final EFA was run on the remaining items and can be found in Appendix B. Communalities for the remaining items were high (0.61–.86), and factor loadings were consistently strong (0.63–.87), indicating that each latent dimension was well represented by its indicators. Inspection of eigenvalues (> 1), the scree plot, and interpretability criteria suggested a four-factor solution, accounting for 75.6% of the total variance after extraction for a 29-item instrument. These four factors corresponded conceptually to (1) EMI speaking competency (SPEAK1–SPEAK8), (2) EMI writing competency (WRITE1–WRITE8), (3) EMI listening competency (LIST1–LIST7), and (4) EMI reading competency (READ1–READ4, READ7, READ8). Factor correlations ranged from 0.55 to 0.75, indicating related but distinct dimensions, consistent with theoretical assumptions about productive vs. receptive academic abilities. Overall, the EFA supported a coherent and interpretable four-factor structure aligned closely with the hypothesized academic English competency constructs.

4.2. Confirmatory factor analysis

A confirmatory factor analysis (CFA) was subsequently conducted in R using the second dataset ($N = 243$) to validate the four-factor measurement model identified in the EFA. Because the items were ordinal, the model was estimated using the WLSMV estimator in lavaan (R), with latent variables standardized (std.lv = TRUE). The CFA was conducted in R using the lavaan package because it allows estimation with the WLSMV estimator, which is recommended for ordinal Likert-type indicators. This estimator is not readily available in SPSS factor analysis procedures.

Inspection of modification indices and standardized residuals identified LIST1 as a poorly performing indicator of the listening comprehension factor due to relatively lower item-level explanatory power and local misfit; therefore, this item was removed. The final model included 30 items loading onto four correlated latent variables and is shown in Fig. 1.

The refined model demonstrated excellent global fit to the data: $\chi^2(344) = 564.75$, $p < .001$, CFI = 0.998, TLI = 0.998, RMSEA = 0.051 (90% CI [.044, 0.059]), and SRMR = 0.043. According to conventional cut-offs (Hu & Bentler, 1999), these values indicate a strong correspondence between the hypothesized model and the observed data. All standardized factor loadings were strong and statistically significant (0.82–.96, all $p < .001$), indicating that each item contributed meaningfully to its latent dimension. Squared multiple correlations (R^2) showed that between 0.62 and 0.92 of the variance in each indicator was explained by its corresponding factor.

Model-based reliability (construct reliability exceeded recommended thresholds (≥ 0.70) for each factor: EMI speaking competency (Factor 1) 0.975; EMI writing competency (Factor 2) = 0.973; EMI listening competency (Factor 3) = 0.932; EMI reading competency (Factor 4) = 0.944 Average Variance Extracted (AVE) also exceeded the 0.50 benchmark, supporting convergent validity: Factor 1 = 0.824; Factor 2 = 0.794; Factor 3 = 0.739, Factor 4 = 0.773. Latent factor correlations ranged from 0.67 to 0.82, indicating substantial yet distinguishable relationships between competencies. Importantly, the AVE for each construct exceeded the squared inter-factor correlations, satisfying the Fornell–Larcker criterion for discriminant validity (Fornell & Larcker, 1981), which indicates that each latent construct shares more variance with its indicators than with other constructs in the model.

Via this two-stage analysis, the EFA and CFA procedures yielded a robust, theoretically coherent four-factor structure representing productive (speaking, writing) and receptive (listening, reading) competencies required for English-medium instructional communication. The removal of two weak EFA items (READ5 and READ6) and one weak CFA item (LIST1) improved parsimony and model fit without compromising conceptual integrity. All retained indicators demonstrated high statistical performance, and the final CFA model exhibited excellent global fit, strong loadings, and high reliability.

4.3. EMI lecturer self-efficacy scale validation

A confirmatory factor analysis (CFA) was conducted on the four teacher self-efficacy items, which explored teacher confidence in delivering teaching (CONF1), interacting with students (CONF2), assessing learning (CONF3), and preparing for classes (CONF4) (see Appendix A), specifying a single latent factor. Standardized loadings were very strong ($\lambda = 0.85$–.97), indicating that all items contributed substantially to the underlying construct. Given the ordinal response format, the model was estimated using DWLS, and model evaluation focused on loadings and reliability rather than global fit indices. Internal consistency was excellent ($\alpha = 0.92$; ordinal $\alpha = 0.95$; $\omega = 0.93$), and the average variance extracted (AVE = 0.83) also indicated strong convergent validity. The four items were averaged to form a composite self-efficacy score used as a single variable in subsequent analyses.

4.4. Scale validation summary

The resulting instrument provides a psychometrically sound measurement framework suitable for subsequent structural modelling, group comparisons, or institutional diagnostics of EMI-related academic English abilities. Together, these findings provide strong empirical support for the validity and reliability of the proposed four-factor measurement model for the *EMI competencies scale* and a one-factor model for the *EMI lecturer self-efficacy scale* and justify its use in subsequent analyses of English language demands on lecturers working within EMI contexts.

5. TECS in Sweden and Spain

To demonstrate the practical utility of the validated instrument, the factor scores derived from the final CFA model were used to examine group differences between EMI lecturers in Spain and in Sweden. This step illustrates how the instrument can be applied in analyses beyond the validation process, providing evidence of its sensitivity to analyze distinctions in the data. Demonstrating such group-based comparisons supports the construct validity of the measure (i.e., whether it behaves as expected across relevant populations) and highlights its potential for use in future empirical research employing the same constructs. It should be noted that these differences relate to mean levels of reported ease rather than to the measurement structure of the instrument. The factor structure of TECS was established through exploratory and confirmatory factor analyses based on covariance patterns among items across the full dataset. For this exercise, we used the total dataset with the exception of the cases at ceiling, which were removed in the validation process. As a first exploration, mean scores were computed for the cases from Spain (SPA, $n = 182$) and Sweden (SWE, $n = 307$) on four factors and self-efficacy. Across all five areas, Swedish EMI lecturers reported significantly higher levels of ease with EMI tasks than Spanish EMI lecturers (Table 2).

A one-way MANOVA was conducted to examine differences between Spain and Sweden on the five composite variables: perceived ease with listening, writing, reading, and speaking in EMI, and EMI self-efficacy. The multivariate test was significant, Pillai's Trace = 0.12, $F(5, 483) = 13.37$, $p < .001$, indicating overall group differences across the five dimensions. Follow-up univariate analyses showed significant effects of country for all variables (all $p < .001$), with effect sizes ranging from small to large (partial $\eta^2 = 0.03$–.12). As shown in Table 1, lecturers in Sweden reported higher mean scores than lecturers in Spain on listening, writing, reading, speaking, and self-efficacy. The largest differences were observed for listening and speaking, followed by smaller but still meaningful differences in writing, reading, and self-efficacy. It is noteworthy that this significance remained even with our conservative decision to remove the cases at ceiling; 42 of the 44 removed cases were from Sweden, meaning that the difference in *EMI competencies* would have been even more pronounced had the total dataset been used.

In order to further investigate whether these differences were country differences, or another variable such as English proficiency (i.e., is the difference simply because EMI lecturers in Sweden are more proficient in English) further analysis was performed. A hierarchical multiple regression was conducted to examine whether lecturers' country of employment (Sweden vs. Spain) predicted their perceived ease of speaking in English, after controlling for self-rated English proficiency. The independent (predictor) variables of country and proficiency were included primarily to contextualize respondents and to enable subgroup analyses on the outcome (dependent) variable of EMI speaking competency. Because the present study focuses on instrument validation, only this limited set of illustrative predictors were used in the regression analyses. In Model 1, proficiency was entered first and accounted for a substantial proportion of variance in speaking ease, $R^2 = 0.397$, $F(1, 459) = 301.78$, $p < .001$. Higher proficiency strongly predicted greater ease with speaking ($B = 0.97$, $SE = 0.06$, $\beta = 0.63$, $p < .001$). In Model 2, the variable of country was added to the model. The addition of this variable produced a small but statistically significant improvement in model fit, $\Delta R^2 = 0.012$, F change $(1, 458) = 9.13$, $p = .003$. Both predictors remained significant: proficiency ($B = 0.91$, $SE = 0.06$, $\beta = 0.59$, $p < .001$) and country ($B = 0.29$, $SE = 0.09$, $\beta = 0.12$, $p = .003$). Thus, even when controlling for proficiency, lecturers in Sweden reported slightly greater ease in speaking in English than those in Spain. Overall, the final model accounted for 40.8% of the variance in perceived speaking ease ($R^2 = 0.408$, adjusted $R^2 = 0.406$), indicating that both individual proficiency and country context contribute to how easily lecturers report being able to use English for speaking.

The same analysis was conducted for the other three competency areas and is summarized in Table 3. Proficiency predicted *EMI writing competency* scores ($R^2 = 0.314$). Adding *country* did not improve the model ($\Delta R^2 = 0.001$, $p = .506$). This suggests that any observed differences between Spain and Sweden in *EMI Writing competency* was already explained by proficiency, with no additional contribution from *country*. Proficiency was associated with *EMI listening competency scores* ($R^2 = 0.286$). Including *country* led to a clearer improvement in model fit ($\Delta R^2 = 0.047$, $p < .001$). Thus, after accounting for proficiency, lecturers in Sweden tended to report significantly higher *EMI listening competency* than those in Spain. Although still modest, this was the largest contribution of the variable *country* among the four competencies. Proficiency also predicted *EMI reading competency* ($R^2 = 0.206$. Adding *country* produced a small but statistically significant improvement in the model ($\Delta R^2 = 0.008$, $p = .028$). This indicates that, once proficiency was controlled, Swedish lecturers still reported significantly higher *EMI reading competency* than Spanish lecturers.

These results show that proficiency is a powerful and significant predictor of all *EMI competencies* for lecturers in this dataset. It also demonstrates that, even when proficiency was controlled for, country differences observed in the initial analysis remained in three of the four EMI competencies. With the statistical results explained, the following section considers the implications and potential

Table 2
Descriptive statistics and effect sizes for Spain and Sweden across the five composite variables.

Variable	Spain *M* (SD)	Sweden *M* (SD)	Mean difference	Partial η^2
EMI listening competency	5.30 (1.10)	6.04 (0.91)	0.74	.115
EMI writing competency	5.47 (1.08)	5.87 (1.08)	0.40	.031
EMI reading competency	5.88 (0.90)	6.26 (0.83)	0.38	.044
EMI speaking competency	5.05 (1.16)	5.72 (1.21)	0.67	.069
EMI lecturer self-efficacy	5.84 (0.92)	6.26 (0.94)	0.42	.044

Note. Means represent composite scores averaged across items for each construct. Mean difference = Sweden − Spain. Partial η^2 values are from the univariate follow-up tests after the significant multivariate effect (Pillai's Trace = 0.12, $p < .001$). $N = 489$ (Spain = 182, Sweden = 307.

Table 3
Model summaries of hierarchical regressions.

Outcome Variable	Model	R^2	Adj. R2	ΔR^2	ΔR^2 *p*
EMI Speaking Competency	M1 Prof	.397	.397	.395	<0.001
	M2 +country	.408	.406	.012	.003
EMI Writing Competency	M1 Prof	.314	.313	314	<0.001
	M2 +country	.315	.312	.001	.506
EMI Listening Competency	M1 Prof	.286	.284	.236	<0.001
	M2 +country	.333	.330	.047	<0.001
EMI Reading Competency	M1 Prof	.206	.204	.206	<0.001
	M2 +country	.214	.211	.008	.028

applications of the TECS scale.

6. Implications and applications of TECS

The illustrative comparison between Sweden and Spain provides a first indication of how TECS can capture variation across EMI contexts. Lecturers in Sweden reported higher perceived ease across all communicative domains, while English proficiency emerged as a strong predictor of EMI competencies. These patterns are broadly consistent with previous research suggesting that lecturers' perceived linguistic competence and teaching self-efficacy play an important role in how confidently they manage instructional communication in English (e.g., Lasagabaster, 2022; Wang, 2021). However, the persistence of some country differences even after controlling for proficiency suggests that contextual factors, such as institutional support for EMI or the broader sociolinguistic environment, may also influence lecturers' experiences of teaching through English. By and large, these findings from our administration in Sweden and Spain indicate that TECS can capture variation associated not only with lecturers' perceived competence but also with the contexts in which EMI is practiced. This responds to calls for comparative methodologies (Macaro, 2018; Molino et al., 2022) and provides a foundation for considering how TECS may be applied more broadly in both research and practice.

From a professional development perspective, TECS could arguably be used as a diagnostic tool. By focusing on classroom communication rather than on general linguistic proficiency, it can identify the specific aspects of teaching-in-English that lecturers may find challenging and provide evidence to guide targeted professional development (Dafouz & Gray, 2022). Administered before and after EMI training, it can help evaluate the impact of such initiatives and show whether lecturers' perceived ease in key communicative areas improves as a result of training. In this way, TECS supports the design of practice-oriented professional development rather than generic language courses.

Another pedagogical use of TECS is to examine how lecturers' perceptions align with their observed classroom behavior. TECS data can be combined with classroom observations, video-based analyses, or student evaluations to explore how lecturers' self-perceived competencies correspond to their actual communicative behavior in EMI settings. Such studies would provide valuable insights into how lecturers' self-efficacy relates to enacted classroom practice and could also inform the refinement of EMI training through evidence of where perceptions and performance may diverge.

From a cross-contextual angle, the instrument also lends itself to comparative research, enabling systematic comparison of EMI lecturers' perceived communicative competence across not only countries, but also across types of institutions (private vs. public), disciplinary areas and university educational levels. As it can be administered in multiple contexts, TECS facilitates the development of cumulative and comparable datasets that can support comparative and meta-analytic work and contribute to more robust cross-context understandings of EMI lecturers' classroom-language practices in EMI contexts (see e.g., Li & Wu, 2025 for the Taiwanese case)

Finally, TECS opens up opportunities for predictive and modelling applications. Given its multidimensional and reliable factor structure, it can be used in multivariate and structural-equation modelling to examine how different aspects of communicative competence and self-efficacy predict other constructs of interest, such as pedagogical choices, interactional patterns, or student engagement. Multilevel models, as pointed above, could incorporate contextual variables (e.g., discipline, institutional policy, level of English exposure) to explore how individual and institutional factors jointly shape lecturers' experience of EMI. In this way, TECS can contribute not only to descriptive mapping but also to theory-building about the mechanisms linking language, confidence, and pedagogy in EMI.

While these applications point to a wide range of possible uses, it is also important to note certain boundaries for TECS' use. One limitation concerns the potential ceiling effect observed in the dataset. A number of participants, primarily from the Swedish sample, selected the highest response option for every item on the scale and also reported C2-level English proficiency. This suggests that the TECS items may have limited discriminatory power among lecturers who report consistently very high levels of ease using English for teaching. Future refinements of the instrument could therefore consider incorporating more fine-grained items capable of capturing variation at higher levels of EMI communicative competence.

7. Concluding remarks

This paper has presented the development and validation of the Teaching-in-English-competencies-scale (TECS), an instrument designed to examine EMI lecturers' perceived communicative confidence, competence, and preparedness when teaching subject content through English. Exploratory and confirmatory factor analyses yielded a robust four-factor structure representing speaking,

writing, listening, and reading competencies, alongside a single-factor measure of teaching self-efficacy. The indicators demonstrated strong statistical performance, supporting the reliability and construct validity of the instrument and confirming its suitability for applications such as group comparisons, institutional diagnostics, and structural modelling of EMI-related teaching competencies.

Although the empirical analyses reported here primarily served to support instrument validation, they also illustrate how TECS can capture variation in lecturers' perceived communicative competence beyond a single EMI context. More broadly, the scale provides a systematic framework for examining how lecturers perceive their ability to use English as a medium for disciplinary teaching and classroom interaction, responding to calls for standardized instruments capable of generating comparable data across EMI settings.

Several limitations of the present study should be noted. First, the sample is unevenly distributed across national contexts, with a larger proportion of respondents from Sweden, which limits the extent to which cross-context generalizations can be made. Second, the gender composition of the sample is not balanced, which may influence response patterns in self-report measures of perceived competence and self-efficacy. Third, although the factor structure of TECS was established across the full dataset, the study does not test whether the instrument operates equivalently across groups (e.g., gender or country). These considerations highlight the need for further validation work before stronger claims about cross-context or cross-group comparability can be made.

Future research should extend the validation of TECS beyond the internal psychometric properties examined here. In particular, studies could investigate how TECS scores relate to other theoretically relevant constructs, such as teacher psychology, professional practices, or pedagogical approaches in EMI. Examining these relationships would help establish the external validity of the instrument and clarify how perceived communicative competence interacts with other dimensions of lecturers' professional experience. In addition, future studies should examine measurement invariance across demographic and contextual variables and apply the instrument in a wider range of institutional and national settings.

In sum, TECS provides stakeholders with a systematic tool to assess and examine lecturers' perceived English language competencies and the factors that shape their teaching in EMI settings. As such, we invite its use in future studies and professional development initiatives aimed at enhancing EMI quality.

Declaration of AI use

During the preparation of this manuscript, the authors made occasional use of OpenAI's ChatGPT to support minor language editing and to ensure stylistic consistency across contributions from multiple authors. All material was subsequently reviewed, edited and approved by the authors, who assume full responsibility for the final article.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT authorship contribution statement

Hans Malmström: Writing – review & editing, Writing – original draft, Project administration, Methodology, Data curation, Conceptualization. **Heath Rose:** Writing – review & editing, Writing – original draft, Methodology, Formal analysis, Data curation, Conceptualization. **Christopher Shepard:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization. **Emma Dafouz:** Writing – review & editing, Writing – original draft, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors gratefully acknowledge the assistance of members of the ICLHE regional group Spain and lecturers in Sweden in Spain who kindly gave their time to complete the questionnaire..

Supplementary materials

Supplementary material associated with this article can be found in the online version, at doi:10.1016/j.rmal.2026.100326.

Appendix A. Full items and codes used in tables and Figures

EMI teaching self-efficacy:

CONF1: … for teaching purposes (e.g., delivering lectures and seminars).
CONF2: … to interact with students (e.g., student consultations and engaging in discussions on course content).
CONF3: …for assessment purposes (e.g., setting assignments, marking, providing feedback).

CONF4: …to prepare for classes (e.g., selecting appropriate reading materials, creating lecture materials).

EMI teaching language competency:
All items lead with: *How easy/difficult is it for you to [statement]?*

SPEAK1: …give verbal instructions to students (e.g., relating to a learning task)?
SPEAK2. …present disciplinary content in a lecture or seminar?
SPEAK3. …develop an argument/teaching point?
SPEAK4: …interact with students in class, e.g., ask or answer questions?
SPEAK5: …speak fluently when teaching (i.e., speak easily, reasonably quickly and without having to stop and pause a lot)?
SPEAK6: …speak accurately (grammatically correct) when teaching?
SPEAK7: …use pronunciation effectively when teaching (i.e., articulate words clearly, and being mindful of word and sentence stress, and intonation)?
SPEAK8: …use appropriate/fit-for-purpose vocabulary when teaching?
WRITE1: …write general (administrative) information (e.g., relating to the course syllabus, or assignment instructions) in writing?
WRITE2: …correspond with students (e.g., via e-mail or messages through a learningmanagement platform)?
WRITE3: …write about disciplinary content (e.g., drafting learning material that you expect the students to read and learn from)?
WRITE4: …provide written feedback on students' assignments?
WRITE5: …write text in a clear, structured, and logical manner?
WRITE6: …write text in grammatically correct English?
WRITE7: …use appropriate/fit-for-purpose vocabulary when writing?
WRITE8: …use appropriate writing style/tone when writing?
LIST2: …understand a classroom discussion?
LIST3: …understand students' (formal) oral presentations about disciplinary content?
LIST4: …understand recorded and broadcast material (e.g., when choosing supplementary learning resources on YouTube)?
LIST5: …understand other speakers' English accents (varieties of English)?
LIST6: …understand specific words used in oral interactions?
LIST7: …understand a specialised lecture given by a teaching colleague from your discipline?
READ1: …understand different types of texts when doing background reading, such as specialized academic (scholarly) or professional publications.
READ2: …read for the purposes of identifying and selecting reading materials for students?
READ3: …"scan" a text by looking through it quickly, searching for specific information?
READ4: …draw conclusions, make predictions, and understand implied information based on the content of a text?
READ7: …understand specific vocabulary used in texts?
READ8: …read with accuracy at an adequate pace?

Items removed in validation process (not to be used in future applications of TECS):
LIST1*: understand questions and/or comments from students in class? (*Item removed in CFA)
READ5*: read students' assignments and assessments? (*Item removed in EFA)
READ6*: read correspondence - e.g., e-mails - from students (or lecturer colleagues)?
(*Item removed in EFA).

Appendix B

Factor loading and structure from final EFA.

	Factor			
	1	2	3	4
SPEAK1	.562	.177	.021	.193
SPEAK2	.520	.079	.038	.313
SPEAK3	.701	.034	.132	.077
SPEAK4	.561	.262	−.052	.144
SPEAK5	.813	.068	−.010	.056
SPEAK6	.804	.029	.060	.084
SPEAK7	.904	.085	.011	−.074
SPEAK8	.757	−.054	.119	.111
WRITE1	−.065	.132	−.047	.895
WRITE2	.070	.074	.005	.789
WRITE3	−.034	.012	.020	.935

(*continued on next page*)

(*continued*)

	Factor			
	1	2	3	4
WRITE4	−.037	.093	−.002	.878
WRITE5	−.029	−.041	.149	.787
WRITE6	.237	−.109	.040	.730
WRITE7	.230	−.080	.152	.640
WRITE8	.227	−.009	.088	.615
LIST1	−.091	.940	−.036	.083
LIST2	.064	.888	−.057	.048
LIST3	.001	.832	.019	.076
LIST4	.123	.635	.196	−.042
LIST5	.143	.656	.053	−.014
LIST6	.297	.482	.202	.002
LIST7	.057	.605	.271	−.037
READ1	−.087	.058	.857	−.019
READ2	−.083	.027	.826	.085
READ3	.102	.010	.713	.040
READ4	.036	−.008	.826	.048
READ7	.179	.106	.604	.052
READ8	.226	.035	.578	.067

References

Abdeljaoued, M. (2023). English-medium instruction in Tunisia: Perspectives of students. *Frontiers in Psychology, 14*, Article 1112255. https://doi.org/10.3389/fpsyg.2023.1112255

Agnew, M., & Neghina, C. (2021). *The changing landscape of English-taught programmes*. British Council & Studyportals.

Aguilar, M., & Rodríguez, R. (2012). Lecturer and student perceptions on CLIL at a Spanish university. *International Journal of Bilingual Education and Bilingualism, 15*(2), 183–197. https://doi.org/10.1080/13670050.2011.615906

Airey, J. (2011). Talking about teaching in English: Swedish university lecturers' experiences of changing teaching language. *Ibérica, 22*, 35–54. https://revistaiberica.org/index.php/iberica/article/view/317.

Balarajan, M., Gray, M., & Mitchell, M. (2011). Monitoring equality: Developing a gender identity question. *London: Equality and Human Rights Commission*.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Chang, A. M., Chau, J. P., & Holroyd, E. (1999). Translation of questionnaires and issues of equivalence. *Journal of Advanced Nursing, 29*(2), 316–322. https://doi.org/10.1046/j.1365-2648.1999.00891.x

Chapman, D. W., & Carter, J. F. (1979). Translation procedures for the cross cultural use of measurement instruments. *Educational Evaluation and Policy Analysis, 1*(3), 71–76. https://doi.org/10.3102/01623737001003071

Council of Europe. (2023). *Common European framework of reference for languages: Learning, teaching, assessment. companion volume with new descriptors*. Council of Europe Publishing. https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989.

Croasmun, J. T., & Ostrom, L. (2011). Using Likert-type scales in the social sciences. *Journal of Adult Education, 40*(1), 19–22.

Curle, S. M., & Derakhshan, A. (2021). Trends in using questionnaires for EMI research: Suggestions for future improvements. In J. K. H. Pun, & S. M. Curle (Eds.), *Research methods in English medium instruction* (pp. 32–45). Routledge.

Dafouz, E. (2018). English-medium instruction and lecturer education programmes in higher education: Ideological forces and imagined identities at work. *International Journal of Bilingual Education and Bilingualism, 21*(5), 540–552. https://doi.org/10.1080/13670050.2018.1487926

Dafouz, E., & Gray, J. (2022). Rethinking the roles of ELT in English-medium education in multilingual university settings: An introduction. *ELT Journal, 76*(2), 163–171. https://doi.org/10.1093/elt/ccab096

Dang, T. K. A., Bonar, G., & Yao, J. (2023). Professional learning for educators teaching in English-medium-instruction in higher education: A systematic review. *Teaching in Higher Education, 28*(4), 840–858. https://doi.org/10.1080/13562517.2020.1863350

Dearden, J., & Macaro, E. (2016). Higher education lecturers' attitudes towards English medium instruction: A three-country comparison. *Studies in Second Language Learning and Teaching, 6*(3), 455–486.

Dimova, S., Hultgren, A. K., & Jensen, C. (2015). English-medium instruction in European higher education: Review and future research. In S. Dimova, A. K. Hultgren, & C. Jensen (Eds.), *English-Medium instruction in European higher education* (pp. 317–324). De Gruyter Mouton.

EF Education First. (2024). *EF english proficiency index 2024: A ranking of 116 countries and regions by English skills*. EF Education First. https://www.ef.com/assetscdn/WIBIwq6RdJvcD9bc8RMd/cefcom-epi-site/reports/2024/ef-epi-2024-english.pdf.

Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong undergraduates. *Journal of English for Academic Purposes, 6*(1), 3–17. https://doi.org/10.1016/j.jeap.2006.11.005

Evans, S., & Morrison, B. (2011). Meeting the challenges of English-medium higher education: The first-year experience in Hong Kong. *English for Specific Purposes, 30*(3), 198–208. https://doi.org/10.1016/j.esp.2011.01.001

Fernández-Costales, A., & Lasagabaster Herrarte, D. (2024). English-Medium instruction in Spanish universities: A systematic review. *Revista de Educación, 403*, 225–248. https://doi.org/10.4438/1988-592X-RE-2024-403-612

Finstad, K. (2010). Response interpolation and scale sensitivity: Evidence against 5-point scales. *Journal of Usability Studies, 5*(3), 104–110.

Fornell, C., & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39–50. https://doi.org/10.1177/002224378101800104

Gaffas, Z. M. (2025). Language challenges in medical education: Exploring predictors and variations among EMI students in Saudi Arabia. *AILA Review, 38*(1), 5–44. https://doi.org/10.1075/aila.24013.gaf

Galloway, N., & Ruegg, R. (2022). English medium instruction (EMI) lecturer support needs in Japan and China. *System, 105*, Article 102728. https://doi.org/10.1016/j.system.2022.102728

Greenleaf, E. A. (1992). Measuring extreme response style. *Public Opinion Quarterly, 56*(3), 328. https://doi.org/10.1086/269326

Guarda, M., & Helm, F. (2017). 'I have discovered new teaching pathways': The link between language shift and teaching practice. *International Journal of Bilingual Education and Bilingualism, 20*(7), 897–913. https://doi.org/10.1080/13670050.2015.1125848

Harkness, J., Pennell, B.-E., Schoua-Glusberg, A., et al. (2004). Survey questionnaire translation and assessment. In R. M. Groves, et al. (Eds.), *Methods for testing and evaluating survey questionnaires* (pp. 453–473). Wiley. https://doi.org/10.1002/0471654728.ch22.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Johns, R. (2010). Likert items and scales. *Survey question bank: Methods fact sheet* (pp. 11–28). https://dam.ukdataservice.ac.uk/media/262829/discover_likertfactsheet.pdf.

Kling, J., & Dimova, S. (2015). The test of oral English for academic staff (TOEPAS): Validation of standards and scoring procedures. In A. Knapp, & K. Aguado (Eds.), *Fremdsprachen in studium und lehre chancen und herausforderungen fur den wissenserwerb* (pp. 247–268). Frankfurt/Main: Peter Lang.

Kling, J. (2015). You try with a little humour and you just get on with it": Danish lecturers' reflections on English-medium instruction. In S. Dimova, A. K. Hultgren, & C. Jensen (Eds.), *English-Medium instruction in European higher education* (pp. 116–137). De Gruyter Mouton. https://doi.org/10.1515/9781614515272-011.

Lasagabaster, D., & Ruiz de Zarobe, Y. (2010). *CLIL in Spain: Implementation, results and lecturer training*. Cambridge Scholars Publishing.

Lasagabaster, D. (2022). Lecturer preparedness for English-medium instruction. *Journal of English-Medium Instruction, 1*(1), 48–64. https://doi.org/10.1075/jemi.21011.las

Lasagabaster, D., & Doiz, A. (2021). *Language use in English-medium instruction at university: International perspectives on lecturer practice*. Routledge.

Li, N., & Wu, J. R. (2025). Assessment in EMI in the higher education context of Taiwan. *Journal of EnglishMedium Instruction, 4*(2), 212–237. https://doi.org/10.1075/jemi.24019.li

Lewis, J. R. (1993). Multipoint scales: Mean and median differences and observed significance levels. *International Journal of Human-Computer Interaction, 5*(4), 383–392. https://doi.org/10.1080/10447319309526075

Ley Organica/Organic Law 2/2023, of 22 March, on the Spanish University System. (2023). *Boletín oficial del estado, 70* pp. 1–36). https://www.boe.es/buscar/act.php?id=BOE-A-2023-7500.

Llurda, E., & Mocanu, V. (2024). English in Spain: Education, attitudes and native-speakerism. *World Englishes, 43*, 315–331. https://doi.org/10.1111/weng.12651

Malmström, H., Pecorari, D., & Warnby, M. (2025). Teachers' receptive and productive vocabulary sizes in English-medium instruction. *Journal of Multilingual and Multicultural Development, 46*(7), 1905–1923. https://doi.org/10.1080/01434632.2023.2260781

Malmström, H., & Pecorari, D. (2022). *Språkval och internationalisering: Svenskans och engelskans roll inom forskning och högre utbildning*. Institutet för språk och folkminnen/Språkrådets rapportserie [Language choice and internationalisation: The roles of Swedish and English in research and higher education][Language Council of Sweden reports]. https://www.diva-portal.org/smash/get/diva2:1619750/FULLTEXT02

McGorry, S. Y. (2000). Measurement in a cross-cultural environment: Survey translation issues. *Qualitative Market Research: An International Journal, 3*(2), 74–81. https://doi.org/10.1108/13522750010322070

Macaro, E. (2018). *English medium instruction*. Oxford University Press.

Macaro, E. (2020). Exploring the role of language in English medium instruction. *International Journal of Bilingual Education and Bilingualism, 23*(3), 263–276. https://doi.org/10.1080/13670050.2019.1620678

Mežek, Š. (2024). English in Sweden: Functions, features and debates. *World Englishes, 43*(2), 332–345. https://doi.org/10.1111/weng.12650

Molino, A., Dimova, S., Kling, J., & Larsen, S. (2022). *The evolution of Emi research in European higher education*. Routledge.

Orduna-Nocito, E., & Sánchez-García, D. (2022). Aligning higher education language policies with lecturers' views on EMI practices: A comparative study of ten European universities. *System, 104*, Article 102692. https://doi.org/10.1016/j.system.2021.102692

Pecorari, D., & Malmström, H. (2018). At the crossroads of TESOL and english medium instruction. *TESOL Quarterly, 52*(3), 497–515. https://doi.org/10.1002/tesq.470

Rose, H., Macaro, E., Sahan, K., Aizawa, I., Zhou, S., & Wei, M. (2023). Defining EMI: Striving for comparative equivalence in research design. *Language Teaching, 56*(2), 539–550. https://doi.org/10.1017/S0261444821000483

Shao, L., & Rose, H. (2024). Lecturers' experiences of English-medium instruction in higher education: A cross-case investigation of China, Japan and the Netherlands. *Journal of Multilingual and Multicultural Development, 45*(7), 2801–2816. https://doi.org/10.1080/01434632.2022.2073358

Shepard, C., & Morrison, B. (2021). Challenges of English-medium higher education: The first-year experience in Hong Kong revisited a decade later. In D. Lasagabaster, & A. Doiz (Eds.), *Language use in english-medium instruction at university: International perspectives on lecturer practice* (pp. 167–192). Routledge.

Soruc, A., Altay, M., Curle, S., & Yuksel, D. (2021). Students' academic language-related challenges in English medium instruction: The role of English proficiency and language gain. *System, 103*, Article 102651. https://doi.org/10.1016/j.system.2021.102651

Thompson, G., Aizawa, I., Curle, S., & Rose, H. (2022). Exploring the role of self-efficacy beliefs and learner success in English medium instruction. *International Journal of Bilingual Education and Bilingualism, 25*(1), 196–209. https://doi.org/10.1080/13670050.2019.1651819

Tschannen-Moran, M., & Hoy, A. W. (2001). Lecturer efficacy: Capturing an elusive construct. *Teaching and Lecturer Education, 17*(7), 783–805. https://doi.org/10.1016/S0742-051X(01)00036-1

Tsui, A. B. M (2018). Lecturer efficacy: A case study of faculty beliefs in an English-medium-instruction lecturer training program. *Taiwan Journal of TESOL, 15*(1), 1–33. https://doi.org/10.30397/TJTESOL.201804_15(1).0004

Vander Borght, M., Malmström, H., Martinez, R., & Pecorari, D. (2025). English proficiency, pedagogy, and confidence: What really matters in EMI teaching? *Journal of Multilingual and Multicultural Development*, 1–16. https://doi.org/10.1080/01434632.2025.2537740

van Splunder, F., Verguts, C., De Moor, T., & De Paepe, S. (2022). Testing lecturers' English language skills in a non-English-dominant academic context. *Journal of English-Medium Instruction, 1*(2), 255–274. https://doi.org/10.1075/jemi.21007.van

Wang, X. (2021). English language proficiency and science lecturers' self-efficacy in English-medium instruction. *Frontiers in Psychology, 12*, Article 611743. https://doi.org/10.3389/fpsyg.2021.611743

Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing, 27*(3), 236–247. https://doi.org/10.1016/j.ijresmar.2010.02.004

Wingrove, P., Zuaro, B., Yüksel, D., Nao, M., & Hultgren, A. K. (2025). English-medium instruction in European higher education: Measurement validity and the state of play in 2023/2024. *Applied Linguistics*. https://doi.org/10.1093/applin/amaf020. Advance online publication.

Yu, D. S., Lee, D. T., & Woo, J. (2004). Issues and challenges of instrument translation. *Western Journal of Nursing Research, 26*(3), 307–320. https://doi.org/10.1177/0193945903260554