Contents lists available at ScienceDirect

System

journal homepage: http://ees.elsevier.com

Measuring effective teaching: Student perceptions of their modern languages lessons in England

Laura Molway

Department of Education, University of Oxford, 15 Norham Gardens, Oxford, OX2 6PY, UK

ARTICLE INFO

Article history:
Received 12 September 2020
Received in revised form 30 November 2020
Accepted 15 December 2020
Available online xxx

Keywords:
Student perception survey
Teacher effectiveness
Teacher evaluation
Modern languages

ABSTRACT

Evaluating the quality of classroom teaching is a complex task. Alongside observation tools and value-added student attainment measures, there is growing international interest in the potential of student perception surveys (SPS), which gather data about teaching practices from whole classes of first-hand observers. This paper reports on the refinement and trial of an existing SPS (available from the Colorado Education Initiative) to make it specific to the teaching of modern languages. Using both student (n = 1370) and teacher (n = 41) surveys, and follow-up interviews with Heads of Department (n = 6), the paper investigates students' classroom experiences of French lessons across fourteen secondary schools. The SPS_ML was was found to have high levels of internal reliability and provided data that correlated with a range of other known measures of effective teaching. Three distinct, languages-specific factors emerged from a factor analysis, which is interpreted as strong support for a subject-specific approach to developing measures of effective teaching. The three development areas suggested by the data relate to 1) the teachers' responsiveness to the needs and interests of their students; 2) the motivation and engagement of students in language lessons; and 3) the extent to which students are encouraged to adopt strategic approaches to language tasks.

1. Introduction

High-quality teaching is the most important school-based factor associated with student achievement scores (Schleicher, 2016). However, the identification and development of good classroom practice is known to be complex and difficult (Kane, Kerr, & Pianta, 2014). No wonder then, that significant time and resources have been invested in the identification of hallmarks of quality in teaching to inform quality assurance and programmes of initial and in-service teacher education in education systems around the world (OECD, 2017).

In much of the Western world, particularly in English-speaking jurisdictions such as Australia, England and the United States of America, high-stakes accountability systems evaluate teaching quality predominantly via students' test scores, sometimes alongside structured observations of teachers' classroom practice (Bowe & Gore, 2017). However, the large-scale Measures of Effective Teaching (MET) project questioned the validity and reliability of these measures and urged a more balanced approach to teacher evaluation involving a combination of multiple data sources (Bill & Melinda Gates Foundation, 2012). One currently underutilised source of information about teachers' work is students' own ratings of the effectiveness of their lessons. Student perception surveys (SPS) are "a cheap and easy source of good feedback about teaching behaviours from a range of observers who can draw on experience of many lessons" (Coe, Aloisi, Higgins, & Major, 2014, p. 4). Beyond their use for accountability purposes, there has also been much research interest in the potential for well-designed feedback and evaluation systems, including SPS, to support teacher's professional learning (Bill & Melinda Gates Foundation, 2013; Schleicher, 2011).

The study described in this paper involves the design and testing of a student perception survey tool (SPS_ML) to provide data about the effectiveness of teaching and learning in modern languages (ML) classrooms in England. An analysis of the students' responses highlights both strengths and areas for development in the subject-specific pedagogies employed by ML teachers across four-

E-mail address: laura.molway@education.ox.ac.uk (L. Molway)

https://doi.org/10.1016/j.system.2020.102440
0346-251/

teen secondary schools. I make the case that the utility and validity of tools for teacher evaluation can be enhanced by adapting such tools to gather data of relevance to subject specialist teachers wishing to use targeted feedback to improve their practice.

2. Literature review

2.1. Frameworks to identify high-quality teaching

High-quality teaching is commonly defined as teaching practices that are associated with improved student outcomes, encompassing both student attainment and also harder-to-measure outcomes such as student enjoyment and wellbeing (Kennedy, 2016). Many attempts have been made to define the attitudes, knowledge and skills that teachers require in order to be effective professionals. It is, difficult to identify what makes 'great teaching' but in recent years researchers have arrived at a degree of international consensus about what some of the indicators of high-quality teaching might be, leading to the development of 'quality teaching' frameworks and measures to guide teacher education and evaluation (Bowe & Gore, 2017; OECD, 2017; Raudenbush & Marshall, 2014; Schleicher, 2011, 2016; Tschannen-Moran & Hoy, 2001). In England, where this study is set, teachers' competencies across all phases and subject specialisms are assessed against the Teachers' Standards (Department for Education, 2013), which define the minimum level of practice expected for the award of qualified teacher status. These standards include eight domains, covering planning, teaching, subject knowledge, knowledge of child development, psychology and classroom management. Elsewhere, more specific competency frameworks have been produced in relation to ML teaching: The European Portfolio self-assessment tool for Student Teachers of Languages (EPOSTL) contains a list of 193 core competencies across seven domains (Newby et al., 2007) and the British Council (2015, p. 5), sets out twelve key competencies for high-quality teaching of English as a foreign language.

The construction of teacher competency frameworks such as those described above has been criticised for turning the complex and reciprocally interacting skills of teaching into a tick list of separate skills to be demonstrated. However, such frameworks can offer a transparent and often evidence-informed description of what is considered to be high-quality work in teaching, which can subsequently be used to devise evaluation tools such as 1) teacher self-assessments; 2) structured observation protocols; 3) student perception surveys; and 4) parent/carer perception surveys.

2.2. Teacher evaluation practices in England

In England, teacher evaluation exercises most commonly rely on student attainment and growth data from standardised national examinations alongside formal observations as part of an annual performance management system. The current norm is far from satisfactory: Kennedy (2015) warns against an over-reliance on student attainment and growth data in teacher evaluation exercises. The usefulness of such data is, of course, "highly dependent on the availability of good outcome measures" (Coe, Aloisi, Higgins, & Major, 2014, p. 4) and a narrow focus on student attainment gives "insufficient regard […] to the inherent complexity in the system and the social and cultural differences that exist more widely" (Ellis & McNicholl, 2015, p. 16). ML teachers may be particularly disadvantaged by the use of standardised measures of attainment to evaluate individual teacher performance due to issues with severe grading of ML examinations in England (Ofqual, 2014). There are also substantial differences in individual school ML policy (variations in the number of school hours and years of mandatory study; access to resources such as language assistants etc.) that are likely to impact both uptake and attainment in the subject. Finally, student attainment data does not provide much in the way of formative feedback to help the teacher to make improvements to their instructional practices.

Observations as a method of teacher evaluation may include more scope for feedback to support teacher development but when relied upon as a tool for accountability they are also highly contentious: The Office for Standards in Education abandoned the practice of grading individual lesson observations in 2014, acknowledging issues with the subjectivity of one-off observations conducted by a single observer. Indeed, even when observers are carefully trained in the use of a research-informed observation protocol, the ratings provided by a single observer are often highly volatile and unreliable (Bill & Melinda Gates Foundation, 2012 (Steinberg and Garrett, 2016)). Another common issue with observations is the practice of senior staff members conducting high-stakes observations in subject areas for which they lack relevant expertise. These observers may offer valuable perspectives on generic aspects of the lesson but there will be inevitable limitations to their understanding and interpretation of the instructional decisions made by the teacher and the extent to which they are able to offer meaningful feedback about them.

2.3. Subject-specific student perception surveys as teacher evaluation tools

Teacher evaluations that rely on student attainment data and lesson observations ignore the insights and voices of some of the key stakeholders in the system: teachers, students and parents. To build a more accurate picture of teacher quality, a combined approach involving multiple methods is needed, including some that can offer information about teachers' effectiveness over time time (Bill & Melinda Gates Foundation, 2013). SPS can form a useful part of such a combined approach since students have accumulated experience of their teacher over many lessons; it is relatively simple and cheap to collect their views; and SPS data is collected in a form that can easily be used as formative feedback for teacher development. Most importantly, student perception data has been found to be ***more*** reliable than most other forms of teacher evaluation: Students' own ratings of the effectiveness of their lessons were significantly more accurate in predicting student attainment on standardised tests than the ratings given by teachers or principals (Wilkerson, Manatt, Rogers, & Maughan, 2000).

Each subject discipline taught within secondary schools presents unique challenges (Thompson, 2015) and generic approaches to the identification of high-quality teaching are therefore insufficient. For example, as a ML specialist I have noticed that observation protocols developed around general indicators of teaching quality are unable to capture useful information about the target language use of teachers and students in the ML classroom. Given the central importance of this aspect of acquiring a language for communicative purposes (Ellis, 2010), this is a serious flaw. Researchers have already developed ML-specific observation protocols (i.e. Allen, Fröhlich, & Spada, 1984) and ML teacher self-assessment tools (i.e. Borg & Edmett, 2019, or the Teacher Effectiveness for Language Learning (TELL) project - www.tellproject.org). However, I am unaware of any existing SPS tools that have been designed to gather feedback from students in ML classrooms. There is therefore a need to adapt and test existing high-quality SPS tools to facilitate the gathering of nuanced and actionable, domain-specific data for ML teachers. This need is underlined by a number of reviews about teachers' development, which conclude that subject-specific professional learning is generally more beneficial than generic (Cordingley et al., 2015, 2018; Desimone, 2009; Schleicher, 2016).

3. Research questions

In order to explore the performance and perceived usefulness of a newly designed, subject-specific student perception survey (the SPS_ML), I constructed the following research questions:

- Does the SPS_ML offer reliable and valid measures of effective ML teaching?
- What does the SPS_ML reveal about students' experiences of ML lessons?
- To what extent is the data gathered by the SPS_ML considered useful for teacher learning?

4. Methods

4.1. Instrument development

The student survey adapted for use in this study is the Colorado Student Perception Survey (CSPS), which was developed and rigorously tested as a part of the MET project (Bill & Melinda Gates Foundation, n.d.). Responses from one set of classrooms were used to predict value-added gains in other classrooms taught by the same teacher, thus controlling for the possible effects of individual student prejudices and other confounding factors (Colorado Education Initiative, 2013). The CSPS is a 34-item instrument with a 4-point Likert response scale (Never; Some of the time; Most of the time; Always) that asks students about how frequently they experience teaching behaviours in the classroom that are known to create a positive classroom culture. The instrument has been found to measure four underlying dimensions of expert teaching (see Table 1).

In order to create a subject-specific version of the CSPS (the SPS_ML), I drew on recent second language acquisition research cataloguing approaches to the teaching of ML that are known to be effective within a communicative languages teaching (CLT) paradigm (Dörnyei, 2009; British Council, 2015; Ellis, 2005; Freeman, 2002; Lightbown, 2003; Macaro, Graham, & Woore, 2016; Nation & Macalister, 2010; Smith & Conti, 2016). I found a high degree of convergence across these studies, with the following areas considered universally important: grammatical awareness; phonology; speaking; listening; writing; reading; communication strategies. I chose to base the ML items for the student survey on eight research-informed principles for languages teaching in instructed settings (Macaro et al., 2016). These principles (reproduced in Table 2) were recently and extensively researched and developed with 241 ML teachers in England (see www.pdcinmfl.com). The eight principles were operationalised as 24 sub-statements and randomly spaced within the existing CSPS, with the aim of minimising order effects (Siminski, 2008). Table 3 contains examples of two items created in order to operationalise the first of the eight principles.

Table 1
The Colorado Student Perception Survey dimensions and example items.

Underlying Dimension of Expert Teaching	Example CSPS Items
Student Learning How teachers use content and pedagogical knowledge to help students learn, understand, and improve.	• My teacher makes learning enjoyable • If I don't understand something, my teacher explains it a different way
Student-Centred Environment How teachers create an environment that responds to individual students' backgrounds, strengths, and interests.	• My teacher respects my opinions and suggestions • Students feel comfortable sharing their ideas in this class
Classroom Community How teachers cultivate a classroom learning community where student differences are valued.	• I feel like an important part of this classroom community • My teacher would notice if something was bothering me
Classroom Management How teachers foster a respectful and predictable learning environment.	• Students in this class treat the teacher with respect • The students behave the way my teacher wants them to

Table 2
Eight Research-informed principles for languages teaching from Macaro et al., 2016.

Language teaching principles (source: Macaro et al., 2016)
1 Oral interaction: Target language input is essential for learning, but it can be made more effective if learners are encouraged to check the understanding of it by asking questions of what the teacher is saying or asking the teacher to repeat.
2 Oral interaction: Learners need to be encouraged to speak spontaneously and to say things that they are not sure are correct
3 Oral interaction: Less spontaneous oral interaction should nevertheless be of high quality. By high quality we mean including substantial student turns; adequate wait time; cognitive challenge [e.g. by requiring a verb phrase or subordinate clause]; appropriate teacher feedback; nominating students rather than waiting for volunteers.
4 Oral interaction: Students should be explicitly taught strategies to use when faced with communication difficulties. These should be used alongside techniques for developing their oral fluency, such as repetition of tasks and chunking of pre-learnt words into whole phrases.
5 Reading and listening: Learners need to be taught how to access a greater range of more challenging spoken and written texts, through explicit instruction in comprehension strategies and in the relationship between the written and spoken forms.
6 Strategic approaches: Learners need to develop their self- confidence and see the link between the strategies they use and how successful they are on a task.
7 Writing: Writing should be developed as a skill in its own right, not just as a consolidation of other language skills. For this to happen students should frequently write using the language and strategies they already know rather than resources provided by the teacher (e.g. textbooks, writing frames, dictionaries, etc.)
8 Skills focus: The principal focus of pedagogy should be on developing language skills and therefore the teaching of linguistic knowledge (knowledge of grammar and vocabulary) should act in the service of skill development not as an end in itself.

Table 3
Exemplar items from the refined SPS_ML.

Principle 1 - Oral Interaction	Items within the SPS_ML instrument
Target language input is essential for learning, but it can be made more effective if learners are encouraged to check the understanding of it by asking questions of what the teacher is saying or asking the teacher to repeat.	• My French teacher expects us to understand instructions in French. • I feel like I can stop my French teacher and ask a question if I don't understand something

A small pilot of the SPS_ML instrument was conducted with a convenience sample of one class of 23 Year 8 pupils of French in a grammar school in London (grammar schools differ from comprehensive secondary schools in that they select pupils based on their academic achievement). These pupils matched the target sample in terms of language learned and year group. However, a limitation of this pilot group is that grammar school pupils may not be representative of my target sample in terms of their level of attainment and therefore perhaps their capacity to understand and interpret the language and meaning of the questions. The pilot pupils completed the questionnaire and were invited to add comments about the process. Their responses were used to eliminate unclear wording and/or redundancy in the questionnaire. For example, some pupils were unsure about how to interpret the term 'strategies' in the item "In this class we learn strategies to help us when we are having a conversation in French". In response to this, a gloss was added to the item: "In this class we learn strategies to help us when we are having a conversation in French. (For example, asking someone in French to repeat their question or to explain what something means)".

4.2. Participants

The study draws on a combination of quantitative and qualitative methods to enable a triangulation of teachers' and students' perceptions. Online teacher surveys were used as an efficient method to collect a wide range of empirical data to complement student survey data (Gillham, 2000) whilst follow-up interviews with Heads of Department (HoDs) enabled a fuller investigation of the impact of the SPS_ML tool on teachers' thinking and practice.

Students of French (n = 1370) in thirteen publicly funded secondary schools local to the author's institution (in South-East England) were purposely recruited (see Table 4 for school profiles). The student participants were members of 59 intact classes, taught by 41 different teachers (see details in Table 5). Students in Year 8 (aged 12–13) were targeted because every school in the study had classes learning French in this year group and students in Year 8 are not yet in preparation for high-stakes examinations (whereas some schools in the study began their GCSE courses with Year 9 students). Twenty-nine of the 41 French teachers completed the teacher survey and six of the thirteen HoDs completed follow-up interviews, providing data for triangulation with the students' perceptions.

Table 4
Demographic data for the schools in the sample at the time of data collection.

School	School type	Age range	Ofsted rating	Location	Students on roll (to nearest 10)	ML teachers	% GCSE 5x a*-c	% students who achieved the EBacc	% students entered for the language area of the EBacc	% students with a statement of SEND	% students with EAL	% students in receipt of pupil premium grant
National average	State-funded schools				990		53.8	22.9	49.3	1.8	15	29.4
Aspen	Community School	11–18	2	Town	820	3	54	15	27	1.6	5.7	21.6
Birch	Academy	11–18	1	City	1880	11	74	43	65	2.3	22.5	19.7
Cedar	Academy	11–18	2	Town	1040	5	67	34	51	3.8	2.6	12.8
Dogwood	Academy	11–18	1	Town	2090	10	52	29	53	2.1	2.8	10.5
Elm	Community School	11–18	2	Town	1130	6	63	32	54	2.5	14.6	11
Fig	Academy	11–18	2	Village	1160	7	78	45	64	1.1	2.7	14
Ginkgo	Academy	11–18	2	City	1050	8	65	30	46	1	9.6	13.9
Hawthorn	Community School	11–16	2	Village	600	3	65	16	37	1.5	1	8.7
Ivy	Academy	11–18	2	Town	1150	5	61	18	47	1	7.3	13.6
Juniper	Academy	11–18	2	City	1380	7	63	34	60	1.2	32.1	33.3
Katsura	Academy	11–18	2	Town	870	3	59.7	25.4	22	0.5	9.2	21.4
Laurel	Academy	11–18	2	Town	1290	5	63	17	30	0.6	2.5	16
Nutmeg	Academy	11–18	2	Town	1000	5	64	24	45	1.1	6.2	24.8

Table 5
Demographic data for SPS_ML participants.

	Female	Male	Undisclosed	Total
Aspen	49	47	3	99
Birch	81	92	8	181
Cedar	35	30	6	71
Dogwood	72	64	9	145
Elm	34	30	4	68
Fig	50	45	6	101
Ginkgo	62	38	17	117
Hawthorn	55	42	2	99
Ivy	62	57	6	125
Juniper	50	57	3	110
Katsura	25	23	5	53
Laurel	14	19	1	34
Nutmeg	83	79	5	167
Total	**672**	**623**	**75**	**1370**

4.3. Data collection

Data collection took place from June to December 2017. Student and teacher surveys were administered in June and July. Following analysis of the survey data, HoDs were invited to participate in a follow-up interview. These interviews were conducted at a time and place of the HoDs' choosing between Oct–Dec.

4.3.1. Student surveys

Hard-copy SPS_ML surveys were delivered to the HoD in each school. The HoD then organised for surveys to be distributed to Year 8 students of French during the course of their usual language lessons. Surveys were completed anonymously but were marked with an identifier for each class. Asking students to give feedback on their enjoyment and progress is within the scope of schools' 'normal' practice (McIntyre et al., 2005) but pupils and parents were informed about the study and given the opportunity to opt out. It was made clear that participation was entirely voluntary, that students were free to withdraw themselves and their data at any time, and that doing so would incur no kind of penalty and have no effect on the assessment of their progress in school.

I anticipated that teachers could feel anxious about their students' responses as the SPS_ML represents yet another form of teacher evaluation within an education system that includes increasing levels of high-stakes teacher assessment and accountability (see Evetts, 2011). To counter this, assurances were given that the findings of the SPS_ML relating to individual teachers would not be accessible to anyone except the main researcher and no individual would be identifiable in subsequent reports or publications.

4.3.2. Teacher surveys

Self-report data were collected via an online survey using the Qualtrics platform. Unique survey links were emailed to all ML teachers in each school to enable linkage between student and teacher data. The data collected relate to teachers' sense of self-efficacy for teaching (Tschannen-Moran & Hoy, 2001) and the extent to which they felt their practice aligned with the same research-informed principles for instructed ML teaching that were investigated via the student questionnaire (Macaro et al., 2016). For each principle, teachers were asked to indicate whether they agreed with the following statements:

(1) This is a principle that I believe has some value
(2) This is a principle that I have experimented with in my own practice
(3) This is a principle that I regularly enact in my own practice

4.3.3. Head of department interviews

Average results of the SPS_ML were communicated via email to the HoD for languages in each school. Each school received a report in which the school's mean scores were presented alongside the whole-study averages (see Table 6 for an extract of one of the reports). Follow-up interviews were conducted 1–2 months after the reports were disseminated, at a time and place of the HoD's choosing. Interviews aimed to gather detailed information about HoDs' perceptions of the results and impact of the SPS_ML. All interviews were recorded and transcribed. The transcripts were imported into the NVivo 11 software package for analysis and coded to identify which SPS_ML items were mentioned by HoDs and to identify themes arising from the discussions. The key questions examined in this paper are listed below:

- Was there anything that surprised you about your students' results?
- Have you discussed your students' results with anyone else? If so, who and why?
- Have the students' responses helped you to identify any areas for professional learning this year?

5. Results and discussion

Firstly, I explore the performance of the SPS_ML and the extent to which there is alignment between teacher and student perceptions. The overall results of the survey are reported and the implications of these for ML teachers' development are discussed. Finally, HoDs' responses to the SPS_ML and their views as to the potential of the tool to inform teacher learning are discussed.

Table 6
Example feedback report for ML departments.

ML Teaching Principles www.pdcinmfl.com	Responded Favorably (Percentage of responses in top two categories)	
	Your ML Department	The Study Average
Principle 1: Target language input is essential for learning but it can be made more effective if learners are encouraged to check the understanding of it by asking questions of what the teacher is saying or asking the teacher to repeat.		
My French teacher expects us to understand instructions in French.	%	**66.8%**
I feel like I can stop my French teacher and ask a question if I don't understand something	%	**49.9%**
Principle 2: Learners need to be encouraged to speak spontaneously and to say things that they are not sure are correct		
In this class I feel safe to practise speaking in French.	%	**65.7%**
My French teacher encourages me to have a go and say things I am not sure are correct	%	**68.4%**
My French teacher expects us to speak some French in every lesson	%	**78.6%**
If I need to say something to the teacher, the teacher encourages me to say it in French	%	**39.2%**

5.1. The performance of the SPS_ML

Feedback from teachers regarding the administration of the SPS_ML was largely positive, with just two teachers commenting that the administration had taken longer than the anticipated 20 min. Responses to each item of the SPS_ML are summarised in Table 7. Of 1370 surveys, 1323 responses (96.6%) were complete across all items. One of the original CSPS items had a notably lower response rate than the average: 80 students (= 5.8%) did not respond to the item '*my teacher respects my cultural background*'. This may indicate an issue with the clarity of the term 'cultural background', suggesting a need for a future iteration of the SPS_ML to include some concrete examples of what it might mean to respect someone's cultural background in an ML lesson. Given that this was the penultimate item in the survey, it may indicate respondent fatigue (although a more general pattern of increasing non-response towards the end of the survey was not observed).

Cronbach's Alpha indicating the internal reliability for the SPS_ML and for the set of additional ML-specific items was exceptionally high at 0.97 and 0.91 respectively (see Table 8). Although the scale as a whole performs well, the clusters of items for each of the eight ML principles were found to have internal reliabilities of less than 0.7, meaning that separate analysis of the items underlying each ML principle is not statistically defensible.

The wording of the ML principles themselves includes some overlap (i.e. 'feedback' is highlighted in both principle 3 and 6; strategic approaches are mentioned in principles 4, 5 and 6). For this reason, it was deemed likely that a few underlying constructs would be found. Exploratory factor analysis was used to identify for each item the factor with which it is most strongly associated. A principal axis factor analysis was conducted on the 24 additional items with oblique rotation (direct oblimin). The Kaiser–Meyer–Olkin measure confirmed the sampling adequacy for the analysis, KMO = 0.93 ('marvellous' according to Hutcheson & Sofroniou, 1999). Three factors had eigenvalues over Kaiser's criterion of 1 and in combination explained 44.2% of the variance. Based on items with a loading of 0.25 and above (Stevens (2002) recommends factor loadings of 0.16 can be considered statistically meaningful for this sample size), the items that cluster on the same factor suggest three underlying constructs that may measure aspects of an effective ML classroom, strengthening the case for a domain-specific instrument (factor loadings can be seen in Table 9):

(1) Conditions of low anxiety & high linguistic support (9 items)
(2) High-frequency and quality target language interaction (5 items - relates to principles 1, 2 and 3 (Macaro et al., 2016))
(3) Strategic approaches to challenging language learning tasks (4 items - relates to principles 5 and 6 (Macaro et al., 2016))

The first factor identified above combines elements of several elements of the language teaching principles (Macaro et al., 2016) that focus on teacher actions to create conditions of low anxiety and high support. The need for teachers to identify and address language learning anxiety in the ML classroom is widely acknowledged due to the nature of the subject: the performative nature of speaking in the target language clearly involves an element of risk taking and researchers have also pointed to the ways in which engaging in target language tasks can be experienced by some students as a threat to their established L1 identities (Horwitz, Horwitz, & Cope, 1986).

The second and third ML-specific factors relate to discrete elements of the language teaching principles that teachers were asked to consider in their questionnaires, and it was therefore possible to investigate the extent to which teacher and pupil perceptions of these aspects of their French lessons were similar. Statistically significant correlations were found between teacher and student perceptions of the languages classroom in relation to both a) high-frequency and quality target language interaction (rho = 0.42, p = .02) and b) strategic approaches to challenging language tasks (rho = 0.35, p = .04). The correlation between teacher and student perceptions of the way in which target language is used is positive, particularly as it has previously been found that ML teachers' own estimates of the levels of target language use in their classrooms can be inaccurate (Copland & Neokleous, 2011).

The SPS_ML data was also triangulated with additional teacher self-report measures of ML teaching and school-level attainment data for languages at GCSE. Since each of these measures was intended to capture data related to the quality of ML teaching and learning I expected to find positive associations and high levels of covariance between them. Mean SPS_ML scores for each school are positively associated with the percentage of pupils achieving a pass grade in a language GCSE (rho = 0.34, p = .008), which suggests that positive experiences of languages teaching at Key Stage 3 support strong uptake of and performance in languages at GCSE level. A moderate positive correlation was also found between teachers' sense of self-efficacy for teaching (measured via the TSES) and their students' overall scores on the SPS_ML (rho = 0.40, p = .05). This means that the general confidence of teachers in this sample in their capacity to teach effectively was associated with their pupils' perceptions of a range of positive teaching strategies within their lessons.

5.2. Overview of development needs suggested by the SPS_ML results

The proportion of students who gave positive responses ('often' or 'always') to each item of the SPS_ML ranged from 9.8% to 78.6%, with an average of 55.8% (see Table 7). Across all survey items, the most frequent response category was 'often' (33.7%), followed by 'sometimes' (29.5%), 'always' (23.6%) and 'never' (13.2%). 86.8% of students' responses indicated that they experienced the range of effective French teaching behaviours and learning conditions contained within the SPS_ML at least 'sometimes'.

Table 7
Overall results for the SPS_ML.

	Responded Favorably (Percentage of responses in top two categories measured on a scale from 1 = never to 4 = always; n = 1370)	Mean	SD	Missing N
Student Learning				
My French teacher makes learning enjoyable.	55.1%	2.66	0.989	15
What I learn in this class is useful to me in my real life.	37.4%	2.26	0.953	8
My French teacher teaches things that are important to me.	58.1%	2.74	0.907	31
My French teacher knows the things that make me excited about learning.	32.4%	2.11	1.028	37
In this class, we learn a lot every lesson.	69.8%	2.85	0.757	0
When the work is very difficult, my French teacher encourages me to keep trying.	66.9%	2.98	0.980	17
My French teacher accepts nothing less than my best effort.	71.3%	3.05	0.907	39
My French teacher knows when we understand the lesson and when we do not.	55.8%	2.68	0.924	26
If I don't understand something, my French teacher explains it a different way.	56.3%	2.68	0.957	14
My French teacher explains difficult things clearly.	64.8%	2.83	0.865	23
In this class, we have a say in what we learn and do.	29.9%	2.10	0.897	37
My French teacher gives feedback on my work to help me to improve.	69.3%	3.01	0.912	28
When we study a topic, my French teacher makes connections to other subjects or classes.	34.1%	2.19	0.894	15
Student-Centred Environment				
My classroom is organised, and I know where to find what I need.	71.6%	3.06	0.915	21
Students feel comfortable sharing their ideas in this class.	60.6%	2.78	0.853	19
My French teacher respects my opinions and suggestions.	73.4%	3.12	0.915	19
My French teacher cares about me.	58.2%	2.79	1.056	37
My French teacher pays attention to what all students are thinking and feeling.	45.9%	2.48	0.972	52
My French teacher respects my cultural background.	64.8%	3.06	1.094	80
My French teacher respects me as an individual.	67.4%	3.01	0.990	37
Classroom Community				
My French teacher would notice if something was bothering me.	43%	2.40	0.986	67
Our classroom materials (books, articles, videos, art, music, posters, etc.) show that French is spoken by people from all over the world.	50.5%	2.61	1.040	35
In this class, I feel like I fit in.	63.8%	2.85	0.991	27
I feel like an important part of this classroom community.	47.8%	2.48	0.965	30
My French teacher knows what my life is like outside of school.	9.8%	1.45	0.751	29
My French teacher knows what is important to me.	22.4%	1.86	0.905	34
I ask for help when I need it.	67.2%	2.99	0.904	33
I feel like I work well in this class.	71.2%	3.00	0.951	32
Classroom Management				
My class gets on with the work and does not waste time.	39.5%	2.33	0.798	28
Students in this class treat the French teacher with respect.	64.5%	2.80	0.840	28
The students behave the way my French teacher wants them to.	47.2%	2.48	0.870	27
Students in this class respect each other's differences.	64.4%	2.82	0.890	25
Language Teaching Principles				
Principle 1: Target language input is essential for learning, but it can be made more effective if learners are encouraged to check the understanding of it by asking questions of what the teacher is saying or asking the teacher to repeat.				
My French teacher expects us to understand instructions in French.	66.8%	2.89	0.836	31
I feel like I can stop my French teacher and ask a question if I don't understand something	49.9%	2.57	0.998	26
Principle 2: Learners need to be encouraged to speak spontaneously and to say things that they are not sure are correct				
In this class I feel safe to practise speaking in French.	65.7%	2.96	0.964	14
My French teacher encourages me to have a go and say things I am not sure are correct	68.4%	2.97	0.909	31
My French teacher expects us to speak some French in every lesson	78.6%	3.24	0.867	20
If I need to say something to the teacher, the teacher encourages me to say it in French	39.2%	2.33	1.020	37
Principle 3: Less spontaneous oral interaction should nevertheless be of high quality. By high quality we mean including substantial student turns; adequate wait time; cognitive challenge [e.g. by requiring a verb phrase or subordinate clause]; appropriate teacher feedback; nominating students rather than waiting for volunteers.				
Sometimes in French lessons the teacher asks me an unexpected question and I have to think on my feet to try to answer in French	48.4%	2.53	0.875	16

	Responded Favorably (Percentage of responses in top two categories measured on a scale from 1 = never to 4 = always; n = 1370)	Mean	SD	Missing N
When we are speaking in French, our teacher expects us to try to use whole sentences	81%	3.22	0.782	13
When the teacher asks me a question in French they always give me enough time to think about my answer	60.3%	2.78	0.951	41
Principle 4: Students should be explicitly taught strategies to use when faced with communication difficulties. These should be used alongside techniques for developing their oral fluency, such as repetition of tasks and chunking of pre-learnt words into whole phrases.				
In this class we learn how to cope when we have difficulty finding the right words to say	53.1%	2.59	0.875	28
In this class we learn strategies to help us when we are having a conversation in French. (For example, asking someone in French to repeat their question or to explain what something means)	46%	2.44	0.972	16
Principle 5: Learners need to be taught how to access a greater range of more challenging spoken and written texts, through explicit instruction in comprehension strategies and in the relationship between the written and spoken forms.				
In this class we read and listen to real French songs, poems, books and videos.	40.3%	2.44	0.927	17
In this class we learn to recognise patterns in how French sounds are written down	57.8%	2.71	0.899	32
The teacher gives us things to read or listen to that I find challenging	52.9%	2.65	0.867	37
When we are reading or listening to French that is challenging, the teacher shows us strategies to help us understand	54.3%	2.64	0.949	32
Principle 6: Learners need to develop their self-confidence and see the link between the strategies they use and how successful they are on a task.				
My French teacher talks to me about my work to help me to understand my mistakes.	53.1%	2.62	0.941	22
In this class I am learning how to be independent (For example I know how to use a dictionary and my notes to help me)	71.2%	2.98	0.861	2
My teacher helps me to think about the different strategies that I can use to tackle different tasks	45.3%	2.48	0.894	41
Principle 7: Writing should be developed as a skill in its own right not just as a consolidation of other language skills. For this to happen students should frequently write using the language and strategies they already know rather than resources provided by the teacher (e.g. textbooks, writing frames, dictionaries, etc.)				
When writing in French in this class it is more important that my teacher can understand what I mean than to get my spellings perfectly correct	63.3%	2.84	0.872	45
The teacher asks us to try writing in French without using our books and dictionaries for help	53.7%	2.65	0.868	35
My French teacher encourages me to have a go and write down things I am not sure are correct	71.8%	3.06	0.893	18
Principle 8: The principal focus of pedagogy should be on developing language skills and therefore the teaching of linguistic knowledge (knowledge of grammar and vocabulary) should act in the service of skill development not as an end in itself.				
In this class, it is more important to understand and communicate in French than to memorise all of the grammar and vocabulary.	51.9%	2.60	0.861	24
The things I learn in my French class would help me to speak to a real French person in the future	58.8%	2.74	0.972	32

Table 8
Reliability analysis for the main domains of the SPS_ML.

	Number of items	N	Mean	SD	alpha
Overall SPS_ML instrument	56	1336	2.7	.55	.97
Student learning	15	1320	2.62	.62	.91
Student-centred environment	7	1287	2.89	.72	.86
Classroom management	4	1312	2.44	.63	.80
Classroom community	8	1319	2.59	.67	.82
Language teaching principles	24	1134	2.8	.06	.91

The aim of the SPS_ML tool is to provide teachers with a diagnostic assessment of their students' perceptions of the classroom to enable the identification of areas to work on. A study such as this one also provides an opportunity to examine students' perceptions of language learning across several different schools. At the level of the whole teacher cohort, 15 survey items received mean scores of 2.5 or less, indicating that most students responded with the lowest two response categories ('never' or 'some of the time'). These lowest-scored items include a series of statements about the degree to which the French teacher knows about and is responsive to the needs and interests of their students:

My French teacher knows what my life is like outside of school (mean = 1.5)
My French teacher knows what is important to me (mean = 1.9)
My French teacher knows the things that make me excited about learning (mean = 2.1)

Table 9
Summary of SPSS factor analysis results for the ML-specific items on the SPS_ML (N = 1336).

Item	Conditions of low anxiety & high linguistic support	High-frequency and quality target language interaction	Strategic approaches to challenging language learning tasks
When the teacher asks me a question in French they always give me enough time to think about my answer	**.71**	-.08	-.09
I feel like I can stop my French teacher and ask a question if I don't understand something	**.70**	-.07	-.02
My French teacher encourages me to have a go and say things I am not sure are correct	**.68**	.06	.00
My French teacher talks to me about my work to help me to understand my mistakes.	**.67**	-.04	-.14
In this class we learn how to cope when we have difficulty finding the right words to say	**.65**	.06	-.14
My French teacher encourages me to have a go and write down things I am not sure are correct	**.65**	.00	.08
In this class we learn strategies to help us when we are having a conversation in French. (For example, asking someone in French to repeat their question or to explain what something means)	**.61**	-.07	.16
In this class I am learning how to be independent (For example I know how to use a dictionary and my notes to help me)	**.52**	.05	.16
When writing in French in this class it is more important that my teacher can understand what I mean than to get my spellings perfectly correct	**.50**	.03	-.04
My French teacher expects us to understand instructions in French.	-.11	**.55**	.06
Sometimes in French lessons the teacher asks me an unexpected question and I have to think on my feet to try to answer in French	-.02	**.42**	-.05
My French teacher expects us to speak some French in every lesson	.20	**.37**	.06
When we are speaking in French, our teacher expects us to try to use whole sentences	.27	**.35**	.25
Usually if I need to say something to the teacher, the teacher encourages me to say it in French	.30	**.25**	-.27
My teacher helps me to think about the different strategies that I can use to tackle different tasks	.58	.02	**-.38**
When we are reading or listening to French that is challenging, the teacher shows us strategies to help us understand	.63	-.05	**-.35**
The teacher gives us things to read or listen to that I find challenging	.11	.35	**-.25**
In this class we read and listen to real French songs, poems, books and videos.	.25	.18	**-.24**

In this class, we have a say in what we learn and do. (mean = 2.1)
My French teacher would notice if something was bothering me (mean = 2.4)
My French teacher pays attention to what all students are thinking and feeling (mean = 2.5)

The development of personal and positive rapport with students is vital when creating a climate conducive to learning (Kaka, 2019) and it is concerning that many students responding to this study felt this element was lacking in a subject focussed, at least to some degree, on developing effective interpersonal communication. Balancing relationship and rapport building (often in the L1) with ensuring that students receive enough input and opportunity to communicate in the ML is also a key pedagogical challenge for ML teachers (Hall & Cook, 2012). A possible reason for the relatively low scores in the items above may be the lack of curriculum time ML teachers are able to spend getting to know their students: the Teaching Schools Council (2016) has recently recommended that, ideally, schools should be offering at least 3 h of ML teaching per week. However, in the majority of English secondary state schools, the learning of a language is allocated less time than recommended and there are frequent reports of 1 h per week or less allocated to KS3 classes (Tinsley & Board, 2017; Tinsley & Doležal, 2018).

Two of the fifteen lowest-scored items indicate issues with students' behaviour and engagement in language lessons. In combination with other low-scored statements, it appears that this may be connected with a prevalent perception that work in ML lessons lacks authenticity and real-life application:

The students behave the way my French teacher wants them to (mean = 2.5)
My class gets on with the work and does not waste time (mean = 2.3)
What I learn in this class is useful to me in my real life (mean = 2.3)
In this class we read and listen to real French songs, poems, books and videos. (mean = 2.4)
When we study a topic in French, my teacher makes connections to other subjects or classes (mean = 2.2)

In the wider literature there is evidence that in England, pupil motivation is at a low ebb (Lanvers, Hultgren, & Gayton, 2019; Tinsley, 2019). In response to this trend, there has been a popular discourse in England emphasising language learning as a practical means to access future employment opportunities. However, as Coffey (2018, p. 475) points out, "this type of instrumentalism posits

a set of motives that are easy for students who do not envisage a future needing this transactional capital to refute". There are therefore current calls for those designing and teaching language curricula and examinations in England to reconsider whether achieving a working level of fluency in communication in a ML should be the only or even the main aim of language tuition. In the last decade, increased globalisation has contributed to a growing recognition of the need for intercultural understanding to be a focus of language education programmes (Moeller & Nugent, 2014). The current Programme of Study for the National Curriculum in England at Key Stage 3 (DfE, 2013) declares in its opening sentence that "learning a foreign language is a liberation from insularity and provides an opening to other cultures". However, cultural knowledge and understanding does not feature prominently in national assessments of ML study and in practice Hennebry (2014) found that it rarely featured in ML lessons in English schools. Another factor that is likely to result in student demotivation is a perception, confirmed by European comparisons, of low levels of progress made in language learning over the course of the secondary school years (Ofsted, 2015, p. 150106).

Finally, there are two low-scored statements that suggest ML teachers in England could further develop their approaches to building students' repertoire of strategies for language tasks:

> We learn strategies to help us when we are having a conversation in French (For example, asking someone in French to repeat their question or to explain what something means) (mean = 2.4)
> My teacher helps me to think about the different strategies that I can use to tackle different tasks (mean = 2.5)

There is a great deal of evidence in the published literature to suggest that explicit teaching of cognitive and metacognitive strategies for tackling language tasks across the four skills (listening, speaking, reading and writing) can be beneficial to learners and can increase their sense of self-efficacy and self-regulatory behaviours (Cohen, 2011; Graham, 2007). However, the results of this study suggest that, at least across this sample of fourteen state schools, approaches to developing strategic approaches to language learning are not yet be firmly embedded in ML teachers' practice.

5.3. The perceived usefulness of the SPS_ML

All six HoDs who were interviewed identified elements of the students' perceptions that they found surprising. The fact that the survey tool did not simply reveal patterns that HoDs were anticipating could be viewed as a positive outcome as the tool provided colleagues with a provocation for reflection and discussion, as highlighted by the comments below:

> We discussed this yesterday and it was a great way to start a conversation about some uncomfortable truths regarding the habits we have fallen into as a department. (TN1, HoD, Nutmeg School)
> I think one that I was disappointed on was […] teachers giving me feedback to help me improve my work because that is something that we started work on last year and […] I did a project myself as well last year as to how we can do it and how best to do it. So we still have work to do. (TI3, HoD, Ivy School)
> […] interesting reading indeed! We need to work on our students' TL use. (TD2, HoD, Douglas-Fir School)

As the above quotations demonstrate, the SPS_ML data that I had collected and collated for HoDs functioned "as a currency that facilitates the access of outsider researchers into contradictory and uncomfortable places in a school community" (Howes, Frankham, Ainscow, & Farrell, 2004, p.243).

Four of the six HoDs had already discussed the SPS_ML results with their colleagues and were using the lower-scoring items in their faculty report to inform their development planning. This is exemplified by the following comment by TB5, HoD at Birch School (TB5 was not interviewed but sent me an email following receipt of the SPS_ML results):

> This is a wonderful document to help us improve. Overall, I am rather pleased but it does highlight some areas for development and this will go into my FDIP [= Faculty Development and Improvement Plan] for 2017–2018.

Of the two HoDs who had not shared the results of the report with their colleagues, one (TD2) explained that they felt some of the low scores could demoralise their team ("You know 'my French teacher makes learning enjoyable'. Only 39%. With the average being 55%. Considering how dynamic the team is and how creative, I think if they saw that they would probably be quite upset by that so I didn't really want to show them"); the other (TN1) highlighted a lack of time at the start of the busy school year as the reason why they had not yet discussed the report with their colleagues. The issue of timing was also raised by TI3, who said "the problem with September has been that […] it's possibly one of the busiest terms of the year so when I circulated it to my colleagues I had a couple of comments back but we haven't really sat down and said actually what do we think?". It seems that receiving the report in September facilitated some departments in planning ahead for their PL but perhaps hindered others due to the slew of conflicting priorities at the beginning of the school year.

Subject-specific items on the report were explicitly highlighted (unprompted) by five of the six HoDs, suggesting that these items were of value and interest to them. The HoD of Elm School (TE2) was not interviewed but in an email said of the report: "it was really, really, really interesting because it's based on evidence about language learning". This aligns with prior research calling for subject-specific and evidence-informed approaches to teacher evaluation and development (Cordingley et al, 2015, 2018).

6. Conclusions

This study set out to answer the following two research questions:

- Does the SPS_ML offer reliable and valid measures of effective ML teaching?
- What does the SPS_ML reveal about students' experiences of ML lessons?
- To what extent is the data gathered by the SPS_ML considered useful for teacher learning?

The SPS_ML tool was found to have high levels of internal reliability and provided data that correlated with a range of other known measures of effective teaching. Three distinct, languages-specific factors emerged from a factor analysis and these offer evidence for the value of a subject-specific approach to developing measures of effective teaching.

The data gathered by the SPS_ML suggest that students often feel that their French teacher does not really know them as individuals. Students' felt that behaviour in ML lessons was frequently not in line with teachers' expectations and there are indications that many students do not see the value in learning French. Aspects of French lessons that students reported infrequent experiences of included exposure to authentic, cultural resources and the development of strategic approaches to language tasks. These findings have implications for language teachers and teacher educators – they suggest aspects of ML teaching that may currently be underdeveloped in England's classrooms and that have the potential to improve students' engagement and attainment in language learning.

The domain-specific student perception survey explored in this project offered teachers a tool with which they could consult their students in order to identify both successes and priorities for future development of their classroom practice. HoDs valued their school-specific survey reports and reported that the tool was useful as a means with which to identify high quality ML teaching and to understand aspects of languages teaching that their team could work to improve.

This research contributes to the field by offering a model for the development of subject-specific measures of effective teaching with high potential to identify good practice and inform teachers' decision making regarding their future development planning. The study provides evidence of the positive impetus for development that such a tool can give to teachers when it is employed in a low-stakes environment.

6.1. Future directions

The study reported here is a first step in the development of a robust, valid and reliable ML-specific student perception survey tool and there is much scope for further refinement. I consider it a regrettable oversight that the SPS_ML reported in this paper does not gather specific information about students' development of intercultural understanding beyond the item relating to classroom resources. Future iterations of the SPS_ML should include additional items focussed on the extent to which students feel they are developing a sense of intercultural understanding in their ML lessons. Future work must also acknowledge the limitations of any single tool for the evaluation of teaching practices. The SPS_ML should be used in combination with other data sources that are more sensitive to contextual variation.

Future studies could look to refine and cross-validate the SPS_ML instrument with a view to enabling language teachers to use it formatively to interrogate their practice. Extensive teacher and student feedback on the survey via interview and think-aloud processes would be helpful. The process would also involve the triangulation of SPS_ML data with teacher self-evaluations and systematic lesson observation data, using pre-existing instruments such as the communicative orientations to language teaching (COLT) protocol developed by Allen et al. (1984).

References

Allen, P., Fröhlich, M., & Spada, N. (1984). The communicative orientation of language teaching; an observation scheme. In Handscombe, D., Orem, R., & Taylor, B. (Eds.), On TESOL '83 (pp. 231–252). TESOL.

Bill Melinda Gates Foundation Measures of effective teaching(n.d.), Retrieved fromhttp://www.metproject.org

Bill, & Melinda Gates Foundation. (2012). Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. MET Project Research Paper, 1–4.

Bill, & Melinda Gates Foundation. (2013). Ensuring fair and reliable measures of effective teaching: Culminating findings from the MET Project's three-year study. MET Project Policy and Practitioner Brief, 1–26.

Borg, S., & Edmett, A. (2019). Developing a self-assessment tool for English language teachers. Language Teaching Research, 23(5), 655–679. doi:10.1177/1362168817752543.

Bowe, J., & Gore, J. (2017). Reassembling teacher professional development: The case for quality teaching rounds. Teachers and Teaching: Theory and Practice, 23(3), 352–366. doi:10.1080/13540602.2016.1206522.

British Council Teaching for success: Continuing professional development (CPD) framework for teachersRetrieved fromwww.teachingenglish.org.uk2015

Coe, R., Aloisi, C., Higgins, S., & Major, L.E. (2014). What makes great teaching? Review of the underpinning research, 1–57. Retrieved from www.suttontrust.com/researcharchive/great-teaching/.

Coffey, S. (2018). Choosing to study modern foreign languages: Discourses of value as forms of cultural capital. Applied Linguistics, 39(4). doi:10.1093/applin/amw019.

Cohen, A.D. (2011). Strategies in learning and using a second language. Harlow, UK: Longman.

Colorado Education Initiative Student perception survey technical reportRetrieved fromhttp://www.coloradoedinitiative.org2013

Copland, F., & Neokleous, G. (2011). L1 to teach L2: Complexities and contradictions. ELT Journal, 65(3), 270–280. doi:10.1093/elt/ccq047.

Cordingley, P., Greany, T., Crisp, B., Seleznyov, S., Bradbury, M., & Perry, T. (2018). Rapid evidence review of subject- specific Continuing professional development in the UK. ((February)).

Cordingley, P., Higgins, S., Greany, T., Buckler, N., Coles-Jordan, D., & Crisp, B., et al. (2015). Developing Great Teaching: Lessons from the international reviews into effective professional development. Teacher Development Trust, 1–21. (Retrieved from educationnext.org).

Department for Education Languages programmes of study: Key stage 3 national curriculum in England (Issue DFE-00174-2013 3)Retrieved fromhttps://www.gov.uk/government/uploads/system/uploads/attachment_data/file/239083/SECONDARY_national_curriculum_-_Languages.pdf2013

Desimone, L.M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. Educational Researcher, 38(3), 181–199. doi:10.3102/0013189X08331140.
Dörnyei, Z. (2009). The psychology of second language acquisition. Oxford University Press.
Ellis, R. (2005). Principles of instructed language learning. System, 33(2), 209–224. doi:10.1016/j.system.2004.12.006.
Ellis, R. (2010). Second language acquisition, teacher education and language pedagogy. Language Teaching, 43(2), 182. doi:10.1017/S0261444809990139.
Ellis, V., & McNicholl, J. (2015). Transforming teacher education : Reconfiguring the academic work. Bloomsbury.
Evetts, J. (2011). A new professionalism? Challenges and opportunities. Current Sociology, 59(4), 406–422. doi:10.1177/0011392111402585.
Freeman, D. (2002). The hidden side of the work: Teacher knowledge and learning to teach. A perspective from north American educational research on teacher education in English language teaching. Language Teaching, 35(1), 1–13. doi:10.1017/S0261444801001720.
R.C. Furger L.E. Hernández L. Darling-Hammond California way the golden state's quest to build an equitable and excellent education system(February). Retrieved fromhttps://learningpolicyinstitute.org/product/california-way-equitable-excellent-education-system2019
Gillham, B. (2000). Developing a questionnaire. London: Continuum.
Graham, S. (2007). Learner strategies and self-efficacy : Making the connection. Language Learning Journal, 35(1), 81–93. doi:10.1080/09571730701315832.
Hall, G., & Cook, G. (2012). Own-language use in language teaching and learning. Language Teaching, 45(3), 271–308. doi:10.1017/S0261444812000067.
Horwitz, E.K., Horwitz, M.B., & Cope, J. (1986). foreign language classroom Anxiety. The Modern Language Journal, 70(2), 125–132. doi:10.1111/j.1540-4781.1986.tb05256.x.
Howes, A., Frankham, J., Ainscow, M., & Farrell, P. (2004). The action in action research: Mediating and developing inclusive intentions. Educational Action Research, 12(2), 239–258. doi:10.1080/09650790400200247.
Hutcheson, G., & Sofroniou, N. (1999). The multivariate social scientist. London: Sage.
Kaka, S.J. (2019). Classroom culture in the social studies classroom : The abilities of preservice teachers. Councilor: The Journal of Social Studies, 80(2), 1–18.
Kane, T., Kerr, K., & Pianta, R. (2014). Designing teacher evaluation systems : New guidance from the measures of effective teaching project. San Francisco, CA: Jossey-Bass.
Kennedy, A. (2015). "Useful" professional learning... useful for whom? Professional Development in Education, 41(1), 1–4. doi:10.1080/19415257.2014.983787.
Kennedy, M. (2016). How does professional development Improve teaching? Review of Educational Research, 20(x), 1–36. doi:10.3102/0034654315626800.
Lanvers, U., Hultgren, K., & Gayton, A.M. (2019). "People can be smarter with two languages": Changing anglophone students' attitudes to language learning through teaching linguistics. Language Learning Journal, 47(1), 88–104. doi:10.1080/09571736.2016.1196384.
Lightbown, P.M. (2003). SLA research in the classroom/SLA research for the classroom. Language Learning Journal, 28, 4–13. doi:10.1080/09571730385200151.
Macaro, E., Graham, S., & Woore, R. (2016). Improving foreign language teaching: Towards a research-based curriculum and pedagogy. New York: Routledge.
Moeller, A., & Nugent, K. (2014). Developing intercultural competence in the language classroom: 161 (pp. 1–18). Faculty Publications: Department of Teaching, Learning and Teacher Education. doi:10.1002/9781118784235.eelt0282.
Nation, I., & Macalister, J. (2010). Language curriculum design. Routledge.
Newby, D., Allan, R., Fenner, A.-B., Jones, B., Komorowska, H., & Soghikyan, K. (2007). The European portfolio for student teachers of languages. Council of Europe. Retrieved from http://archive.ecml.at/mtp2/fte/pdf/C3_Epostl_E.pdf.
OECD. (2017). How can professional development enhance teachers' classroom practices? Teaching in Focus, 16, 1–5. doi:10.1787/2745d679-en.
Ofsted (2015). Key stage 3: The wasted years? (p. 150106). Issue September.
Raudenbush, S.W., & Marshall, J. (2014). To what extent do student perceptions of classroom quality predict teacher value added? In Kane, T., Kerr, K., & Pianta, R. (Eds.), Designing teacher evaluation systems: New guidance from the measures of effective teaching project (pp. 170–202). San Francisco, CA: Jossey-Bass.
Schleicher, A. (2011). Building a high-quality teaching profession: Lessons from around the world. OECD Publishing. doi:10.1787/9789264113046-en.
Schleicher, A. (2016). Teaching excellence through professional learning and policy reform: Lessons from around the world. International Summit on the teaching profession. Paris: OECD Publishing. doi:10.1787/9789264252059-en.
Siminski, P. (2008). Order effects in batteries of questions. Quality and Quantity, 42(4), 477–490. doi:10.1007/s11135-006-9054-2.
Smith, S., & Conti, G. (2016). the language teacher Toolkit. CreateSpace Independent Publishing Platform.
Steinberg, M.P., & Garrett, R. (2016). Classroom composition and measured teacher performance: What do teacher observation scores really measure? Educational Evaluation and Policy Analysis, 38(2), 293–317. doi:10.3102/0162373715616249.
Stevens, J.P. (2002). Applied multivariate statistics for the social sciences (4th ed.). Hillsdale, NJ: Erlbaum.
Teaching Schools Council Modern foreign languages pedagogy review.https://www.tscouncil.org.uk/modern-foreign-languages-report/2016
Thompson, I. (Ed.). (2015). Designing tasks in secondary education : Enhancing subject understanding and student engagement. Abingdon: Routledge.
Tinsley, T., & Board, K. (2017). Language trends 2016/17: Language teaching in Primary and secondary schools in England. British Council. doi:10.1016/j.pcl.2008.07.007#.
Tinsley, T., & Doležal, N. (2018). Language trends 2018: Language teaching in Primary and secondary schools in England. British Council.
Tschannen-Moran, M., & Hoy, A.W. (2001). Teacher efficacy: Capturing an elusive construct. Teaching and Teacher Education, 17(7), 783–805. doi:10.1016/S0742-051X(01)00036-1.
Wilkerson, D.J., Manatt, R.P., Rogers, M.A., & Maughan, R. (2000). Validation of student, principal, and self-ratings in 360 degree feedback (registered) for teacher evaluation. Journal of Personnel Evaluation in Education, 14(2), 179–192. doi:10.1023/a:1008158904681.