

AN INSTRUMENT FOR ENGLISH MEDIUM INSTRUCTION (EMI) CLASSROOM OBSERVATION IN HIGHER EDUCATION

UM INSTRUMENTO PARA OBSERVAÇÃO DE AULA EM CONTEXTOS DE INGLÊS COMO MEIO DE INSTRUÇÃO NO ENSINO SUPERIOR

Ron Martinez¹¹³

Francisco Fogaça¹¹⁴

Eduardo Henrique Diniz de Figueiredo¹¹⁵

ABSTRACT: Classes taught through English in higher education (in countries where English is not an official language) is a growing phenomenon worldwide. In Brazil the trend has only emerged in the last decade, and has faced some resistance on many fronts, including among professors. One of the concerns raised by professors is related to their identity: essentially, are instructors who teach through a foreign language delivering a class that is qualitatively different? For example, are they as able to interact with the students in the same way they would in their L1? In order to move beyond mere conjecture regarding these and related questions, the present study describes the development and validation of a classroom observation instrument designed to be used (or adapted for use) by researchers wishing to investigate issues surrounding, especially, interactivity in English Medium Instruction in higher education settings.

Keywords: EMI; classroom interaction; bilingual education; internationalization.

RESUMO: Aulas ministradas no ensino superior por meio de língua inglesa (em países onde o inglês não é um idioma oficial) é um fenômeno crescente no mundo acadêmico. No Brasil, a tendência só surgiu na última década e tem enfrentado certa resistência em muitas frentes, inclusive entre os professores. Uma das preocupações levantadas por docentes é relacionada à identidade: isto é, será que os professores que ensinam através de uma língua estrangeira ministram uma aula qualitativamente diferente? Por exemplo, será que eles conseguem interagir com os alunos da mesma maneira que eles conduzem uma aula na primeira língua? Para ir além da mera conjectura sobre essas e outras questões relacionadas, o presente estudo descreve o desenvolvimento e a validação de um instrumento de observação de aula projetado para ser usado (ou adaptado para uso) por pesquisadores que desejam investigar questões relacionadas, principalmente, à interatividade didática quando em contextos de Inglês como Meio de Instrução em inglês no âmbito do ensino superior.

Palavras-chave: EMI; inglês como meio de instrução; ensino bilíngue; internacionalização.

1 Introduction

It is fairly well established that English as a medium of instruction (henceforth, EMI) in higher education is a growing global phenomenon (DEARDEN, 2014; WÄCHTER & MAIWORM, 2014), though that growth has often met with challenges (DOIZ, LASAGABASTER, & SIERRA, 2012). The decision to implement (or not) EMI—whether at a particular institution or across an entire country—is often fraught with complexity on a number of levels (Bradford, 2016), including issues of language policy (JENKINS, 2013; NUNAN, 2003), sociolinguistics (HU & LEI, 2014; LUEG & LUEG, 2015), teacher and student attitudes and beliefs (DEARDEN & MACARO, 2016; KIRKGÖZ, 2009), proficiency in English (JENSEN,

¹¹³ Graduado em Letras pela West Virginia University, Mestrado em Applied Linguistics and Second Language Acquisition pela University of Oxford. Doutor em Linguística Aplicada pela University of Nottingham e pós-doutorado pela University of Nottingham. Professor Adjunto da Universidade Federal do Paraná-UFPR.

¹¹⁴ Doutor em estudos da Linguagem pela Universidade Estadual de Londrina-UEL. Pós-doutorado na San Francisco State University. Professor no Curso de Letras e do programa de Pós-graduação em letras da Universidade Federal do Paraná-UFPR.

¹¹⁵ Doutor em Linguística Aplicada pela Arizona State University. Professor da Universidade Federal do Paraná-UFPR.

DENVER, & MEES, 2013; LEI & HU, 2014), assessment of proficiency (DIMOVA, 2017; KLAASSEN & BOS, 2010), and possible effects on learning (DAFOUZ, CAMACHO, & URQUIA, 2014; HERNANDEZ-NANCLARES, 2017). Broadly speaking, it can thus be said that extant literature on EMI has generally reflected that inherent complexity.

An area of EMI that has received somewhat less attention, however, is what actually happens inside EMI classrooms. There is a considerable amount of research that *alludes* to or even reports on concerns about classroom behavior in some form. For example, in a survey of 115 lecturers at a public university in Italy (HELM & GUARDA, 2015), most respondents voiced apprehensions related to teaching methodology (though many had never actually taught in English), and their (in)ability to be as spontaneous in English, vis-à-vis Italian. Similarly, lecturers at a private university in Ukraine reported having to adjust their speaking pace and classroom discourse style when delivering EMI classes (Goodman, 2014), though their classes were not actually observed. Indeed, there are similar misgivings that surface in most contexts, especially when EMI is new at an institution, and concerns range from lecturers' ability to deliver content as effectively as they would in their first language, to worries that students might be unable to participate as much in EMI classrooms, thus potentially hindering their learning (WÄCHTER, 2008).

However, with few exceptions, most research on actual classroom behavior has relied primarily on personal accounts (i.e. what lecturers and students report) regarding the nature of their interaction in class, as opposed to actual primary data collected from within the EMI classroom. An example of one of the few exceptions is Floris (2014), who observed the classes of 13 different EMI lecturers at a university in Indonesia, making field notes on classroom interaction and various aspects of the use of English (lexis, grammar, pronunciation, etc.). Floris noted that students seemed to be reluctant to participate in many of the classes observed. Another notable example can be found in Kling (2015), who used recorded observation at a university in Denmark for the purpose of stimulated recall and found that, unlike the respondents in the Helm and Guarda (2015) survey, the Danish lecturers did not feel that teaching in English somehow changed who they were as teachers in any fundamental way.

One could argue that the issue with all the aforementioned examples of research related to actual pedagogy in EMI classrooms is that it is difficult to infer very much regarding any differential effects of teaching in English when there is no real observed point of comparison, and when observations are based mainly on the impressions of one researcher. If students seem reluctant to speak in an EMI class, how is one to know whether that reluctance is most likely due to the language of instruction itself? (And not, for example, simply a usual characteristic of that subject, irrespective of language.) Likewise, if Italian lecturers report that they feel they are (or will likely be) less spontaneous and engaging when teaching in English, how can one ascertain the extent to which that is true (or not, as in the case of the Danish EMI lecturers)?

In the Brazilian context in which the research herein presented occurred, such question may be particularly relevant. Current legislation in Brazil provides that mandatory undergraduate courses may not be offered in a language other than Portuguese. One strategy Brazilian universities have adopted to circumvent this limitation is to offer a “mirrored” version of such courses in English, thus fulfilling legal requirements and meeting internationalization objectives (MARTINEZ, 2016). At the same time, it is important to be able to evaluate the extent to which such offerings are qualitatively comparable in their respective linguistic iterations.¹¹⁶

¹¹⁶ As researchers in the Global South, we feel it is important to note that it is not our intention here is to not preferentially endorse English *per se*. We are aware of the complex discourses of hegemony and imperialism that can be associated with the English language (as well as potentials and affordances). Instead, the research herein described emerged as a function of a perceived need to critically examine the rapidly growing phenomenon of courses in Brazilian higher education being taught through English. It is our hope that, in the future, more careful consideration

Dearden (2014) lists a number of pedagogically-related questions that have yet to be fully addressed in EMI research, including the following: "How does classroom interaction change as the medium of instruction changes? Does it become more interactive or less interactive?" (p.3). It was this question that drove the initial research herein described, which in turn eventually culminated in a related question: *Can a research instrument measure [E1] differences in classroom interaction behavior in EMI settings?*

2 Background

The present research began approximately three years before the time of writing, when the first author learned of an EMI program that had been instituted at a private university in the south of Brazil. What interested the researcher especially was the fact that this particular program often involved dual-version EMI courses in which the same lecturer would deliver the same content in Portuguese to one class, and in English to another. As mentioned in the previous section, such a scenario might provide especially valuable insight into questions of the extent to which the language of instruction affects the instruction itself as there would be a point of comparison. Permission was sought and granted to conduct observations, and several different lecturers were invited to participate in the study. Only one lecturer was willing and able to take part: Beatriz (a pseudonym), a Brazilian lecturer of civil engineering.

The first author began observations upon the second class meeting of the semester as a non-participant observer, making field notes and recordings of four total initial classes: two delivered in Portuguese (on Fridays), and their counterparts, with the same content, taught to an EMI class on the following Mondays. Those first notes already began to reveal some interesting points. For example, the researcher noticed that, often, Beatriz would closely mirror on a Monday that which she had taught on the Friday before, including elements of humor and even board work. On the other hand, field notes were made regarding some differences among the two groups of students, with the Friday (Portuguese) students often "tuning out" (as evinced by behaviors such as checking social media apps on their smartphones), while the Monday (EMI) students seemed less inclined to do so. Moreover, the EMI students even seemed to engage in the lecture as much as their Portuguese-language peers, answering and asking questions with roughly the same apparent frequency. The researcher's initial impressions, therefore, based on this very small sample size, was that Beatriz's EMI teaching and the interaction resulting from that teaching seemed to be not very different (or at least, not detrimentally so) from her instruction in Portuguese. Naturally, additional questions also abounded (e.g. previous EMI experience of the learners, the teaching experience of Beatriz herself, etc.), but the initial observations aroused enough curiosity to enlist the help of two researcher colleagues (both lecturers of English applied linguistics), who in turn joined the first author on a number of further non-participant observations using the same instruments (field notes and video recording), with the same objective: become familiar with Beatriz and her mirrored (Portuguese-EMI) teaching context.

Following a total of eight more observations (four classes in Portuguese, four in English of the same content), it was decided that, while we all concurred that there were a number of interesting insights to be gleaned, our interpretations of our observations were still based on general impressions. We knew we needed to gather more data —from more lecturers— but we needed to try to incorporate more systematicity into those observations to be able to establish, for instance, a degree of inter-rater agreement and reduce suspicions of researcher bias when making assertions (CRESWELL, 2013).

be given to other languages of instruction, for which the instrument presented here could be considered for use, perhaps in adapted form.

One commonly-used tool to introduce such rigor into classroom observation is the *observation schedule* (or *observation protocol/scheme*). Such instruments have been used extensively since at least the late 1960s to observe interaction (Flanders, 1968), and generally consist of a list of pre-defined behaviors (e.g. "teacher asks a question," "students read silently") that are either quantified in the form of a kind of tally over an entire lesson (*event-sampling*), or registered at regular intervals (*time sampling*) (DORNYEI, 2007). However, upon surveying the possibilities of adapting an existing schedule, we realized we had a problem. On the one hand, there are a number of existing schedules that are designed for general content classrooms, and especially for younger students (e.g. KADERAVEK & SULZBY, 1998; MERRETT & WHELDALL, 1986; MYHILL, 2002; RUBIE-DAVIES, 2007); on the other hand, there is an abundance of such instruments meant for language learning classrooms (ALLEN, 1983; E.G. ELDER, 1993; FRÖHLICH, SPADA, & ALLEN, 1985; LYSTER & RANTA, 1997; NABEI & SWAIN, 2002) — but the EMI classroom is not really meant to be a language-learning context *per se*. Yet it is a context in which language clearly matters.

3 Method

3.1 The EMICIOS: The First Version

Rather than attempt a Frankensteinish patchwork derived from instruments from both camps (general education and language education), we went back to our collective field notes and began to tease out recurrent themes/behaviors that seemed to be important. Those themes resulted in the original instrument (Figure 1 - translated into English), which we ultimately called the *EMI Classroom Interaction Observation Schedule* (henceforth, EMICIOS).

Figure 1 - First version of the EMI Classroom Interaction Observation Schedule (EMICIOS)

Time	Behavior	Com.?	Tallies	Total
1. VERBAL INTERACTION - ENGLISH				
	Lecturer asks a referential question			
	Lecturer asks a rhetorical question			
	Lecturer answers question			
	Lecturer uses humor ¹			
	Student answers question			
	Student asks a question			
	Student makes a comment			
	Students talk among themselves			
	Lecturer calls students' attention			
2. VERBAL INTERACTION - PORTUGUESE				
	Lecturer asks a referential question			
	Lecturer asks a rhetorical question			
	Lecturer answers question			
	Lecturer uses humor ¹			
	Lecturer uses Portuguese to translate			
	Student answers question			
	Student asks a question			
	Student makes a comment			
	Students talk among themselves			
	Lecturer calls students' attention			
NON-VERBAL INTERACTION				
	Lecturer writes on board			
	Lecturer draws on board			
	Lecturer monitors students			
	Students on mobile devices, etc.			

As pointed out by Mackey and Gass (2015), instruments can be developed to suit particular contexts in such a way, especially when they are driven by data recorded in the classroom. The format and even a few items (those related to questions) in the observation schedule shown in Figure 1 were in some part inspired by a similar instrument in Nunan (1991), but most of the instrument is unique and reflective of our collected field notes on Beatriz's classes. Linguistically speaking, we had noticed, for example, that Beatriz asked many rhetorical questions and often added question tags ("*But this is easy, right?*"), and wondered if she produced roughly the same number of such devices in both languages. Also on the language level, we felt the need to include separate sections for English and Portuguese. We also noticed, however, a number of interesting non-verbal behaviors that seemed relevant (and not included in existing observation protocols we had seen), such as students going "off-task" (e.g. looking at their phones), and the apparent importance of drawing on the board. Most of the items could be considered *low-inference* (i.e. not requiring much subjective interpretation), however, there is one notable high-inference item: "Lecturer uses humor." Obviously, what counts as "humor" is highly subjective, but on the form this was operationalized in a footnote as "action or behavior whose main purpose seems to be to make others smile or laugh." This behavior was included since it was one that seemed to be used by Beatriz on a number of occasions and, again, we wondered if its incorporation was roughly equal in both versions of her class, since the use of humor has increasingly been shown to be a kind of pedagogical "safe house" for multilingual teachers and students who can use such spaces to build alternate identities together (POMERANTZ & BELL, 2011). Moreover, as discussed in Section 1, it is a characteristic that has been variously reported in the EMI literature as being lost (HELM & GUARDA, 2015) or otherwise important (Kling, 2015).

3.2 EMICIOS: Pilot of first version

In order to test the effectiveness of the instrument, we recruited three additional research assistants (two undergraduates and one graduate) and piloted the EMICIOS using the video-recorded lectures of a different Brazilian EMI professor who the lead author had observed earlier in the year. After just two hours of using the instrument, we evaluated the level of inter-rater agreement among the six judges (which was very low, below 20%), and overall impressions. A number of problems surfaced, and Table 1 shows what those issues were, and how the instrument was changed as a result.

Table 1. Issues arising from first pilot, and alterations made

ISSUE	CHANGE MADE
The category "students talk among themselves" was found to be confusing as it could include both pairwork/groupwork and casual conversation not related to the lesson.	The title of the category was altered to "Students talk among themselves about the class."
There was a perceived need to include "display questions" (questions a lecturer knows the answer to) as a category (Long & Sato, 1983).	A separate category ("Lecturer asks display question") was created on the revised form.

It was also noticed that the type of follow-up move (Sinclair & Coulthard, 1975) offered by the lecturer to questions/comments could be subcategorized as "evaluative" (e.g. "right," "interesting") and "discursive" (commenting and expanding on student contributions) (Cullen, 2002), but also could simply be an "echo" (repeating what the student had said back to the student/class), and this seemed potentially interesting.	Referential questions received a three-box format in several rows for tallying different types of I-R-F instances (see Appendix).
The pilots noted that some referential and display questions were answered by students, and others not, which seemed worth keeping track of.	The "Display question" category section was formatted similarly to the "Referential" one above it (see Appendix), but also allowed for "no response."
There was some disagreement regarding what constituted a rhetorical question (but we also agreed it was worth keeping in the instrument).	No change.
In addition to "using" the board, all observers noticed that the lecturer also sometimes <i>referred</i> to the board (e.g. "If you look at...").	A section was added titled "Lecturer refers to the board/visuals."
A category for using other media (such as PowerPoint) was also indentified as important.	A section titled "Lecturer uses visuals" was added to the instrument.

In addition to the changes listed in Table 1, the entire section 2 ("Verbal Interaction - Portuguese") was removed. The main reason for this exclusion was a practical one: we needed space. However, it was also found to be somewhat superfluous since, in reality, both sections (English and Portuguese) were exactly the same but for one item ("Uses Portuguese to translate"). Moreover, there were observed instances of the students using Portuguese as well, and not for the purposes of translation. Thus, two new fields were added, simply, "Lecturer uses Portuguese" and "Students use Portuguese." The sections would allow for tallying instances of such events, and we agreed that elaboration on what and why the L1 use occurred could be jotted down on an attached "Comments" sheet. These new additions and alterations generated Version 2 of the EMICIOS.

3.3 Version 2: Second and third pilots

The researchers observed three more video-recorded hours of the same lecturer watched in the first pilot, after which inter-rater reliability increased to over 70% among all judges. There were in fact two separate instances of using the instrument during this pilot phase, with the first exercise reaching 78% reliability, but this figure actually decreased to 72% in the second use. The cause was identified as the sole inclusion that was made after the first exercise, a suggestion made by one of the researchers, who felt the need to include a field labeled "Lecturer comments on/explains topic." This category, when added to Version 2, caused some confusion among the researchers regarding how exactly it should be interpreted, thus reducing reliability figures.

Until this point, it should be noted, the EMICIOS had only actually been tested on the observation of two EMI lecturers. That changed in the final pilot, which ultimately resulted in our current version of the instrument.

4 Piloting for the current EMICIOS version

For what would be the final EMICIOS pilot, the researchers were invited to observe a new EMI course titled "Basic Concepts in Rural Science," a 30-hour course (the equivalent of an

elective for one semester) that would be delivered intensively by five different lecturers over a period of five days. None of the lecturers had ever taught an EMI course before. The enrolled students were senior undergraduate crop science students from the same university (n=18), with the only language requirement being "intermediate level of English" (student levels were determined based on students' self-reporting, rather than on a specific testing instrument). The class was roughly evenly divided in terms of gender, 10 male and 8 female.

The lecturers were observed at all times, and always by at least two researchers, but the researchers took the observations in shifts. Version 2 was used on Day 1 (Table 2), but an updated version — the one shown in the Appendix — was already created by Day 2 as a result of a number of issues (Table 2) encountered when using Version 2 with the new lecturers. (The lectures were also video-recorded.) These issues came to light when the researchers compared comments (written on a separate page) at the end of Day 1. We had noted that there were a number of questions of a type we had not yet commonly encountered in our observation of the other two lecturers. Examples are "Anyone else?" and "What do you think?" We determined that the function of these questions, in reality, was to elicit participation — not referential, display or rhetorical questions. In addition, we noticed yet another type of question that was identified as a comprehension check (e.g. "So, who can give me an example of...?").

The aforementioned question types were common enough on that first day that there was little doubt we needed to adjust our instruments right away. Adjustments were therefore made on the instrument by the end of Day 1. However, what is perhaps more interesting is a possible reason these two new question types surfaced. It happens that Chris and Victor (the Day 1 lecturers) had taken part in a 40-hour one-week intensive EMI lecturer training course just the week before, and on that course (delivered by EMI specialists from a British university) there was a focus on such techniques as elicitation and checking comprehension.

One more adjustment was made by the end of Day 1, which also emerged as a result of comparing notes. Specifically, we noticed that there was clear disagreement on what should constitute "Lecturer comments on/explains topic," and when (and how) it should be counted. It was therefore dropped as a category, and the new version (see Appendix) reflects this elimination, and the addition of the two new question types ("Elicits" and "Checks comprehension").

5 Results

Table 2 presents the key information regarding the observations conducted in the final pilot. (There was a fifth lecturer whose data was not included and not shown in Table 2 because the vast majority of her class occurred in a field outside the university building.)

Table 2. Details of the final pilot

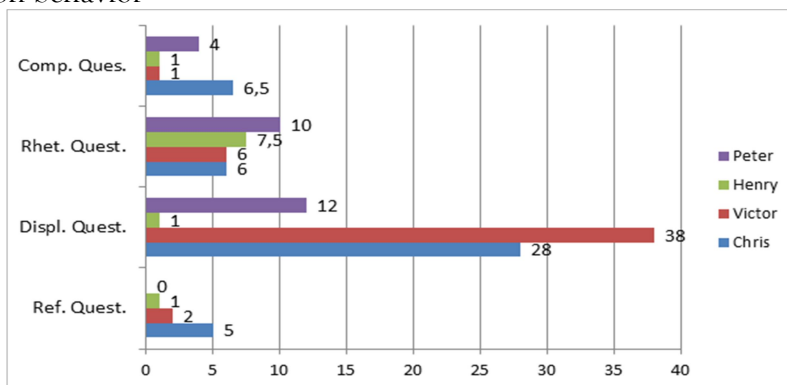
Lecturer	Days taught	Hours observed	Observers	Observer agreement	Comment
Chris	Day 1	3	R1, R3, R4	72%	Participated in EMI training
	Day 2	3	R1, R2	88%	
	Day 5	n/a	n/a	n/a	
Victor	Day 1	2.5	R1, R5	76%	Participated in EMI training
	Day 2	3	R2, R3	88%	
	Day 5	n/a	n/a	n/a	
Henry	Day 3	5	R1, R2, R3	94%	
Peter	Day 4	3	R3, R5	91%	

It should be noted that, although there was a Day 5, it was a day mostly devoted to student presentations, making use of the instrument less relevant. Nonetheless, 12 students were able to be interviewed about their impressions of the course on that final day, and these data would later prove useful for providing further insight.

A number of patterns emerge when looking at the results presented in Table 2, especially when taken together with the tallies and notes from the EMICIOS. First, the removal of the "Lecturer comments on/explains topic" item seems to have had a positive effect on inter-rater reliability. It also appears that the agreement percentages may have increased as a function of practice in use of the instrument. However, perhaps most interesting is looking at the differences among the scores for the different lecturers.

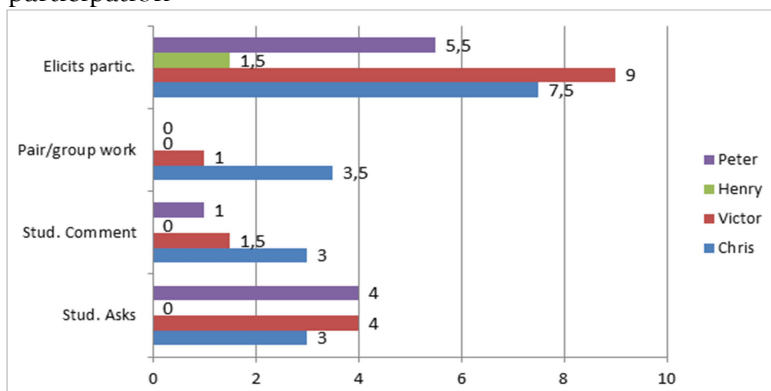
The question that drove this study was related to interaction, specifically, *Can a research instrument measure differences in classroom interaction behavior in EMI settings?* The EMICIOS seems to have been able to detect such differences. Perhaps where that ability is shown most clearly is in the differences between Henry and Victor. Figures 2, 3 and 4 present the totals for key interaction variables controlled for in the EMICIOS for the observed lecturers, averaged between total observer scores and adjusted per class-hour.

Figure 2 - Question behavior



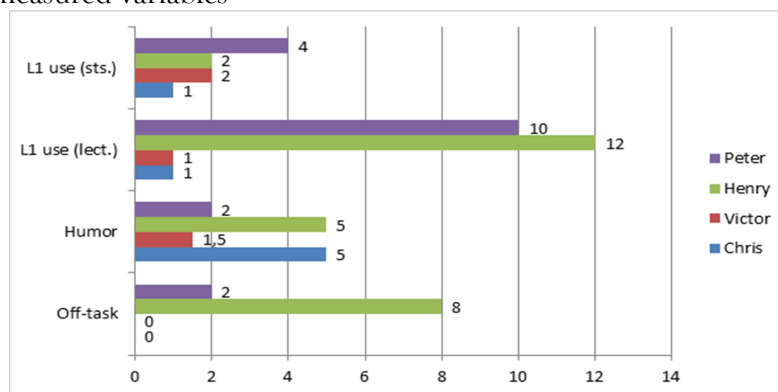
Victor clearly stands out as a prolific producer of questions in Figure 2, as does Chris, but especially of the display variety. Peter also produced a number of display questions, but no reference questions at all. Henry asked almost no display or reference questions the entire 5 hours he taught. Answers and follow-ups (R-F) were also accounted for. Chris and Victor nearly always received a reply from students (92 and 95%, respectively). Peter also received responses from students, but at a slightly lower rate (68%), and Henry, on the rare occasions he asked questions, never once received a reply. (Possible reasons for this are discussed later.) The number of rhetorical questions did not seem to vary much among the lecturers.

Figure 3 - Student participation



In Figure 3, Henry's data seems almost absent, with no pair or group work observed, nor student comments or questions of any kind. The same can be said of Peter's class regarding pair/group work. All lecturers elicited participation at some point, with Victor and Chris again standing out as the top two in that regard.

Figure 4 - Other measured variables



Henry is much more present in Figure 4, but arguably not always for positive reasons. He was recorded as using humor fairly consistently, with Victor making the fewest attempts. However, perhaps the variable that stands out most is "off-task." Whereas Victor and Chris have no such incidents recorded across 11.5 total hours of instruction, Henry averaged 8 (i.e. an average of 8 different students who "checked out" at some point during the class). In terms of L1 language use, Henry's class also is notable as having more than the others, though not much more than Peter.

6 Discussion and conclusion

As stated in the introduction, there is a need for more in-class observational research, and especially research that moves beyond individual impression and interpretation. The observation schedule is one instrument that can help address that need. However, as also pointed out in the introduction, the EMI classroom needs to be conceived of as a kind of hybrid of both content-focused and language-focused classrooms. It is both of those in some regards, but in the end it is its own unique context that merits its own special considerations when choosing data collection instruments. The EMICIOS was constructed driven by field notes from non-participant observers, and coincidentally incorporates some behaviors already featured in some existing observation schedules, such as I-R-F sequences, L1 use, and even off-task activity —yet no existing instrument was found that brought all these elements together in one observation schedule.

The question we sought to address in this research was *Can a research instrument measure differences in classroom interaction behavior in EMI settings?* We feel that the answer is yes, though there are many questions that still need to be answered, and much more use (i.e. trialing and analysis) in a wider diversity of contexts is required before the instrument in the Appendix can be said to have been thoroughly tried and tested. The experience of piloting the EMICIOS on the four lecturers proved insightful, but the results also carry some caveats. On the surface, it appears that Victor and Chris delivered lessons that were more interactive. This assertion is based partly on such variables as greater volume of questions (Figure 2), and more apparent student engagement (as measured by students not looking at their mobile phones, for example) (Figure. 4). However, strength is added to that assessment of how relatively interactive Victor and Chris's classes were when considering the interviews that were recorded from students on the last day of the course. When students were asked the question "What did you like about the course?", 8 of

the remaining 12 students (some students had left early) in focus groups mentioned Victor and Chris by name, stating that they had enjoyed the way those lecturers (Victor especially) had encouraged collaborative learning activities in class, and therefore felt more involved. Conversely, without eliciting any specific names, several students mentioned Henry as an exception. One obvious explanation for the apparent greater engagement in Victor and Chris's classes would be that they had received training just one week before —but this claim cannot be sustained by the EMICIOS data alone.

However, we did also interview all the lecturers. It is important to remember that the original question that drove the quest to develop the EMICIOS was not "Are some teachers 'better' than others?" but rather "Are EMI teachers in any way different in their classroom interactions when switching to English?" Interestingly, according to Chris and Victor, they felt that the EMI preparation course had "shaken up" their teaching, and reported they had become more interactive when teaching in English as a result. This assertion finds evidence in the observations made on the comments sheets of two of the observers, who noted that Victor and Chris seemed to purposely provide wait time after asking questions, and wait time was in fact one of the points emphasized on the EMI training course they took. (This may also explain their superior response rate noted in the Results section.) In this case, therefore, the switch to English appears to have led to a change and, indeed, similar accounts can be found elsewhere (e.g. GUARDA & HELM, 2017). Perhaps equally as interesting, when Henry was asked about his EMI class, he reported that he did "absolutely nothing" different from what he usually does in class, stressing that he felt no change in his identity or competence as a lecturer (though he did report sometimes struggling with the language). Peter made the same claim, with the only difference being that he felt students seemed to pay more attention in his EMI class than his students usually do when he teaches in Portuguese. In any case, to achieve a more valid understanding of whether Henry and Chris are really the same when teaching in either language, or whether the benefits the switch to EMI seems to have brought to Chris and Victor also extend to teaching in Portuguese, one would need to observe the lecturers in both teaching contexts. And that was the original reason the EMICIOS was designed.

In terms of the practical use and application of the instrument, there are a few important points. First, the EMICIOS from its inception was designed to be used together with complementary notes of qualitative nature. For example, all the lecturers used Portuguese at various points in their lessons, but those uses were not all functionally the same (some for example, were concept check questions, such as "How do you say it in Portuguese?" while others seemed to be uttered for some other effect). (Instead of tallies, in these cases the instances were assigned a number, which could later be referenced in the notes.) The instrument does not and cannot capture everything deemed relevant that happens in a class, so, for example, when a lecturer invited a student to come to the board to write his answer, it was recorded on that separate "Comments" sheet. In short, the quantitative and more objective aims of the EMICIOS are better served when *qualitative and subjective observation is included*. Moreover, it goes without saying that the use of EMICIOS must be triangulated with the use of other instruments that can elicit participants' perspectives, mainly qualitative interviews - as was the case in this study.

Although the EMICIOS could likely be employed in a variety of educational contexts (and not only EMI) we feel that the instrument and the particular constructs it aims to measure serve the EMI context best. The EMICIOS has been shown to be able to reliably measure constructs related to student engagement and lecturer confidence, which may later help shed light on the extent to which student lack of interest may be attributable to English proficiency, or pedagogy. If the case of Henry (who used more Portuguese than any other lecturer) is any indication, there is evidence that pedagogy would seem to matter more. This kind of evidence, in turn, is potentially

useful in the debate regarding whether or not EMI lecturers should be tested for proficiency. There is evidence that such testing can have negative effects (DIMOVA, 2017), and may prove to be not as important as the way one conducts a class in the first place. In that sense, too, the EMICIOS may prove useful. If one is to take it at face value that the EMICIOS was able to capture, on some level, the effectiveness of the training Victor and Chris had received, then the instrument may prove useful to help lecturers (and EMI training stakeholders) perceive the extent to which such courses can make a difference.

In the immediate future, we intend to continue trialing the EMICIOS in the original context for which the EMICIOS was designed: comparing differential interaction when the same lecturer delivers the same or similar content to two different groups of students in two different languages. We unfortunately do not yet know the extent to which there may be differences, but we now feel we at least and at last have a reliable instrument that can provide data to help us find out. On a last note, we reiterate that although the instrument was first used in EMI contexts, it may be adapted to settings in which other languages of instruction are utilized; and we certainly hope that we can go beyond EMI into more multi/pluri/translingual perspectives in classes on a number of contexts and areas of knowledge.

References

- ALLEN, P. The Communicative Orientation of Language Teaching: An Observation Scheme. 1983. Retrieved from <https://eric.ed.gov/?id=ED275155>
- BRADFORD, A. Toward a typology of implementation challenges facing English-medium instruction in higher education: Evidence from Japan. *Journal of Studies in International Education*, v. 20, n. 4, pp. 339-356, 2016.
- CRESWELL, J. *Research design: Qualitative, quantitative, and mixed methods approaches*. Los Angeles: University of Nebraska-Lincoln, 2013.
- CULLEN, R. Supportive teacher talk: The importance of the F-move. *ELT journal*, v. 56, n. 2, pp. 117-127, 2002.
- DAFOUZ, E.; CAMACHO, M; URQUIA, E. 'Surely they can't do as well': a comparison of business students' academic performance in English-medium and Spanish-as-first-language-medium programmes. *Language and Education*, v. 28, n. 3, pp. 223-236, 2014.
- DEARDEN, J. *English as a medium of instruction-a growing global phenomenon*. British Council, 2014.
- DEARDEN, J; MACARO, E. Higher Education Teachers' Attitudes towards English Medium Instruction: A Three-Country Comparison. *Studies in Second Language Learning and Teaching*, v. 6, n. 3, pp. 455-486, 2016.
- DIMOVA, S. Life after oral English certification: The consequences of the Test of Oral English Proficiency for Academic Staff for EMI lecturers. *English for Specific Purposes*, v. 46, pp. 45-58, 2017.
- DOIZ, A.; LASAGABASTER, D.; SIERRA, J.M. (Ed.). *English-medium instruction at universities: Global challenges*. Multilingual matters, 2012.
- DORNYEI, Z. *Research methods in applied linguistics: quantitative, qualitative, and mixed methodologies*. Oxford: Oxford University Press, 2007.
- ELDER, C. How do subject specialists construe classroom language proficiency?. *Language Testing*, v. 10, n. 3, pp. 235-254, 1993.
- FLANDERS, N.A. Interaction analysis and inservice training. *The Journal of Experimental Education*, v. 37, n. 1, pp. 126-133, 1968.
- FLORIS, F.D. Learning subject matter through English as the medium of instruction: students' and teachers' perspectives. *Asian Englishes*, v. 16, n. 1, pp. 47-59, 2014.

- FRÖHLICH, M.; SPADA, N.; ALLEN, P. Differences in the communicative orientation of L2 classrooms. *Tesol Quarterly*, v. 19, n. 1, pp. 27-57, 1985.
- GOODMAN, B.A. Implementing English as a medium of instruction in a Ukrainian University: Challenges, adjustments, and opportunities. *International Journal of Pedagogies and learning*, v. 9, n. 2, pp. 130-141, 2014.
- GUARDA, M.; HELM, F. 'I have discovered new teaching pathways': the link between language shift and teaching practice. *International Journal of Bilingual Education and Bilingualism*, v. 20, n. 7, pp. 897-913, 2017.
- HELM, F.; GUARDA, M. "Improvisation is not allowed in a second language": A survey of Italian lecturers' concerns about teaching their subjects through English. *Language Learning in Higher Education*, v. 5, n. 2, pp. 353-373, 2015.
- HERNANDEZ-NANCLARES, N.; JIMENEZ-MUNOZ, A. English as a medium of instruction: Evidence for language and content targets in bilingual education in economics. *International Journal of Bilingual Education and Bilingualism*, v. 20, n. 7, pp. 883-896, 2017.
- HU, G.; LEI, J. English-medium instruction in Chinese higher education: A case study. *Higher Education*, v. 67, n. 5, pp. 551-567, 2014.
- JENKINS, J. *English as a lingua franca in the international university: The politics of academic English language policy*. Routledge, 2013.
- JENSEN, C.; DENVER, L.; MEES, I. Students' attitudes to lecturers' English in English-medium higher education in Denmark. *Nordic Journal of English Studies*, v. 12, n.1, pp. 87-112, 2013.
- KADERAVEK, J.; SULZBY, E. Parent-child joint book reading: An observational protocol for young children. *American Journal of Speech-Language Pathology*, v. 7, n.1, pp. 33-47, 1998.
- KIRKGÖZ, Y. Students' and lecturers' perceptions of the effectiveness of foreign language instruction in an English-medium university in Turkey. *Teaching in Higher Education*, v. 14, n. 1, pp. 81-93, 2009.
- KLAASSEN, R.; BOS, M. English language screening for scientific staff at Delft University of Technology. *Hermes—Journal of Language and Communication in Business*, v. 45, n.1, pp. 61-70, 2010.
- KLING, J. You try with a little humor and you just get on with it": Danish lecturers' reflections on English-medium instruction. In: DIMOVA, S.; HULTGREN, A. K.; JENSEN, C. (Ed.). *English-Medium Instruction in European Higher Education: English in Europe*. Berlin: De Gruyter, 2015. pp. 201-222.
- LEI, J.; HU, G. Is English-medium instruction effective in improving Chinese undergraduate students' English competence? *International Review of Applied Linguistics in Language Teaching*, v. 52, n. 2, pp. 99-126, 2014.
- LONG, M.; SATO, C. Classroom foreigner talk discourse: Forms and functions of teachers' questions. In: SELIGER, H.; LONG, M. (Ed.). *Classroom Oriented Research in Second Language Acquisition*. Rowley, MA: Newsbury House, 1983. pp. 268-286.
- LUEG, K.; LUEG, R. Why do students choose English as a medium of instruction? A Bourdieusian perspective on the study strategies of non-native English speakers. *Academy of Management Learning & Education*, v. 14, n.1, pp. 5-30, 2015.
- LYSTER, R.; RANTA, L. Corrective feedback and learner uptake. *Studies in Second Language Acquisition*, v. 20, pp. 37-66, 1997.
- MACKEY, A.; GASS, S. *Second language research: Methodology and design*. 2. ed. New York: Routledge, 2015.
- MARTINEZ, R. English as a Medium of Instruction (EMI) in Brazilian higher education: challenges and opportunities. In: FINARDI, K. (Ed.). *English in Brazil: views, policies and programs*. Londrina: Eduel, 2016. pp. 191-228.

- MERRETT, F.; WHELDALL, K. Observing pupils and teachers in classrooms Merret(OPTIC): A behavioural observation schedule for use in schools. *Educational Psychology*, v. 6, n.1, pp. 57-70, 1986.
- MYHILL, D. Bad boys and good girls? Patterns of interaction and response in whole class teaching. *British Educational Research Journal*, v. 28, n. 3, pp. 339-352, 2002.
- NABEL, T.; SWAIN, M. Learner awareness of recasts in classroom interaction: A case study of an adult EFL student's second language learning. *Language Awareness*, v. 11, n. 1, pp. 43-63, 2002.
- NUNAN, D. *Understanding language classrooms: A guide for teacher-initiated action*. Upper Saddle River, NJ: Prentice-Hall, 1991.
- NUNAN, D. The impact of English as a global language on educational policies and practices in the Asia-Pacific Region. *TESOL Quarterly*, v. 37, n. 4, pp. 589-613, 2003.
- POMERANTZ, A.; BELL, N. Humor as safe house in the foreign language classroom. *The Modern Language Journal*, v. 95, n. s1, pp. 148-161, 2011.
- RUBIE-DAVIES, C. Classroom interactions: Exploring the practices of high-and low-expectation teachers. *British Journal of Educational Psychology*, v. 77, n.2, pp. 289-306, 2007.
- SINCLAIR, J.; COULTHARD, M. *Towards an analysis of discourse: The English used by teachers and pupils*. Oxford: Oxford University Press, 1975.
- WÄCHTER, B. Teaching in English on the rise in European higher education. *International Higher Education*, v. 52, pp. 3-4, 2008.
- WÄCHTER, B.; MAIWORM, F. (Ed.). *English-taught programmes in European higher education: The state of play in 2014*. Bonn: Lemmens Medien, 2014.

APPENDIX

The EMI Classroom Interaction Observation Schedule (EMICIOS)

VERBAL INTERACTION	
Lecturer asks REFERENTIAL QUESTION	
Question	
Response	
Follow-up	
Types of follow-up: E – evaluative; D – discursive; -- (no answer)	
Lecturer asks DISPLAY QUESTION	
Question	
Response	
Follow-up	
Types of follow-up: E – evaluative; D – discursive; -- (no answer)	
Lecturer asks RHETORICAL question (please assign numbers and note each instance on "Comments" sheet)	
Lecturer CALLS STUDENTS' ATTENTION	Lecturer CHECKS COMPREHENSION
Lecturer uses HUMOR ¹	Lecturer ELICITS from students
Student ASKS QUESTIONS	
Question	
Answer	
Types of teacher answer: E – evaluative; D – discursive; R – lecturer repeats back what student said	
Student MAKES A COMMENT	
Comment	
Reply	
Types of reply: A – evaluative; D – discursive; R – lecturer repeats back what student said	
Pairwork/Groupwork	
Lecturer REFERS TO BOARD/VISUALS	
Lecturer USES LI ²	Students USE LI
NON-VERBAL INTERACTION	
Lecturer WRITES ON BOARD	
Lecturer DRAWS ON BOARD	
Lecturer MONITORS/APPROACHES STUDENTS	
Students using mobile phone / other off-task activity not related to the lesson	

¹ Delivered with the apparent intention to make student(s) laugh or smile.

² Please try to briefly explain what the nature of the use of Portuguese was, its apparent function.

Recebido em: 30/09/19

Aceito em: 07/11/19