Erasmus+
Enriching lives, opening minds.

In-service Teachers' Experiences with Virtual Exchange and their Attitudes Towards Implementing Virtual Exchange into the Curriculum

- A needs analysis report

INVITED

INVITED - Integrating primary and pre-school virtual exchange projects into language teacher education

Univerza v Ljubljani

Pädagogische Hochschule Freiburg

UNIVERSITY OF WARSAW

UNIVERSIDAD DE MURCIA

CONTENTS

INTRODUCTION **3**

METHODOLOGY **4**

RESEARCH QUESTIONS **6**

PROJECT REPORT **7**

RQ1: What is participants' opinion regarding virtual exchange (VE) and how – if at all – is it affected by their work experience, occupation and (previous) involvement in VE? 10
RQ2: What do the participants believe is the added value of VE in comparison to on-site teaching? 18
RQ3: What do participants not practically experienced in VE consider would be the most challenging aspects of conducting VE projects? How does this compare to reports of those with prior experience? 19
RQ4: How have teachers with previous experience regarding VE organised, designed and executed their projects? 24
RQ5: How do participants perceive their competence regarding VE projects? 38
RQ6: What would be teachers' training history regarding VE and their prospects about improvement? .. 42
RQ7: Which areas of VE projects would the participants like to develop professionally? 43
RQ8: How many participants plan to carry out VE projects in their future? 44

DATA PRESENTATION PER COUNTRY **47**

Added Value of VE 47
Most Challenging in Conducting VE 48
Involvement in VE projects, their averages and descriptions 49
Activities during VE projects 50
Frequency of Organisation Types 52
Preparation for VE sessions 53
Training about the development of VE 54
Carrying out virtual exchange projects with students 55

REFERENCES **57**

Introduction

Virtual exchange (VE) in language education is the process of communicating in a shared foreign language and collaboratively learning with peers from different countries through the use of technology (adapted from Dooly & Vinagre 2021, 393). The most common virtual exchange projects at primary and secondary level of education are carried out through the eTwinning platform, which is part of the Erasmus+ programme under the auspices of the European Commission. eTwinning provides support, tools and services to educational institutions (https://school-education.ec.europa.eu/en/etwinning). In 2019, the Council of the European Union adopted a recommendation calling for the support for foreign language learning of school-age children and the use of innovative teaching methods, including eTwinning (Renard & Milt, 2023). Pennock-Speck and Clavel-Arroitia (2022) argue that primary school pupils are rarely involved in various VE projects and therefore this area is under-researched. They assert that virtual exchanges increase motivation and interest in learning a foreign language (FL) and enhance students' intercultural awareness (Pennock-Speck & Clavel-Arroitia, 2022). Similarly, Nemiña et al. (2023) state that integrating the eTwinning platform into teaching practices could provide various benefits, such as increasing ICT skills, expanding learning opportunities and dialogue with peers in other contexts, cultural awareness and improvement of a FL.

Few studies have been carried out about the impact of virtual exchanges and eTwinning projects on foreign language learning, however, not many have investigated prospective FL (pre-)primary school teachers' views on virtual exchange. This report aims to fill this gap, adding to the body of research in this area. The study presented in this report was conducted transnationally, among 309 in-service (pre-)primary FL teachers in four different teaching contexts, i.e. Spain, Poland, Slovenia and Germany. The study is part of the ***INVITED Erasmus+ project (Integrating primary and pre-school virtual exchange projects into language teacher education)***, which is led by the Pädagogische Hochschule Freiburg and partners from the University of Warsaw, the University of Ljubljana and the University of Murcia. The main goals of the project are to promote the use of VE projects in primary and pre-school language education and to develop primary and pre-school teacher's competencies regarding VE in the language classroom. For this purpose, a questionnaire was developed to explore in-service teachers' experiences with VE and their attitudes towards implementing VE in the curriculum. We also wanted to enquire about their needs for this implementation in order to support them in the process of planning, organising and implementing VE into their future lessons.

Methodology

The study as such is conceptualised as needs analysis related to how ubiquitous and manifested is virtual exchange with in-service teachers of pre-school and primary education.
First phase began in December 2023 with questionnaire drafting, which was after several revisions launched in its final version on February 2nd 2024 and left active until May 21st of the same year.
It consisted of 40 questions that were realised in 165 variables, for which respondents required an average of 13 minutes to answer (median being 11 min 27 sec). This is in congruence with what was declared in the questionnaire's introduction. As per mode is concerned, we opted for online data gathering, made available by University of Ljubljana *Centre for Social Informatics's* online survey services – *One Click Survey* (1ka), available on https://www.1ka.si/.
The questionnaire was administered exclusively in English language, which we justify with the following reasons: (1) targeted teachers' profiles are related to language teaching; (2) virtual exchange as such stipulates the use of shared foreign language (which was in our case English); (3) multi-language translations can impact or change meaning; (4) the selected platform has an English user interface; (5) logistically more feasible as it did not warrant any translations and thus extra costs with potential time delays.
The adopted sampling strategy was non-probability *convenience*, which we deemed suitable due to our overall intent to investigate and discern established attitudes in a rather unfamiliar area. Although no tangible bias can be reported, we are aware that we have obtained merely those who were available at that time, which is why generalizability is arguable. Yet, because of time constraints and deadlines, funds allocation, relatively small researcher group and the global objective to even become acquainted with target audience's perceptions and experience, we were willing to risk representativeness which would otherwise be achievable with large-scale research incorporating random sampling.
The instrument was devised into 6 blocks [sections], in order of appearance these being: *understanding of virtual exchange, experiences with virtual exchange*, *problems with virtual exchange, competences concerning virtual exchange*, *needs regarding virtual exchange* and *demographic information*. Blocks that were presented to every respondent regardless of prior contact with virtual exchange were all but one connected with past experiences and problems, as the former served as a filter for inquiries related to modes of work, class organisation, resource use, etc. Questions combined dichotomous, multiple-choice, Likert-type scales, numerical scales and open format. Open questions were categorised in a multiethnic team of distinct area specialists, which reduced the likelihood of subjective interpretations, whilst coding was mainly closed. The reliability analyses are depicted in the table below; however, their summation is that the instrument can be assessed as of quality. Be that as it may, parts connected to respondents' experience could not be comprehensively verified for reliability due to relatively low valid counts (often ≤ 70) with simultaneous high number of items.

Scale	***Cronbach α***
Opinion regarding VE (before subscale division)	.612
Opinion subscale: Development of Language and Intercultural Skills and VE Variety	.727
Opinion subscale: Difficulties in VE Related to Learners' Age	.832
Opinion subscale: Students' Language and Technology Proficiency	.559*
Opinion subscale: Execution and Communication in VE	.416*
Problems with VE	.729
Resource use in VE	.769
Students' Development	.798
Self-assessed VE competence	.909

*Due to low reliability score, items composing the scale were treated independently

We can argue that we have a reliable instrument as α > .60. This criterion is violated in two instances, for two opinion subscales, which is explained in subsequent sections.

For validity verification, we can report favourable outcomes as we managed to design an instrument that covered a multitude of aspects related to virtual exchange, therefore be regarded as extensive and operational in terms of our purpose. Upon reviewing survey questions, data as such and research questions, our assessment is that we obtained what was envisaged. Consequently, this permitted us to underpin latent constructs (such as *competence*) and uncover relationships between variables.

Data analysis commenced in early-June 2024 and was finalised for in-service teachers in late-August 2024, although proofreading, editing and evaluation ensued. As evident from the results, we initially inspected data as a whole, then proceeded to country-based depictions, where pertinent research questions refer to part one. Analyses were computer-assisted, employing *MS Excel* and *IBM SPSS Statistics 28.0*.

Research Questions

RQ1: What is participants' **opinion regarding virtual exchange** (VE) and how – if at all – is it affected by their work experience, occupation and (previous) involvement in VE?

RQ2: What do the participants believe is the **added value of VE** in comparison to on-site teaching?

RQ3: What do participants not practically experienced in VE consider would be the **most challenging aspects** of conducting VE projects? How does this compare to reports of those with prior experience?

RQ4: How have participants with previous experience regarding VE **organised, designed and executed** their projects?

RQ5: How do participants **perceive their competence** regarding VE projects?

RQ6: What is participants' **training history** regarding VE?

RQ7: Which areas of VE projects would the participants like to **develop professionally**?

RQ8: How many participants plan to carry out **VE projects in their future**?

Project Report

Participants. Participants were 309 pre-primary and primary school teachers of various qualifications, employed in educational institutions across partners' corresponding countries (i.e., Germany, Poland, Slovenia, Spain and United Kingdom). The subsequent charts illustrate demographic data such as gender, age, country, years of professional service, etc. more vividly and scrupulously whilst supplementary comments are provided only for counts-abundant categories or to expound notable differences.

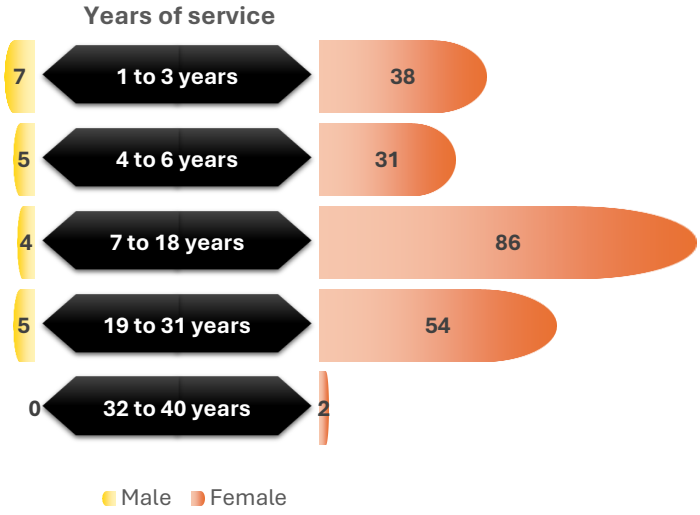

Figure 1: Participants by years of service and gender

Gender was not binary presented in the questionnaire as it is here, however, only 2 (0.8%) people opted for all other possibilities combined, which is why we have not depicted them in graphical form. Upon closer inspection, it was revealed that demographic variables suffered a tremendous drop-out rate (around 23%), which we have interpreted as a consequence of their end-of-survey positioning and nonessential question type. Yet, in instrument design, we have concluded that demographic information is often perceived as tedious and demotivating, which could affect the entire process. By placing them at the end, we have at least hypothetically avoided participants' disengagement.

Viewed separately without any other factor, sample distribution pertaining to gender reflects population characteristics, where around 10 % of the entire workforce are represented by male teachers (the share of male pre-school teachers is admittedly lower and hovers at around 5 %). In concrete terms, 22 respondents (9.3%) have identified as male and 213 (89.9%) as female, the majority of them (65.4%) having at least 7 years of working experience in pedagogical profession, even though the proportion of novice teachers (3 years or less of teaching experience) cannot be construed as negligible with 19.2% with an additional 15.4% of proficient teachers (with more than 3 but less than 7 years of service)[1]. Nevertheless, the

[1] 1-3 Novice teachers, 4-6 Proficient teachers, 7-18 Experienced teachers, 19-31 Veteran teachers, 32-40 Master teachers

scales are tipped towards older generations, indicating a relative lack of new personnel – this will be explored in later segments.

When asked about their qualifications, respondents were instructed to select whatever applied for them, so they could choose more than one answer. 237 people responded and provided us with 304 entries, where the chart below follows the percent of responses and not cases.

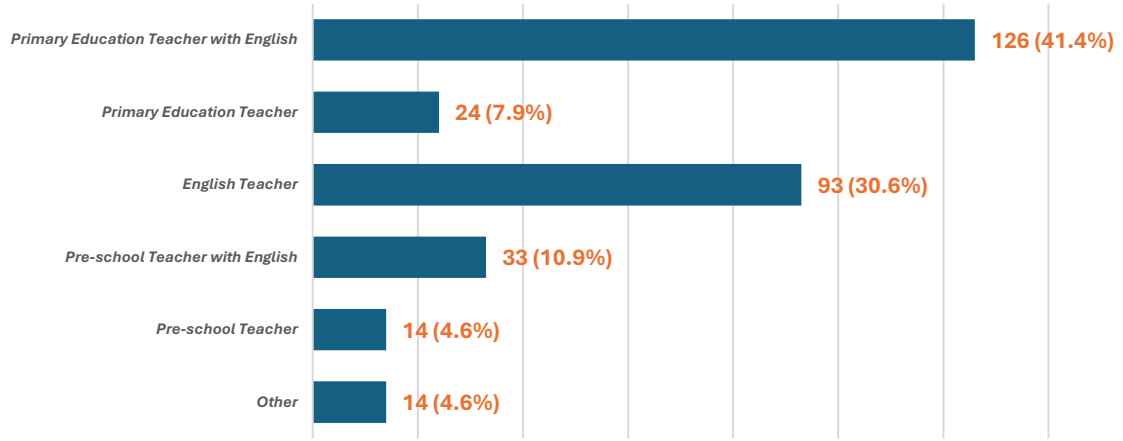

Figure 2: Participants' qualifications

We have incipiently expected a somewhat larger subsample of primary school than pre-school teachers as this is reflective of the population (the difference as such, not these exact proportions). If we profoundly simplify and take a calculated risk of reducing distinct data, we can clarify the present chart as though having 79.9% of teachers whose qualifications correspond with primary school environment and a 15.5% that pertain to pre-schools. When asked about their actual current employment in a follow-up inquiry, it was revealed that

Figure 3: Participants' employment

their qualifications coincide with their present-day occupation, meaning that their expected career trajectories onset by formal education did not veer of course. In other words, sample participants have a direct link [high correlation] between their educational background and existing profession.

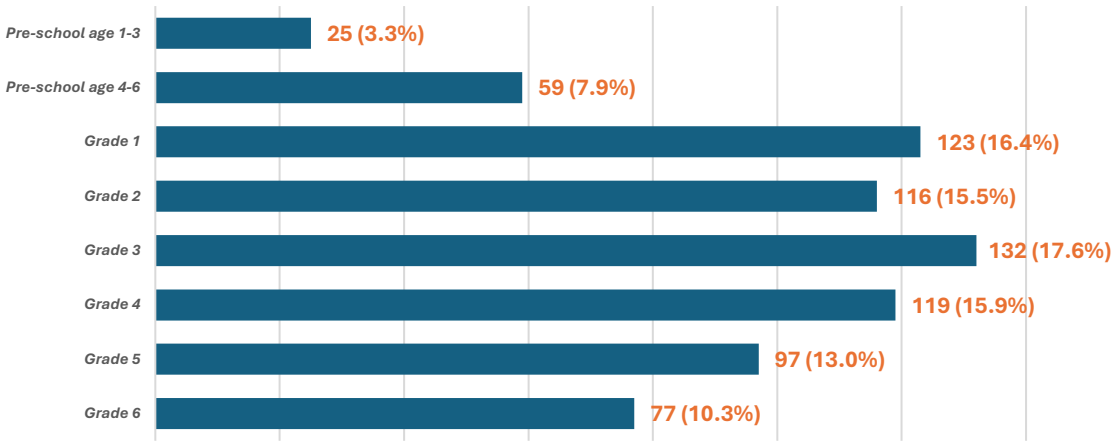

Figure 4: Grades in which participants teach

Without going into any superfluous descriptions, *Figure 4* indicates an overlap where teachers, on any level, teach in several grades. What could be emphasised and is not apparent from the chart is that these numbers are not only the results of teachers being present in several grades, but also teachers being active in different institutions. For example, 5.5% (11 out of 199) of primary school teachers report working with pre-school children ages 1-3 and a further 12.1% (24 out of 199) operate in the cohort of 4–6-year-olds. The reverse, i.e., pre-school teachers in primary schools, was not perceived as apparent in any higher grades and predominantly perceived as a rare exception.

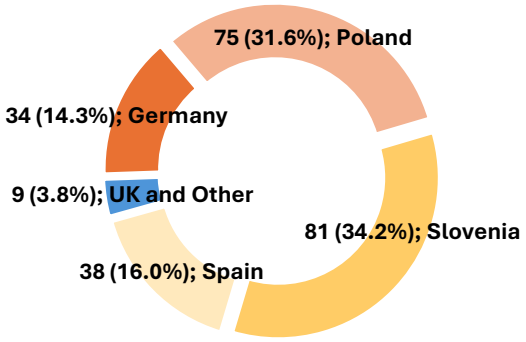

Figure 5: Participants by country

Lastly, the variable of country, depicted separately due to its significance for later analyses and overall impact it may have with its culture on interpretation. Poland and Slovenia have roughly twice the number of cases as Germany and Spain (approx. 30% vs. 15%) with UK and Other being more or less an honourable mention with 3.8%. Due to low counts, *UK and Other* will not be taken into consideration as a separate category.

RQ1: What is participants' opinion regarding virtual exchange (VE) and how – if at all – is it affected by their work experience, occupation and (previous) involvement in VE?

Opinion related to virtual exchange was measured with 18 5-point Likert scales, the first 9 being answered by 308 individuals and the remaining by 284.

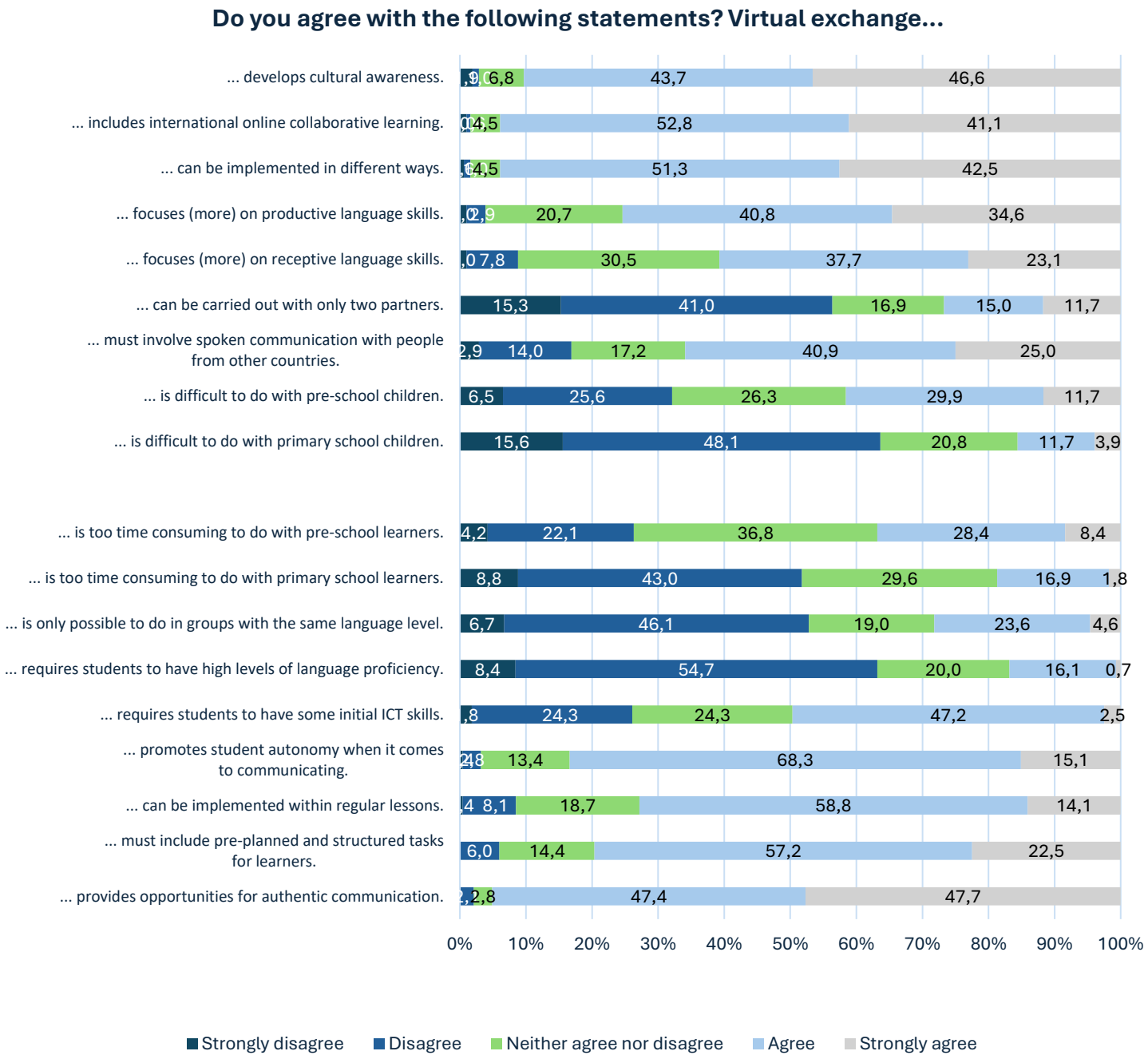

Figure 6: Participants' opinion regarding VE

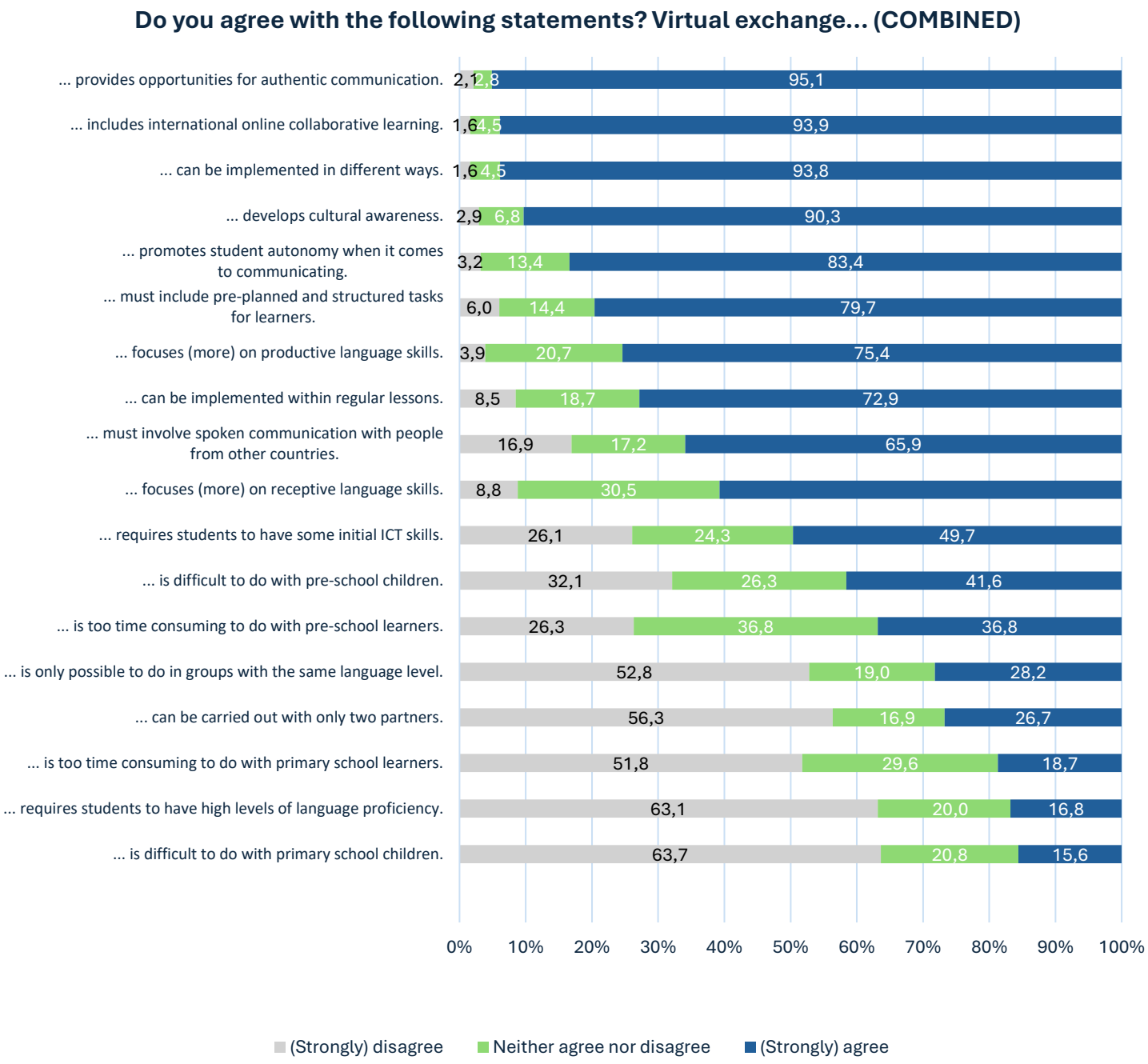

Figure 7: Participants' opinion regarding VE (combined)

We have opted for dimension reduction in terms of exploratory factor analysis (EFA) with principal axis factoring as extraction method and with varimax rotation. Sampling adequacy was met (KMO = .761; Bartlett's $\chi^2 = 1422.285$, $p < .001$), however, not all MSAs in the anti-image correlation matrix were above .5 (we had one value of .477, whilst every other was >.669). That problematic item was *"virtual exchange… can be carried out with only two partners"*, which was removed. Mind that this particular statement was questionable even with the larger sample (N = 501) where we had students as respondents, because of which

we did not hesitate to take such action. After this decision, EFA was launched again and revealed that our sample now fits established criteria.

Based on *eigenvalues*, a 4-factor solution was initially adopted, which explained 41.41% of total variance – altogether consistent with analyses done on student-given data. Yet, rotated factor matrix suggested a somewhat different factor loadings that was established in the study with pre-service students. For the sake of clarity, we provide the original schematic:

Factor 1: **Development of Language and Intercultural Skills with VE Variety**, comprised of:

…focuses (more) on productive language skills.
…focuses (more) on receptive language skills.
…includes international online collaborative learning.
…develops cultural awareness.
…can be implemented in various ways.
Factor 1 Cronbach' alpha for in-service teachers: .727

Factor 2: **Difficulties in VE related to Learners' Age**, comprised of:

…is difficult to do with pre-school children.
…is difficult to do with primary school children.
…is too time consuming to do with pre-school children.
…is too time consuming to do with primary school children.
Factor 2 Cronbach' alpha for in-service teachers: .832

Factor 3: **Students' Language and Technology Proficiency**, comprised of:

…is only possible to do in groups with the same language level.
…requires students to have high levels of language proficiency.
…requires students to have some initial ICT skills.
Factor 3 Cronbach' alpha for in-service teachers: .559

Factor 4: **Execution and Communication in VE**, comprised of:

…provides opportunities for authentic communication.
…promotes student autonomy when it comes to communicating.
…must include pre-planned and structured tasks for learners.
…can be implemented within regular lessons.
Factor 4 Cronbach' alpha for in-service teachers: .416

Dropped with pre-service sample (because of insufficient loadings onto one factor):

...can be carried out with only two partners.

...must involve spoken communication with people from other countries.

Figure 8: Exploratory factor analysis (EFA) with principal axis factoring (4-factor solution)

The color-coding above needs to be explained and is as follows: **BLUE** indicates complete congruence of both studies (i.e., pre-service and in-service), **LIGHT ORANGE** stands for items that did load together in both studies, but not necessarily in the exact same order, whilst **RED** is dedicated to items that had to be dropped in the sample of teachers.

Because pre-service teachers had a larger sample size, we decided to retain this structure, but have adhered to Factors 1 and 2 exclusively in the following procedures. Factor 4 did not achieve a satisfactory degree of reliability in both studies ($\alpha < .6$), whilst Factor 3 was only borderline acceptable in the present case. Moreover, comparability will be determined for those factors that are united by one latent concept and not prone to scrutiny.

		Development of Language and Intercultural Skills with VE Variety	Difficulties in VE related to Learners' Age
N	Valid	309	308
	Missing	0	1
Mean		4.155	2.824
Median		4.200	2.750
Std. Deviation		.0552	.8377
Minimum		1.00	5.00
Maximum		1.00	5.00

Figure 9: Comparison between Factor 1 and 2

Factor	Kruskal-Wallis				Dunn's post hoc test*	
	Mean ranks	χ^2	df.	p	Sample 1-Sample 2	adj. p
1	(n = 45) Novice teachers (1–3 years): 112.78 (n = 36) Proficient teacher (4–6 years): 127.60 (n = 91) Experienced teachers (7–18 years): 121.69 (n = 60) Veteran teachers (19–31 years): 108.03 (n = 2) Master teachers (32–40 years): **135.75**	2.730	4	**.604**	/	/
2	(n = 45) Novice teachers (1–3 years): **147.08** (n = 36) Proficient teacher (4–6 years): 93.86 (n = 91) Experienced teachers (7–18 years): 117.01 (n = 60) Veteran teachers (19–31 years): 112.69 (n = 2) Master teachers (32–40 years): 44.00	15.784	4	**.003**	Novice–Proficient	.004

*Showing only significant pairwise comparisons

Figure 10: Kruskall-Wallis and Dunn's post hoc test for Factor 1 and 2

Differences related to the *Development of language and intercultural skills with VE* variety that would stem from years of service were not ascertained (p = .604), whilst in case of *Difficulties in VE related to Learners' Age* we have found significant differences between classes of teachers (χ^2(H) = 15.784, df. = 4, p. = 0.003). Although somewhat anticipated, novice teachers report the highest concord with statements that relate to problems of VE ($\bar{R} = 147.08$) which we can attribute to the overall lack of experience which is customarily accompanied by an intensified perception of frustration and hardship. As years of service increase, these viewpoints reverse where the degree of agreement diminishes.

Even though qualifications were a more differentiated category, they were ultimately compressed into whether they work as pre-school or primary school teachers.

Factor	Mann-Whitney U Test					
	n	Mean Ranks	U	Z	p	effect size estimate
1.	Pre-school teachers: 38 Primary school teachers: 199	Pre-school: 130.88 Primary school: 116.73	3329.50	-1.175	**.240**	**r = - 0.07**
2.		Pre-school: 91.76 Primary school: 124.20	2746.00	-2.684	**.007**	**r = - 0.17**

Figure 11: Man-Whitney U Test for Factor 1 and 2

What we have determined was that differences between pre- and primary school teachers were present only for *Factor 2* which concerns difficulties and hurdles. Primary school teachers report significantly (p = .007) higher assent than their pre-school colleagues, however, the effect is rather small. We wondered if this could be explained by their involvement in VE projects, where we anticipated that there would be less of that with pre-school children, causing pre-school teachers not to be comprehensively aware of VE-related predicaments.

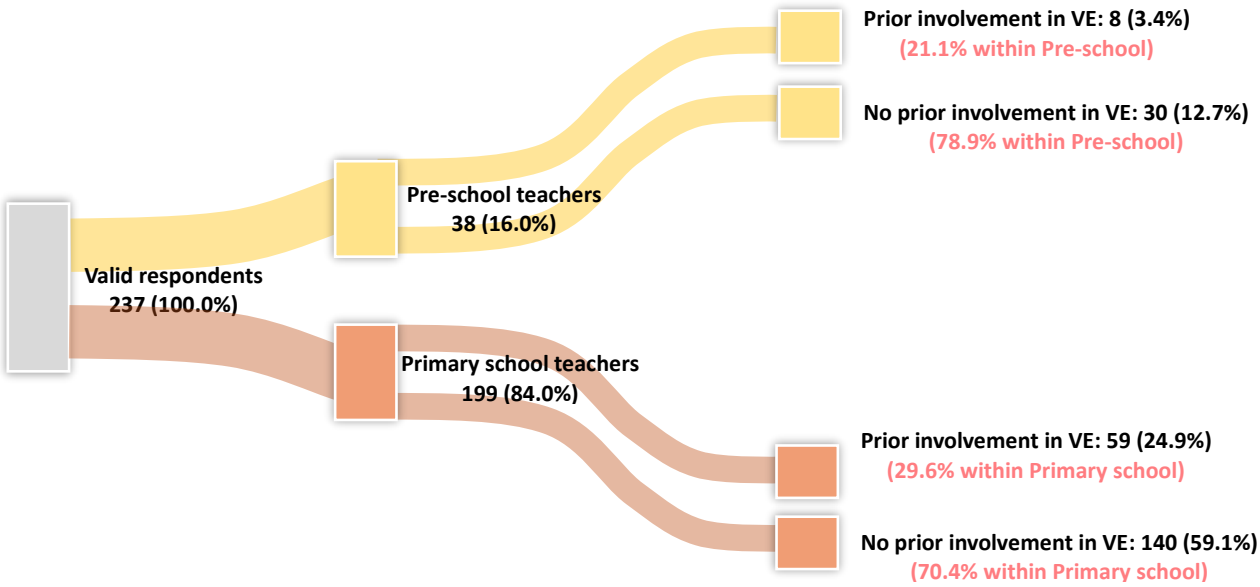

Figure 12: Participants' involvement in VE

First step was to check for differences between pre-primary and primary teachers regarding their (prior) involvement in any VE project. The graph above implies no conspicuous discrepancy as the proportion of those with any kind of experience is around 30% in both subsamples, nevertheless, chi-square test for independence was still performed. Results indicate (χ^2 = .777, df. = 1, p. = .378) that there are no significant differences between pre- and primary school teachers with regard to their involvement in VE projects, be that past or present. The reasons why more than 70% of participants have never been involved will be addressed promptly, yet before we accomplish this, we would like to present findings related to whether their involvement plays any role in how they perceive VE as such – and in doing so, we have preserved the division based on employment, which is shown in the two tables below.

Factor	Mann-Whitney U Test for *Pre-school teachers* (n = 38)					
	n	Mean Ranks	U	Z	p	effect size estimate
1.	Involvement in VE: 8 No involvement in VE: 30	Involvement: 18.81 No involvement: **19.68**	114.50	-0.200	**.841**	**r = - 0.03**
2.		Involvement: 15.69 No involvement: **20.52**	89.50	-1.102	**.270**	**r = - 0.18**

Figure 13: Mann-Whitney U Test for Pre-school teachers regarding their involvement in VE

Factor	Mann-Whitney U Test for *Primary school teachers* (n = 199)					
	n	Mean Ranks	U	Z	p	effect size estimate
1.	Involvement in VE: 59 No involvement in VE: 140	Involvement: **111.73** No involvement: 95.06	3438.00	-1.878	**.060**	**r = - 0.13**
2.		Involvement: 76.02 No involvement: **110.11**	2715.00	-3.831	**< .001**	**r = - 0.27**

Figure 14: Mann-Whitney U Test for primary school teachers regarding their involvement in VE

With pre-school teachers their participation in VE projects exhibited no effect on how they perceive VE, neither for *Development* nor *Difficulties*. However, involvement was deemed significant ($p < .001$) with primary school teachers and their apprehension of difficulties that encompass time requirements and learners' age-wise characteristics (i.e., *Factor 2*). Those without any contact with VE express higher degree of agreement ($\bar{R} = 110.11$) with corresponding statements than those teachers who cooperate(d) in VE ($\bar{R} = 76.02$). The effect size ($r = -.27$) of this interpretation is borderline medium, thus in no way negligible. If we rephrase; teachers who confess to having no experience with VE, believe it to be (too) time consuming and difficult to execute with pupils, or at the very least, think it to be such more prominently than their colleagues who were actually engaged in it.

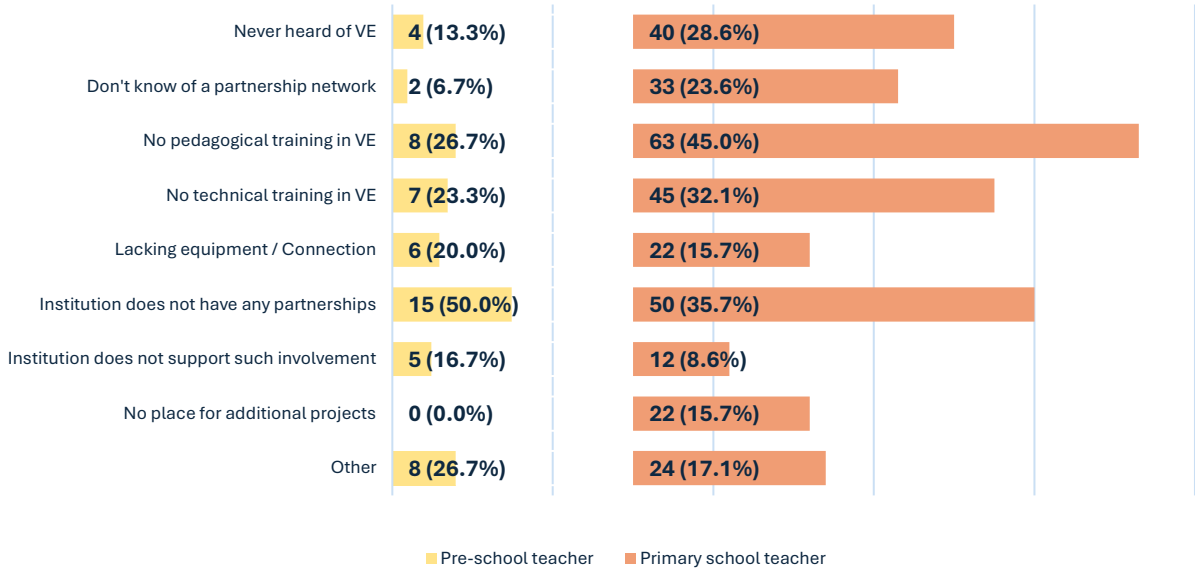

Figure 15: Participants' reasons for non-involvement in VE

As to why they have never been involved, we present the responses of 30 pre-primary and 140 primary school teachers who were asked to report everything that might be applicable. Significant differences have been found between these two groups (χ^2 = 22.771, df. = 9, p = .007), with the most noteworthy discrepancy pertaining to the reasons of *don't know of a partnership network* and *no place for additional projects* which were reported more frequently by primary school teachers. Both groups, however, state that the absence of pedagogical training in VE and non-existent institutional partnerships are the factor that kept them from being engaged in VE activities.

RQ2: What do the participants believe is the added value of VE in comparison to on-site teaching?

We presented this question completely opened as we refrained from confining participants in any way. In return, we received 213 responses which were sorted into 10 distinct categories (content suitability was achieved through researcher triangulation, whilst non-English answers were translated into target language by native speakers).

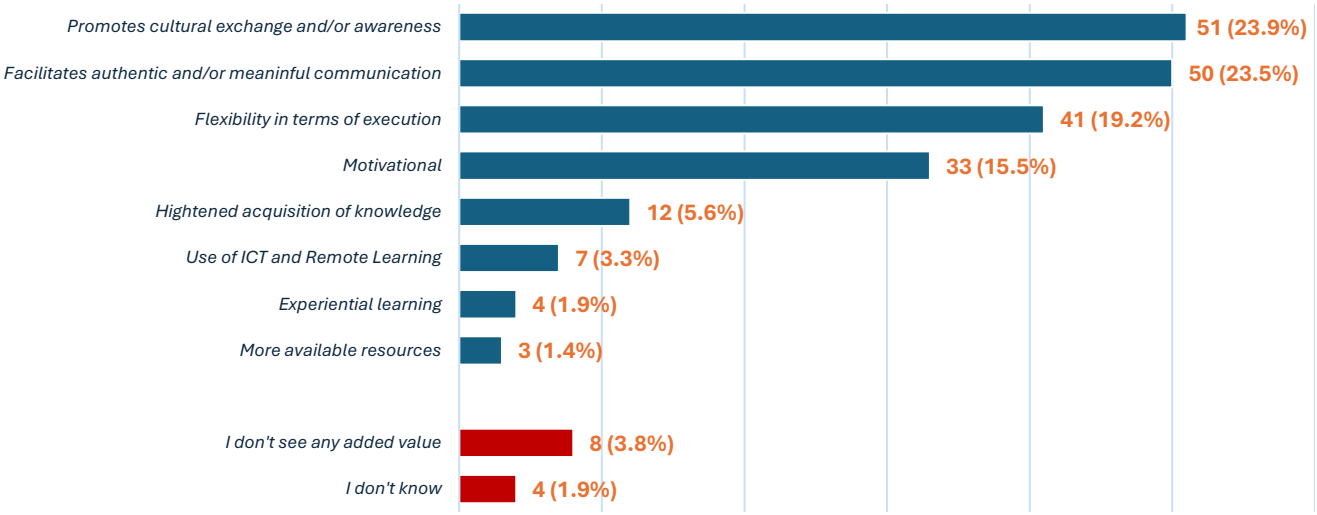

Figure 16: Participants' perceptions of VE's added value

Promoting cultural exchange and *authentic and meaningful communication* are positioned as foremost aspects (approx. 24% each) in attributing added value to virtual exchange with *flexibility in execution* (approx. 20%) as an immediate runner-up. 5.7% of participants either do not see any added value or are unaware of it; alternatively, 94.3% of teachers were in various extents able to justify virtual exchange as having some added value in teaching.

Furthermore, we decided to ascertain whether segments listed below affect perceptions related to virtual exchange's added value. Differences based on:

gender have proved to be **non-significant** (2î = 6.034, df. = 9, p = .737)

country have proved to be **non-significant** (2î = 47.016, df. = 45, p = .390)

years of service have proved to be **non-significant** (2î = 35.666, df. = 36, p = .484)

occupation have proved to be **non-significant** (2î = 15.651, df. = 9, p = .075)

involvement in VE have proved to be **non-significant** (2î = 10.966, df. = 9, p = .278).

RQ3: What do participants not practically experienced in VE consider would be the most challenging aspects of conducting VE projects? How does this compare to reports of those with prior experience?

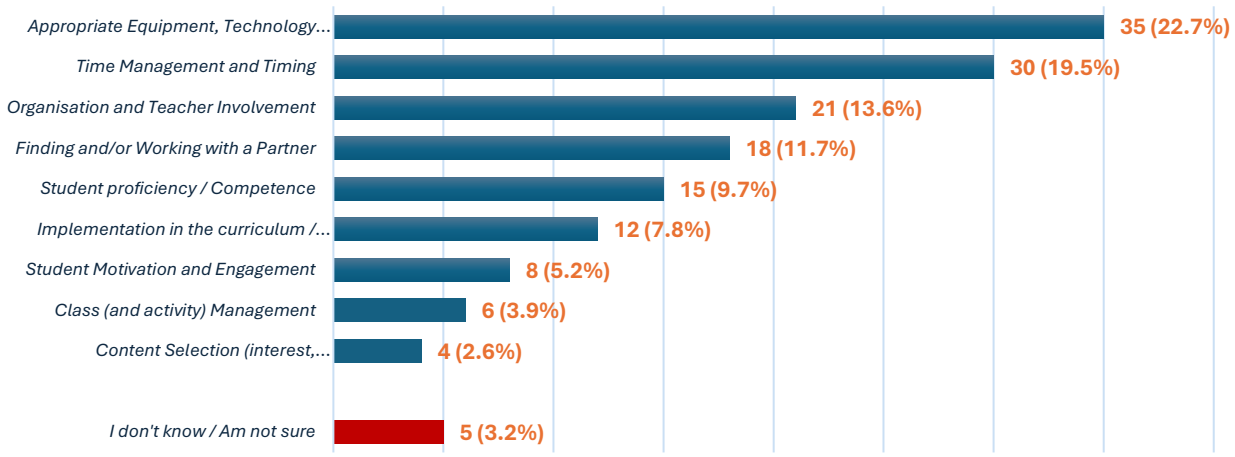

Figure 17: Participants' perceptions of VE challenges

Sample data suggests that when viewed in isolation, study participants recognize *equipment and technology* (in 22.7% of cases) as most challenging with *time management* as a relatively close second (19.5%). However, note that we have asked "*What do you consider would be most challenging in conducting virtual exchange projects?"* only those who previously stated that they have no experience. Yet, if we previously managed to obtain and interpret answers as to why they have never been involved, we could now further contemplate about possible impacts on their impressions.

In continuation we have focused primarily on those who claimed ***they have never heard of VE* (n = 45)**, legitimately serving as the prototypical group of people who were not involved on any level (i.e., practical, contemplative, administrative) and for whom we can thus logically assume that they have become acquainted with VE through our questionnaire (which provides definitions and examples as to what VE alludes).

Participants who *NEVER HEARD OF VE* and their opinion regarding its challenges

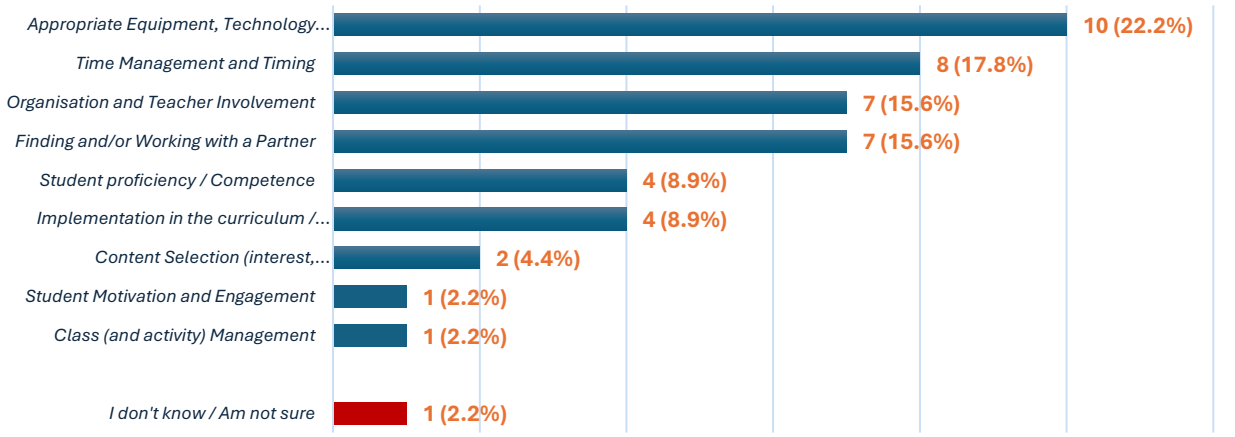

Figure 18: Participants not familiar with VE and their opinion regarding its challenges

Admittedly, the frequencies were low, so this is purely informative in nature, yet, the message we can extract – albeit tentatively – is rather intense. Even though they had no previous conceptions about VE, these selected teachers either surmised from the questionnaire, inferred from similar occasions when they employed digital aids and ICT as such (what immediately came to mind was COVID19 period when they were practically restrained to computer-assisted and enabled teaching and tutoring) or merely guessed that *maintaining equipment, technology and connection* would present itself as the most burdensome dimension of VE.

Now, onwards to those who ***actually were or still are involved in virtual exchange (n = 66)*** and which problems have they encountered. We have designed 12-item list detailing associated complications in advance and asked them to evaluate their occurrence on a 5-point scale ranging from *always* to *never*.

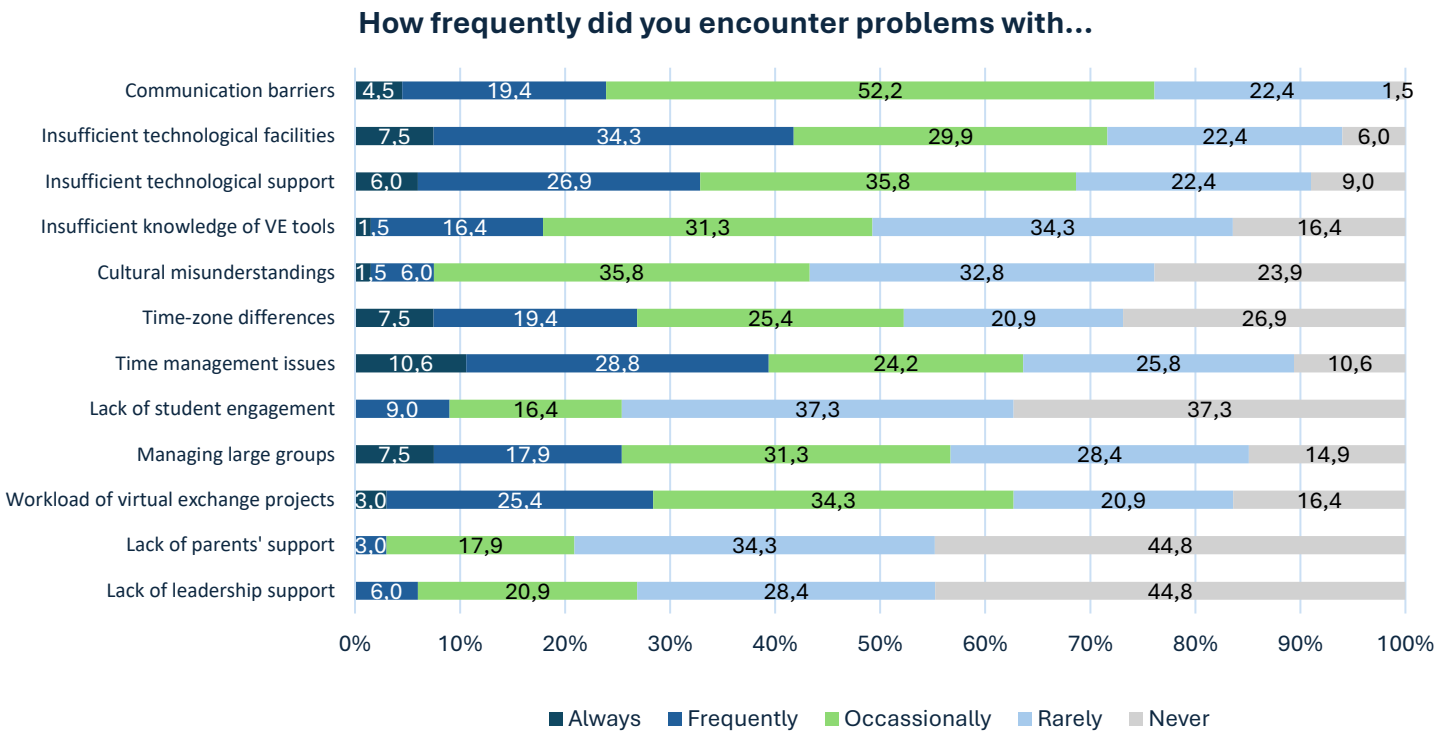

Figure 19: Frequency of participants' encounter with problems regarding VE

Merely for easier interpretation we merged the two sets of options that detail either regular (i.e., *always + frequently*) or scarce (i.e., *rarely + never*) incidence, whilst every additional analysis is done on non-reduced data with all five levels. The percentages are arranged in a descending order based on the degree of regularly befalling problems.

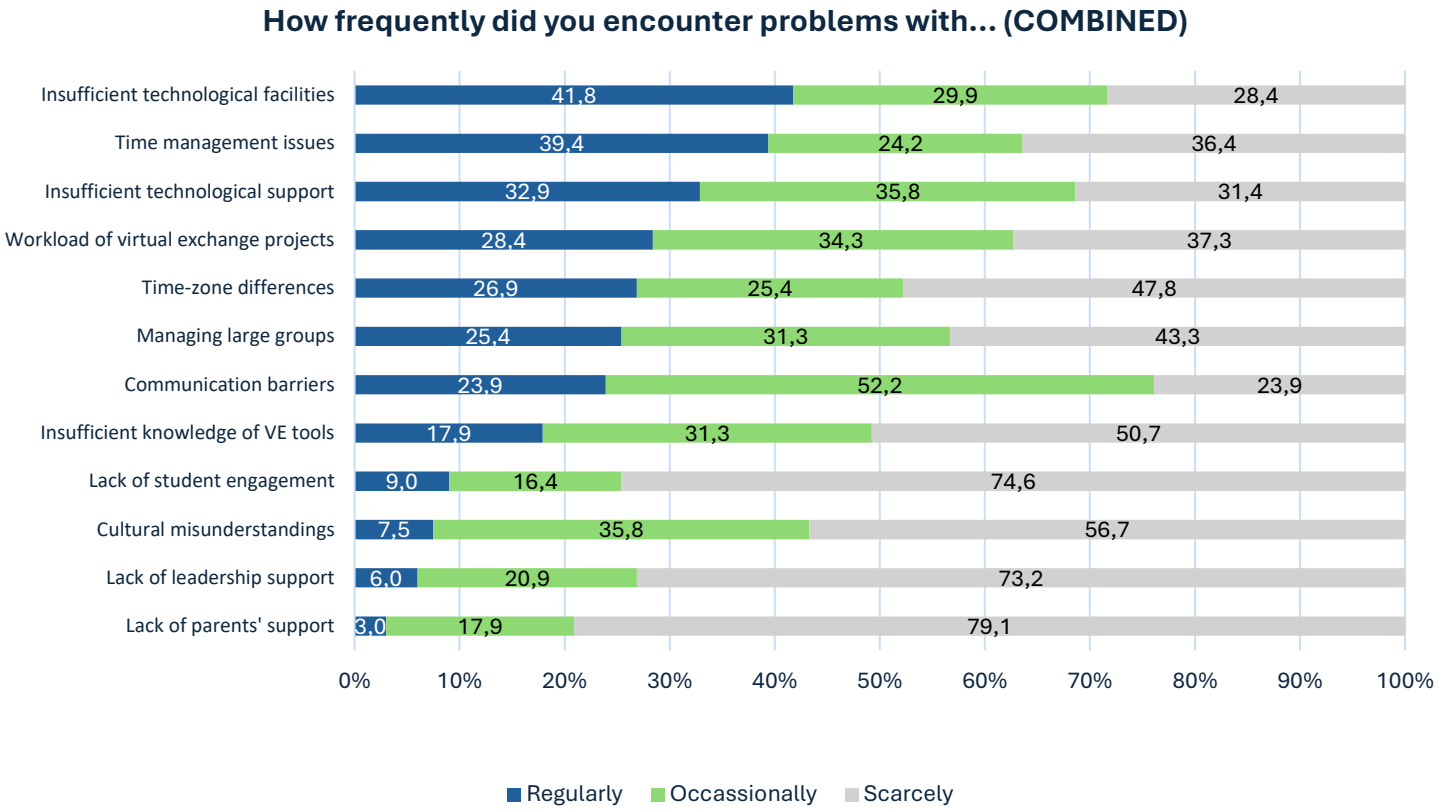

Figure 20: Combined frequency of participants' encounters with problems regarding VE

It is rather obvious that *insufficient technological facilities*, under which we specified poor internet connection, old hardware, and/or lack of computers, is declared as the most recurring problem with 41.8% participants claiming that they confront such issues on a regular basis when involved in a VE project. With 39.4% this was followed by *time management issues*, whilst third place goes to *insufficient technological support* which was selected as most troublesome by 32.9% of teachers. Note that this arrangement is designed by accentuating "regular" occurrence, which we emphasise due to its importance of ascertaining what would be virtual exchange's predominant predicament.

Interestingly, if we compare merely these three most frequently listed obstacles, those who have never been involved in VE as well as those who have some experience both determine *technical difficulties* as most problematic aspect of VE, whilst significant additional strain is caused by issues related to *time management*.

	Friedman's Test (n = 66)				Pairwise comparisons*	
Problem	Mean Ranks	Median	χ^2	p	Sample 1–Sample 2	p
Communication barriers (CB)	4.70	3.00	171.076	< 0.001	IK – ITF	.032
Insufficient technological facilities (ITF)	4.54	3.00			Cultural misunderstandings – ITF, CB, TMI, ITS	all four ≤ .016
Insufficient technological support (ITS)	5.08	3.00			Lack of student engagement – ITF, CB, TMI, ITS, MLG, workload of VE	all six ≤ .005
Insufficient knowledge of VE tools (IK)	6.73	4.00			Lack of leadership support – ITF, CB, TMI, ITS, MLG, workload of VE, time zone differences, IK	all eight ≤ .050
Cultural misunderstandings	7.38	4.00			Lack of parents' support – ITF, CB, TMI, ITS, MLG, workload of VE, time zone differences, IK	all eight ≤ .026
Time-zone differences	6.43	3.00				
Time management issues (TMI)	5.05	3.00				
Lack of student engagement	8.44	4.00				
Managing large groups (MLG)	5.90	3.00				
Workload of virtual exchange projects	5.97	3.00				
Lack of parents' support	8.95	4.00				
Lack of leadership support	8.84	4.00				

*Due to having 66 $\binom{12}{2}$ possible combinations, we decided to display just 27 comparisons which had significant adjusted values.

Figure 21: Participants' reported problems regarding VE

Firstly, we have coded these 12 "problems" from 1 (i.e., always) to 5 (i.e., never), which means that the lower the average, the higher the incidence rate of a specific item. Moreover, the data focuses exclusively on teachers with experience, who assess specific periodicities significantly different (χ^2(F) = 171.076, df. = 11, p < 0.001), with *insufficient technological facilities* ($\bar{R} = 4.54$), *communication barriers* ($\bar{R} = 4.70$), *time management issues* ($\bar{R} = 5.05$) and *insufficient technological support* ($\bar{R} = 5.08$) being significantly lower scored than most other items (see detailed list of comparisons in the table itself and further explanation provided with the asterisk). To measure effect size, we calculated Kendall's coefficient of concordance, whose result was .24 indicating moderate agreement of ordinal assessments made by multiple teachers. In practical terms, these four problems explicitly listed above have been reported by said teachers as most prevalent, whilst those issues related to various lacks in support and cultural misunderstanding were found to be rather singular and uncommon.

We combined these into one variable, clearly indicating that their content relates to VE problems. Additionally, we confirmed our approach with a reliability analysis, where a Cronbach's α score of .729 showed that these 12 items can be considered a group with acceptable internal consistency. The reason for this is rather transparent – we wanted to verify

if the overall problem-related-score is impacted by segments such as gender, country, years of service, occupation, and number of projects.

Variable	Kruskal-Wallis				Dunn's post hoc test*	
	Mean ranks	χ^2	df.	p	Sample 1-Sample 2	adj. p
Gender	(n = 7) Male: 36.00 (n = 59) Female: 34.10 (n =1) Prefer not to say: **14.00**	1.113	2	**.568**	/	/
Country	(n = 11) Germany: 34.09 (n = 19) Poland: 29.00 (n = 17) Slovenia: 36.56 (n = 15) Spain: **27.03**	2.858	3	**.414**	/	/
Years of service	(n = 3) Novice teachers (1–3 years): **19.67** (n = 6) Proficient teacher (4–6 years): 39.25 (n = 29) Experienced teachers (7–18 years): 38.90 (n = 27) Veteran teachers (19–31 years): 28.15 (n = 1) Master teachers (32–40 years): 28.50	6.577	4	**.160**	/	/
Occupation	(n = 8) Pre-school teacher: 45.50 (n = 59) Primary-school teacher: **32.44**	3.175	1	**.075**	/	/
Number of projects	(n = 44) VE Beginner (1–3 projects): 33.06 (n = 12) VE Intermediate (4–8 projects): **32.63** (n = 4) VE Advanced (9–15 projects): 42.63 (n = 7) VE Master (more than 15 projects): 37.36	1.158	3	**.763**	/	/

Figure 22: Participants' reported problems regarding VE analysed by segments of gender, country, years of service, occupation and number of projects

We were not able to find any significant differences in the overall scores related to problem assessment based on selected independent variables. Although we have highlighted the lowest mean ranks, these are in most cases comparable even on the sample level.

What might be worthy of some attention but still construed with caution, namely due to extremely small subsamples, is that participating novice and primary school teachers apparently record the highest rate of problems or tend to experience them more frequently than other groups.

RQ4: How have teachers with previous experience regarding VE organised, designed and executed their projects?

Everything that relates not merely to whether they had any contact with VE, but to genuine collaborations with partners, participant characteristics, activities, developed skills, used resources, etc. originated from the (sub)sample of 87 teachers who initially professed they have been involved in virtual exchange. However, this size unfortunately dwindles as we progress.

When asked about the **number of projects**, the range was between 1 and 200 projects per person, the mode being on 1 project (24 out of 84 participants or 28.6%). Admittedly, 200 projects were reported by just one person, without whom the maximum value fell to 30.

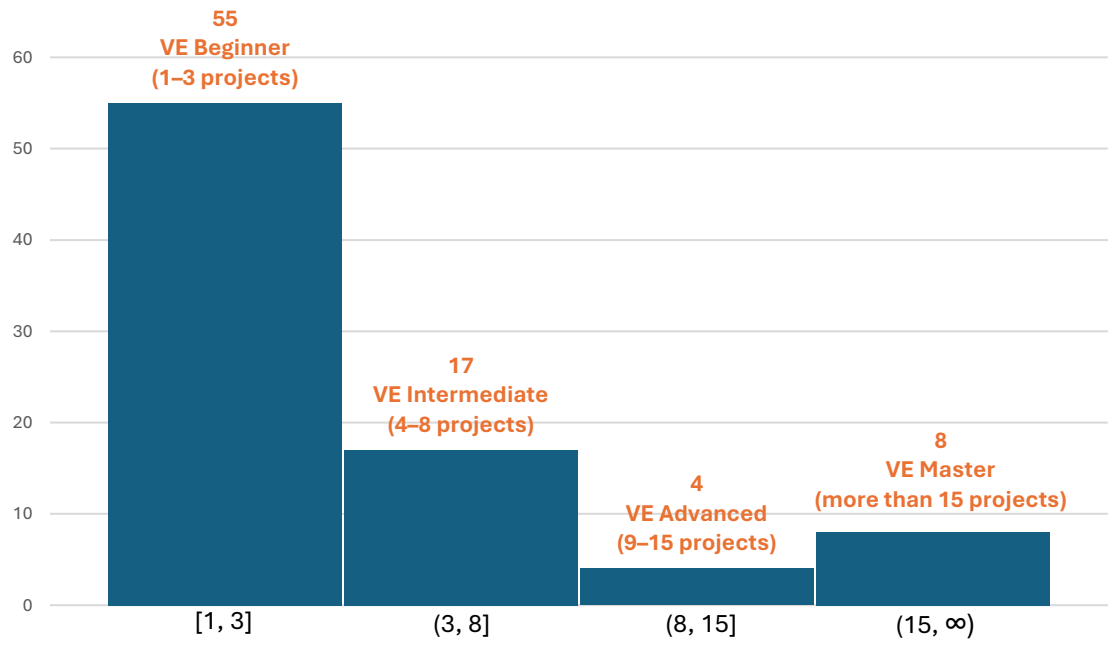

Figure 23: Participants' numerical involvement in VE projects

Based on their numerical answers we have recoded them into four groups as exhibited in the histogram above. Vast majority of teachers (55 out of 84 or 65.5%) belong to the group of "beginners" who are defined by participating in no more than 3 projects. Seeing that the second placed group is that of "intermediates" (17 out of 84 or 20.2%), who completed a maximum of 8 projects, we cannot claim that we were dealing with expert practitioners, but neither with amateurs, thus somewhere in between. We have based this division on existing explorations of virtual exchange and eTwinning (see EVOLVE Project Team, 2020; Lenc et al., 2016; Pettenati et al., 2021), where they speak of *levels* which are conditioned by carried out projects. Additionally, it was established that on average a virtual exchange project is executed in around one school semester. While this might seem unrelated to our survey, it is actually

very relevant. Specifically, when describing the sample, we noted a higher number of teachers with more years of service, meaning they are generally older. To see whether an increase in years of service translates into more projects, we have – due to non-normally distributed values – conducted a Kruskal-Wallis H test. Results indicate ($\chi^2(H) = 10.764$, df. = 4, $p = 0.029$) that there are significant differences between in-service teachers with varying years of service with regard to the number of VE projects. Experienced teachers (7–18 years) have the highest mean rank ($\bar{R} = 37.28$), followed by their "veteran" (19–31 years of experience) colleagues with $\bar{R} = 34.89$. Although it is rather expected that novice teachers first need to find their footing when they are at the start of their careers and are less likely to organise something that is outside the curriculum, what this concurrently alludes to is that virtual exchange with its appropriation of ICT is not something that would be in the exclusive domain of younger generations.

When asked about the nature of said projects, teachers **described** them without any prompts, which is why not all answers are on the same level of construction. That being said, they denoted them as:

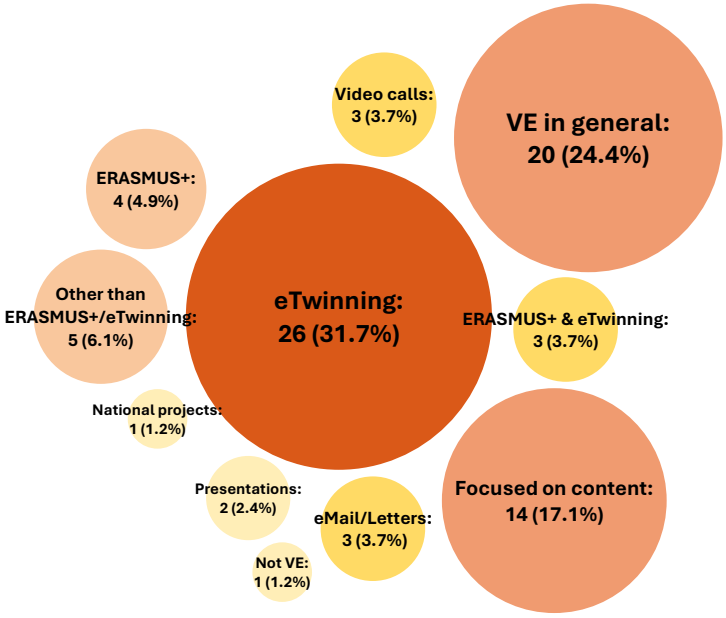

Figure 24: Participants' description of VE projects they were involved in

A majority (26 out of 82 or 31.7%) of projects was described as conducted on the eTwinning platform, whilst the latter appeared in combinations with ERASMUS+ as well (an additional 3.7%). Further strong categories related to VE in general (24.4%) and projects which participants defined they had an emphasis on content (17.1%). As other answers have less than 10 valid answers, we are not describing them in detail.

It was of interest to us how teachers came in **contact with their project partners**, where options were prepared for them, however, they were able to add their own response if nothing fitted. Moreover, they could opt for multiple answers.

Due to distinct institutional dynamics, the data is presented separately for pre-primary and primary teachers (thus, percentages are calculated intra-group), whilst whole sample data is supplied in the last column. Be aware that the latter differs from that of subsamples as not all who have answered about contact with partners shared their occupation.

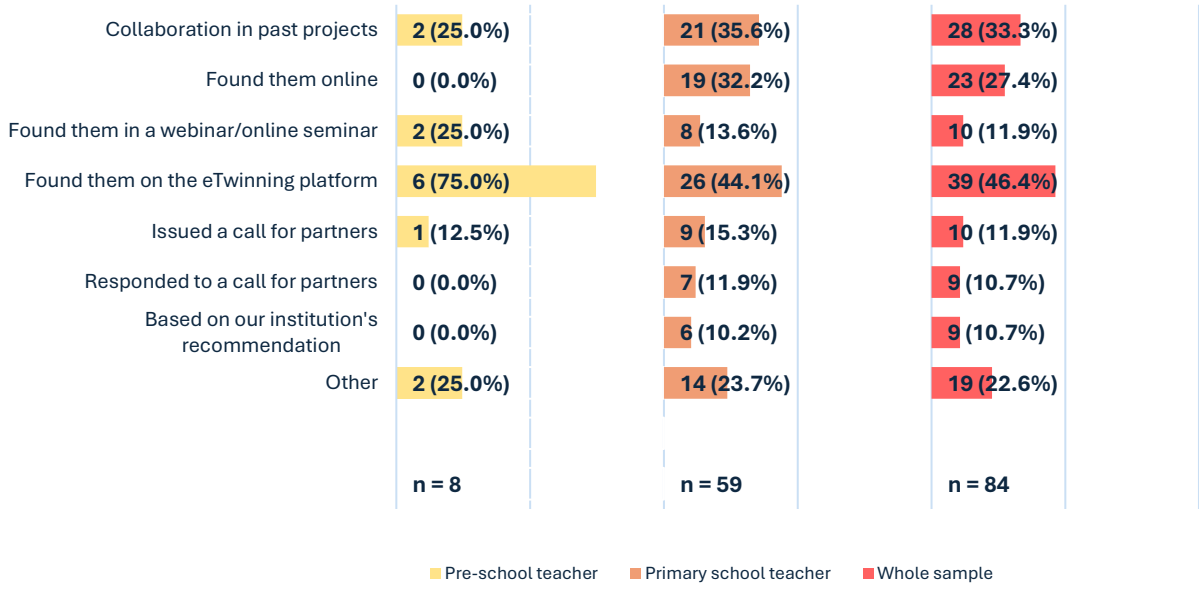

Figure 25: Participants' contact with their project partners

Let us start by saying that significant differences in column proportions were not found for any option (verified with z-tests and/or chi-square test for equality of proportions) between pre-primary and primary teachers, but both groups detail the eTwinning platform as the main route to partners (46.4% of respondents). Given that their own responses make up nearly a fourth of all answers (22.6% to be exact), the relevant analysis is:

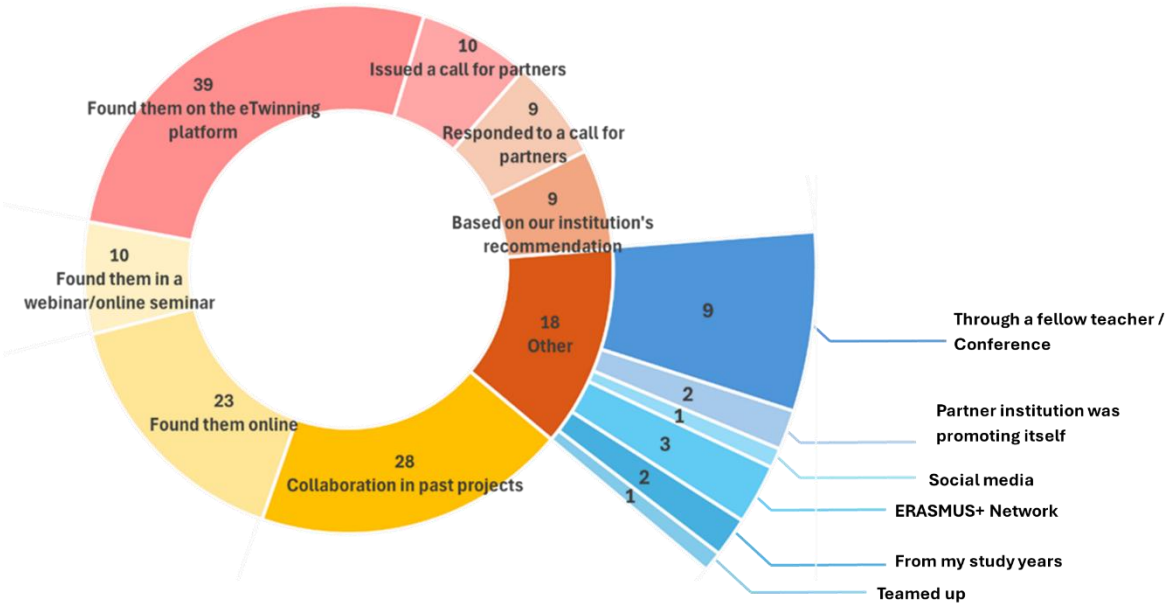

Figure 26: Participants' contact with project partners (whole sample)

We are not offering percentages here as they can be misleading acknowledging that the superordinate category (i.e., *other*) only had 18 counts. However, it appears that several teachers became acquainted with their partners through their own social networks, via colleagues, conferences, symposiums. Among other notable mentions we can include ERASMUS+ network as well.

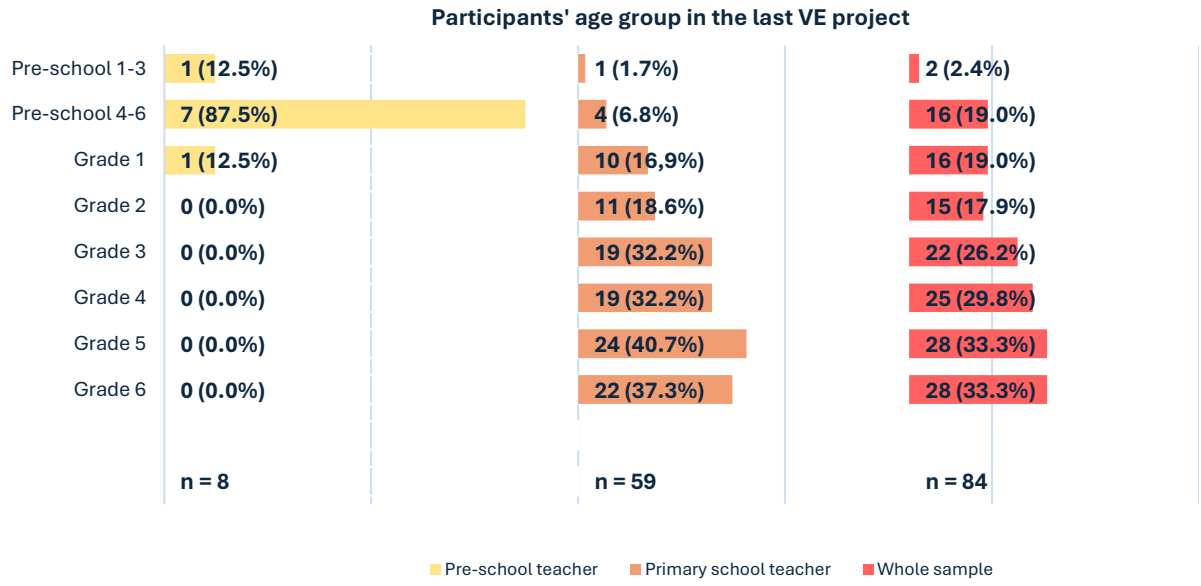

Figure 27: Participants' age group in the last VE project

Although we cannot claim anything for pre-primary teachers as we are severely restricted by their sample size, it might appear that they prefer having VE projects with their older pre-schoolers. Same applies for primary teachers – as students' age increases, so does the number of projects.

Next item on the agenda was to determine how teachers decide on the **content of their projects**.

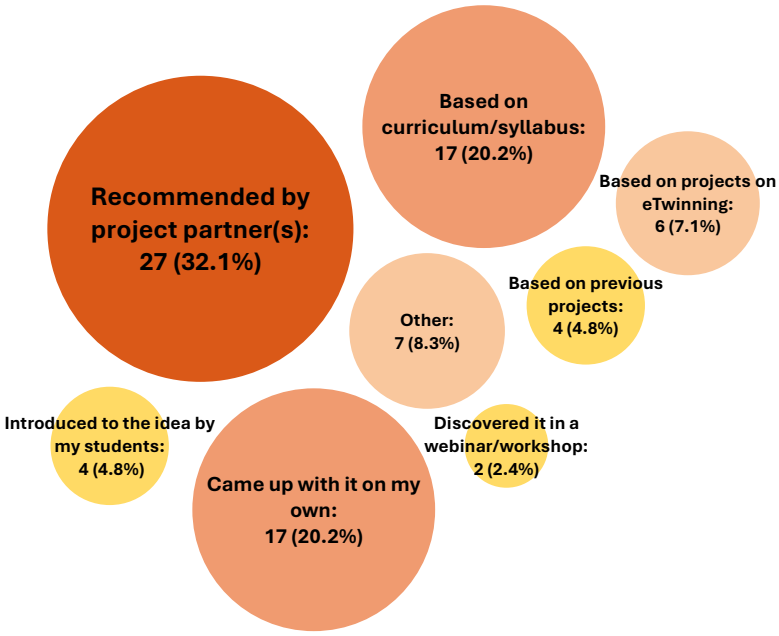

Figure 28: Decision regarding content of participants' VE projects

Almost a third (27 out of 84 or 32.1%) of teachers claim they were acquainted with the content by their partners with a good fifth (17 out of 84 or 20.2%) saying it was their original idea and an additional 20.2% reporting they had established it on curriculum or syllabus.

When asked how often students met online per project per month, the participants answered they met on average 3.98 times (95% CI between 2.45 and 5.50) with 88.9% teachers requiring equal or less than 5 appointments.

To comprehensively investigate how projects were implemented, we scrutinized *the structure of their online meetings*, *organisation, resource use*, *student preparation* and *activities*. Before we comment on these segments, we exhibit relevant data in ensuing charts in hopes they will ease understanding.

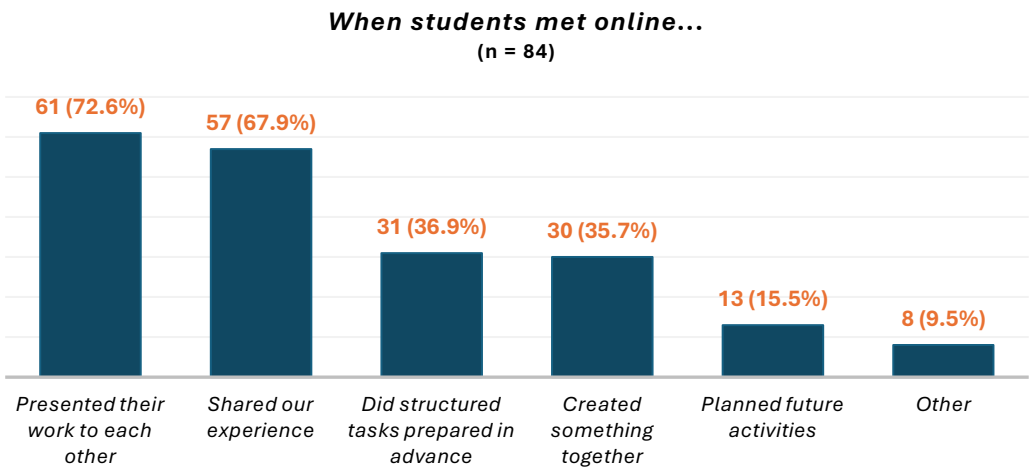

Figure 29: Content/nature of students' online meetings in VE projects

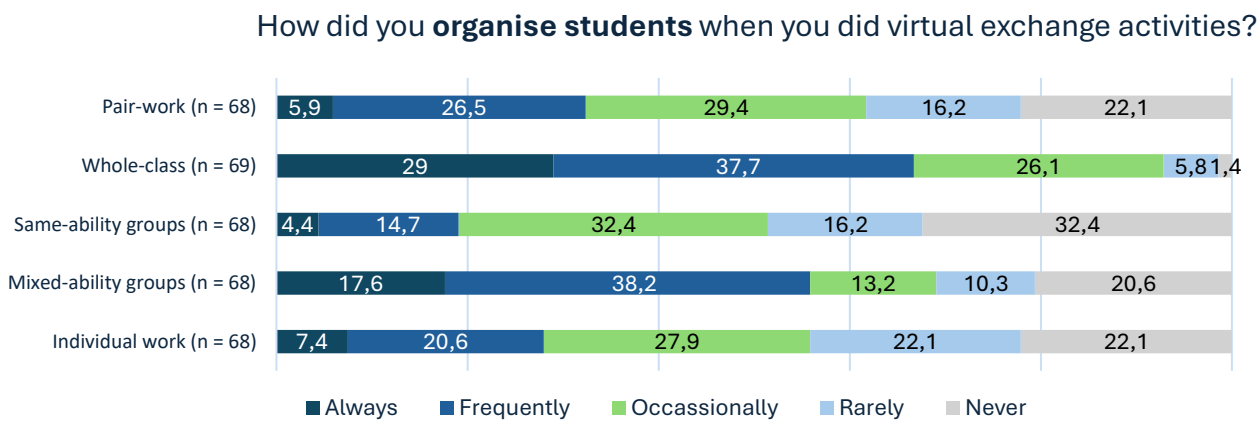

Figure 30: Organisation of students in VE activities

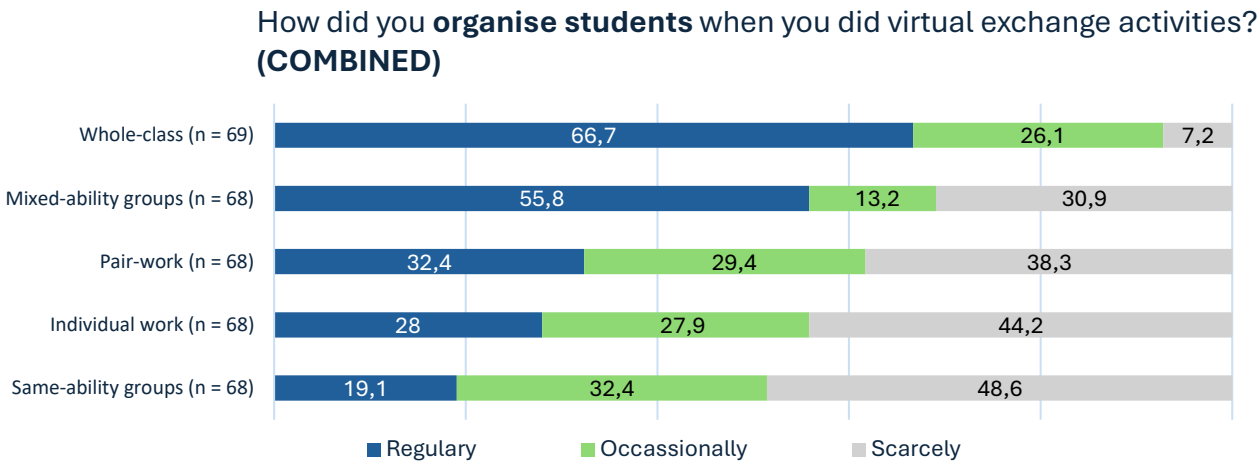

Figure 31: Organisation of students in VE activities (combined)

In terms of organisation, *whole-class* endeavours are preferred amongst teachers as their implementation is at least regular in 66.7%, and at least occasional in 92.8% of interactions. Significant differences were found in frequency of organisation types ($\chi^2(F) = 46.456$, df. = 4, $p < .001$) where Bonferroni corrections for multiple tests revealed that *whole-class* ($\bar{R} = 2.23$) is more recurrent than *pair-work* ($\bar{R} = 3.12$), *individual work* ($\bar{R} = 3.35$) and *same-ability groups* ($\bar{R} = 3.67$). In other words, significantly more prevalent (in all cases $p < .010$) than all other types but *mixed-ability* ($\bar{R} = 2.63$), whose only significant difference is with *same-ability*. Note that *mixed-ability* is regularly organised in 55.8% of activities and in 69.0% if we add occasional employ.

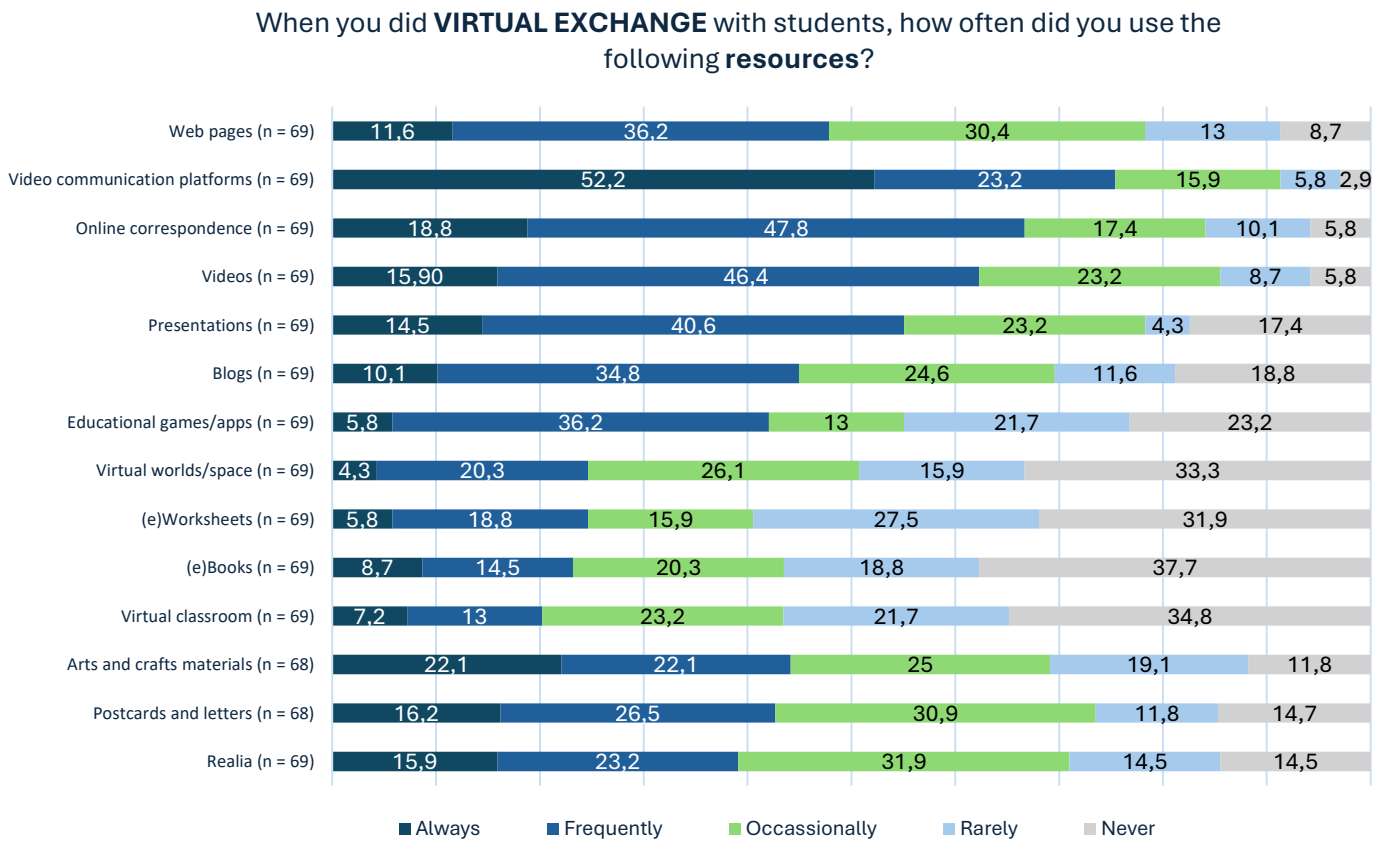

Figure 32: Frequency of using listed resources in VE

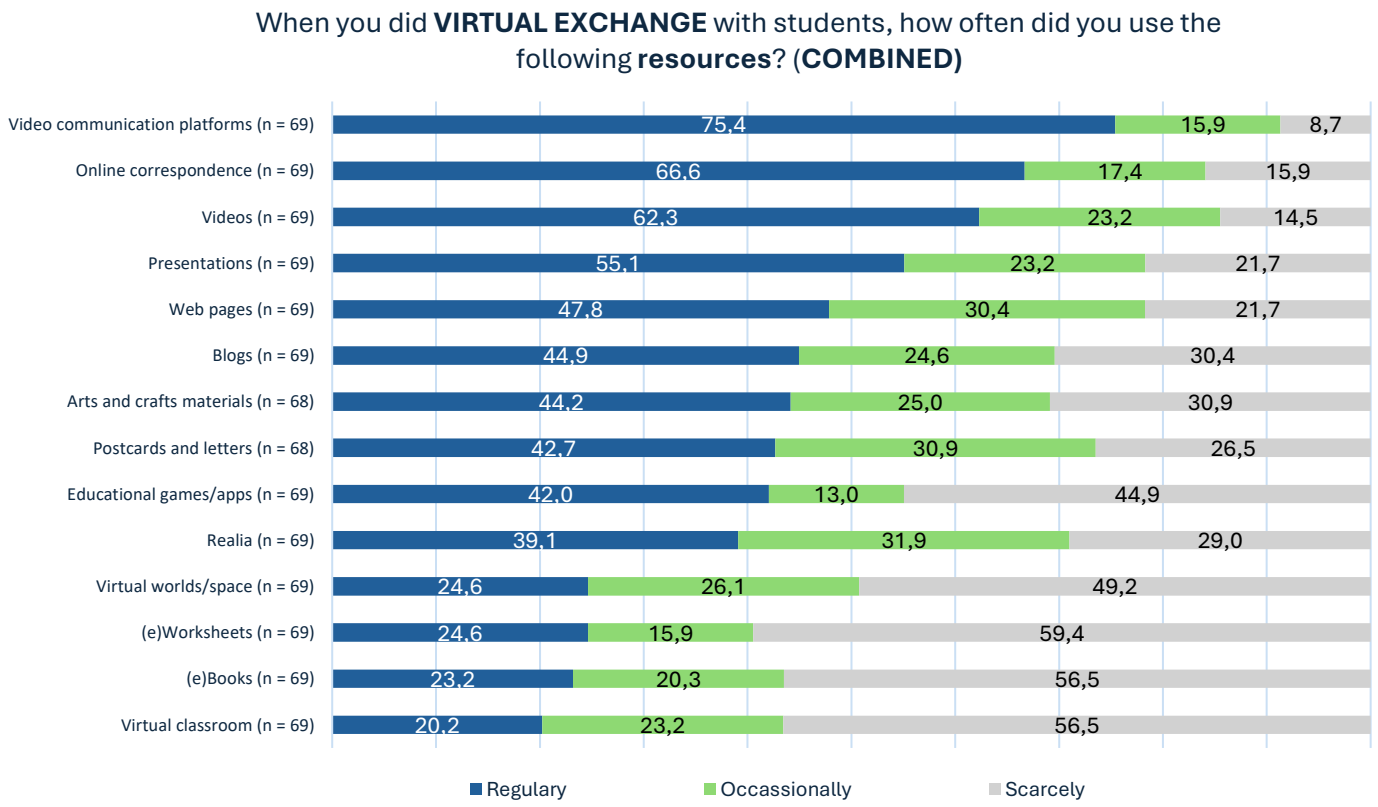

Figure 33: Frequency of using listed resources in VE (combined)

Sample-wise we can legitimately expose *video communication platforms*, such as ZOOM, MS Teams, Skype, Google Meet, etc. as the dominant type or resource in terms of usage as 75.4% of teachers report that they had utilized it regularly with another 15.9% (together 91.3%) declaring themselves as occasional users. *Online correspondence* (explained as texting, email, e-forums, google docs, etc.) and *videos* are comparable with regard to teachers who used them at least occasionally with these proportions being approximately 85%, firmly placed on a bedrock of >60% of regular usage amongst teachers. Apart from these three, every other resource type is not as regularly (i.e., always or frequently) employed, whereas individual differences between varieties ordered consecutively are imperceptible. That is, until we arrive to the lowest, namely *(e)worksheet*, *(e)books* and *virtual classroom*. These were in more than 55% of cases just scarcely used with a third of teachers reporting never to have used them at all.

To see for generalisable differences in scoring, we again conducted Friedman's test, which was significant ($\chi2(F) = 187.907$, df. = 13, $p < .001$). With that we could now say that there are differences in frequency scores between distinct resources, but did not know where exactly these are. For that we needed pairwise comparisons (sample1-sample2 logic), but acknowledging we have 14 items (i.e., resources) that would amount to 91 unique combinations. Out of those, 34 were significant and had to do with either *video communication platforms* being significantly more used, or *(e)worksheets*, *(e)books*, *virtual classroom* (and sometimes *virtual worlds*) being significantly less used than the other item in the pair.

Being adamant that these variables can be surmised under the umbrella term of *resources* without having to verify it with any kind of prior dimension reduction, we directly measured only its reliability which was interpreted as *acceptable* with Cronbach α being .769. We decided to merge these resources into one composite variable, where due to coding lowest mean scores indicated frequent use and high the opposite. And that is what we wanted to see – if a particular group uses more or less than the other.

Variable	**Mann-Whitney Test**				
	Mean Ranks	U	Z	p	effect size estimate
Gender	(n = 7) Male: 35.71 (n = 59) Female: **33.24**	191.00	-.323	**.746**	**r = - 0.04**
Occupation	(n = 8) Pre-school teacher: **33.31** (n = 59) Primary-school teacher: 34.09	230.50	-.107	**.915**	**r = - 0.01**

Figure 34: Differences in combined resources by gender and occupation

Variable	Kruskal-Wallis H				Dunn's post hoc test*	
	Mean ranks	χ^2	df.	p	Sample 1-Sample 2	adj. p
Country	(n = 11) Germany: 34.05 (n = 19) Poland: **26.50** (n = 17) Slovenia: 36.97 (n = 15) Spain: 29.77	3.390	3	**.335**	/	/
Years of service	(n = 3) Novice teachers (1–3 years): 37.17 (n = 6) Proficient teacher (4–6 years): 29.75 (n = 29) Experienced teachers (7–18 years): 35.45 (n = 27) Veteran teachers (19–31 years): 32.65 (n = 1) Master teachers (32–40 years): **11.50**	2.010	4	**.734**	/	/
Number of projects	(n = 46) VE Beginner (1–3 projects): 40.45 (n = 12) VE Intermediate (4–8 projects): 22.00 (n = 4) VE Advanced (9–15 projects): 37.25 (n = 7) VE Master (more than 15 projects): **20.21**	12.318	3	**.006**	VE Intermediate-VE beginner	**.027**

Figure 35: Differences in combined resources by country, years of service and number of projects

Differences in combined resource use were significant only between teachers with varying levels of VE experience (p = .006). Based on *post hoc tests* teacher that had between 4–8 projects, thus those on the intermediate level use resources significantly more than other groups, especially when compared to beginners.

To recapitulate; what seems to be the stimulant (and concurrently a staple) of VE are students' reciprocal presentations of their work and projects, most frequently executed in a form of a whole class activity through video-communication platforms, online correspondence and videos.

However, to achieve this, teachers surely need to first prepare their students, where our respondents usually opt for material presentation and materials-related discussion (69.1%) alongside language structure and vocabulary study (60.3%) as their *modus operandi*. Substantially less (that is 30.9%) devote time to fashioning scripts for dialogues, even fewer (4.4 %) decide to focus on students' interests and wants, whilst only 1.5% of teachers take the spontaneous approach of "no preparation whatsoever". When concretely queried about types of asynchronous activities (i.e. activities not done at the same time with the VE project partners), teachers report resorting to:

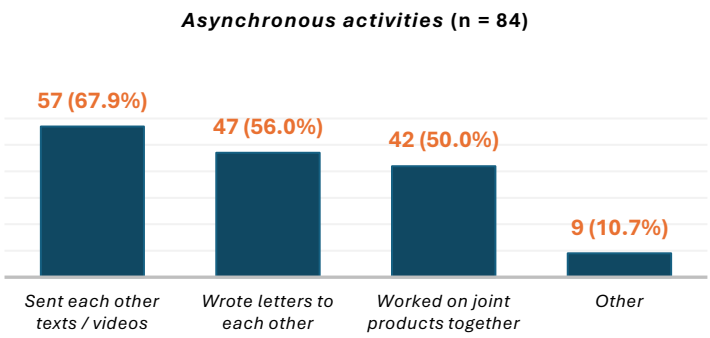

Figure 36: Types of activities used in VE projects

In 'Other' teachers selected sharing documents, listening to podcasts, observing, preparing activities for the other group, sending parcels and similar. In the follow-up open question, they enumerated which activities they considered **most** and **least effective**.

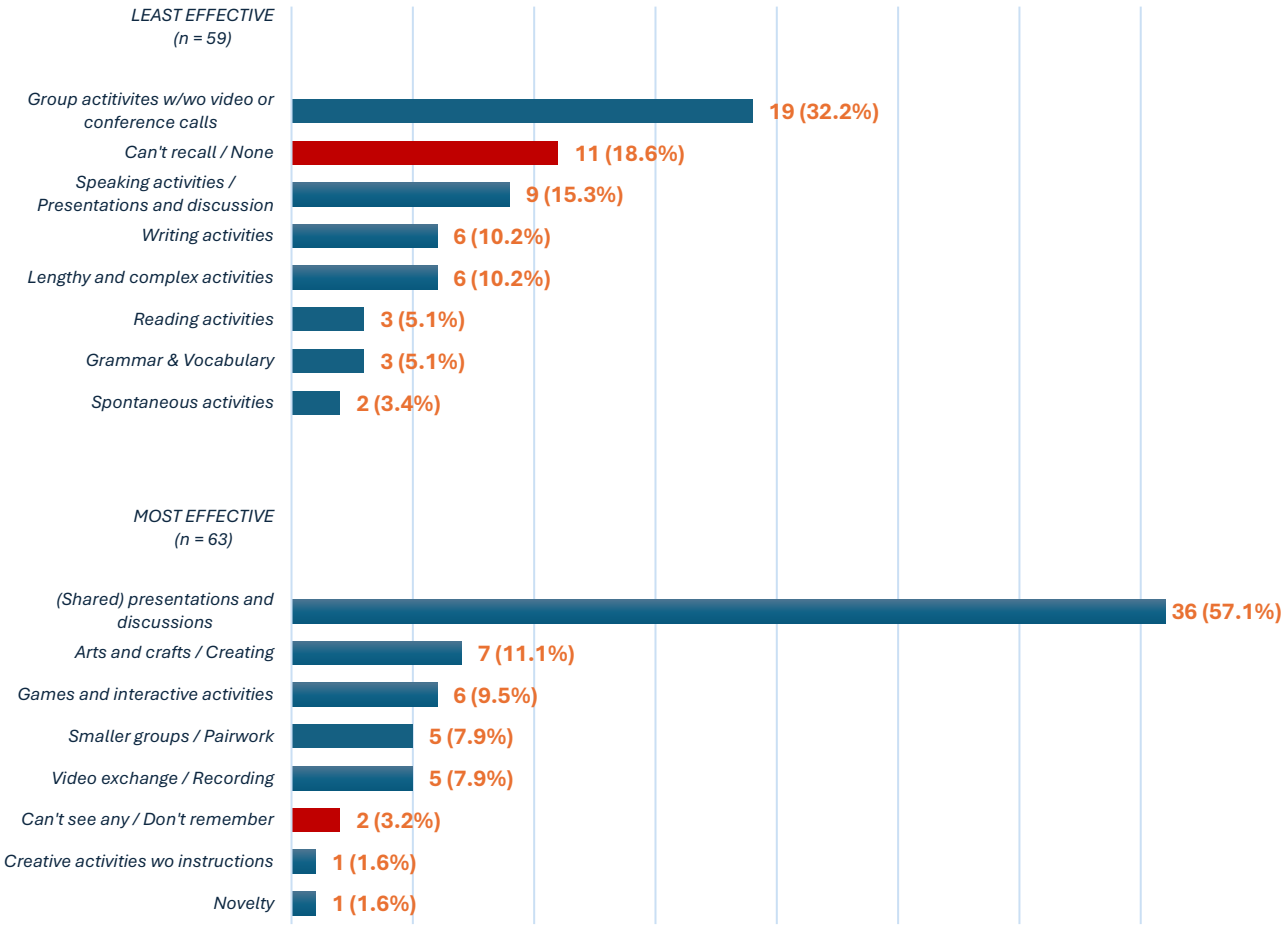

Figure 37: Least and most effective activities used in VE

If we fixate merely on first elements in each category, group activities appear to have yielded the least success, whilst (shared) presentations and discussions were a favourite amongst teachers with 57.1% saying they were most effective activities in virtual exchange projects. However, this exact answer was simultaneously explicitly stated by 15.3% of respondents as ineffective as though thwarting progress, since students often got anxious while waiting for their peers, lost interest, experienced difficulties in following the subject matter. Such justification is not limited to group activities exclusively, but arguably somewhat omnipresent throughout the section of *least effective*.

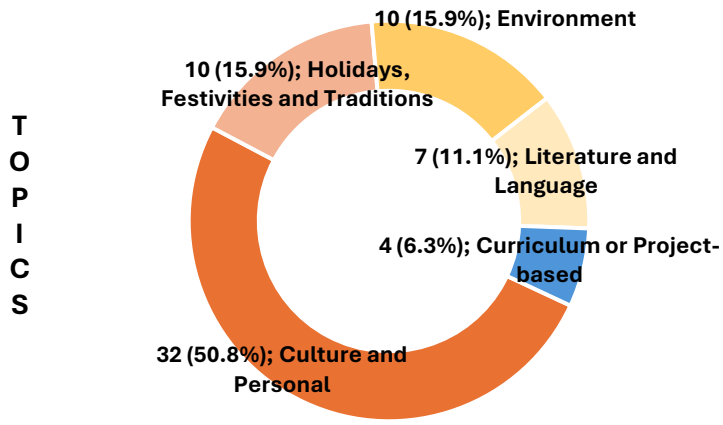

Figure 38: Topics of activities used in VE

According to the sample data, most (50.8%) of these activities were related to themes that are culturally oriented or directly relevant to the students or have some personal meaning for them. In addition, teachers repeatedly expressed that the content focused on the environment (15.9%) or was organised around current holidays, festivities and traditions (15.9%) of both or all participating groups, which could arguably be related to the cultural domain if the answers did not explicitly highlight it as a specific area.

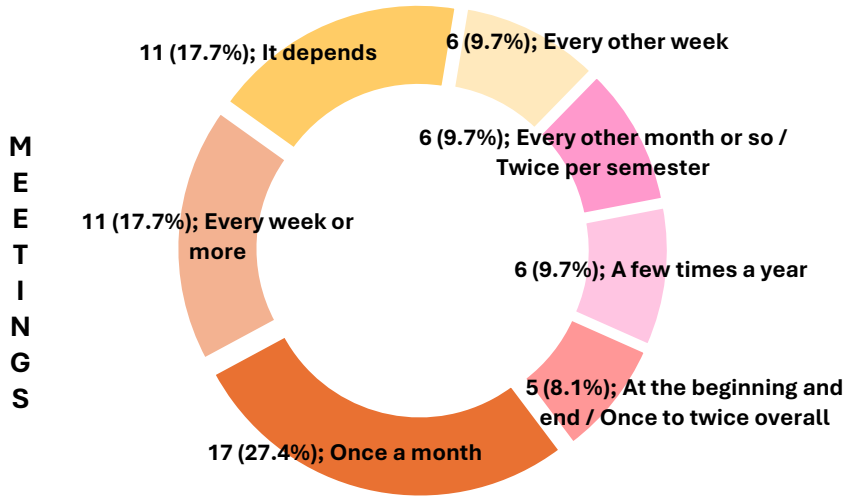

Figure 39: How often should students meet online as reported by teachers

When teachers were asked how often they thought students should meet remotely, the majority (27.4%) said once a month was sufficient. If this is the first ranked answer, the second ranked answers are "every week or more" and "it depends", with 17.7% each, but the fact that it depends on several factors does not really tell us anything and is self-explanatory, obvious and redundant in terms of interpretation, so we can justifiably put "every week or more" in 2nd place. We do not find it worth commenting on the other placements because of the low frequency, but we would just add that an overview of these responses suggests that teachers would prefer to get together more often rather than only intermittently. This also coincides with the problems mentioned by the respondents and discussed above, i.e. that the main obstacle for VE is not related to the number of lessons and consequently the number of meetings, in the sense that these cannot be fitted anywhere in the timetable, since despite the workload teachers still strive to have as many interactions as possible.

Ultimately, we were interested in how all the effort (i.e., the manner, implementation, use of resources, preparation, the activities) results in the students' knowledge or skills. Here, teachers made their assessments on a 5-point scale, ranged from "not a lot" to "a lot".

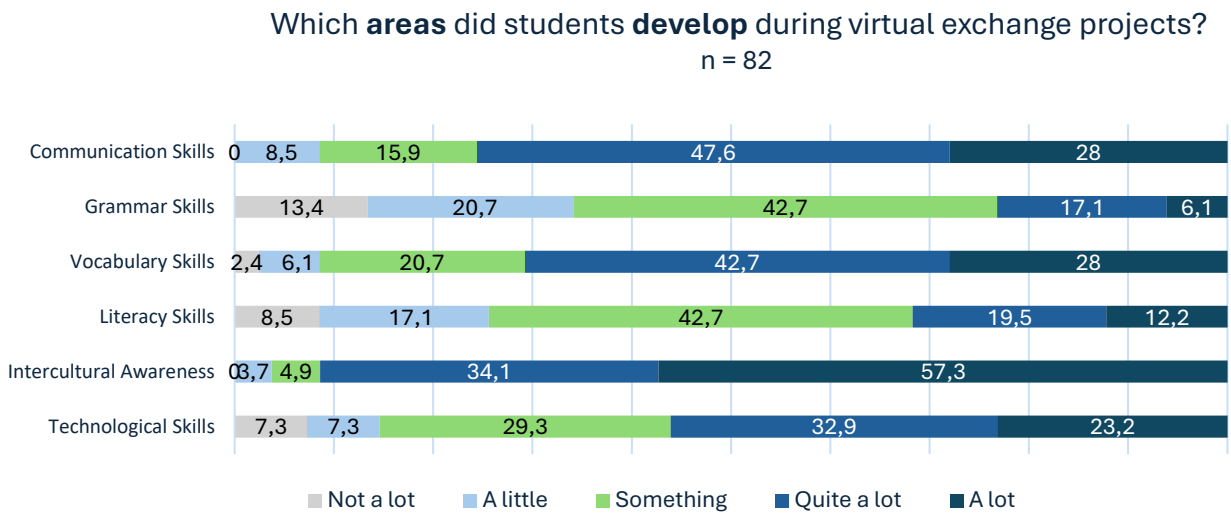

Figure 40: Development of students' skills/areas in VE projects

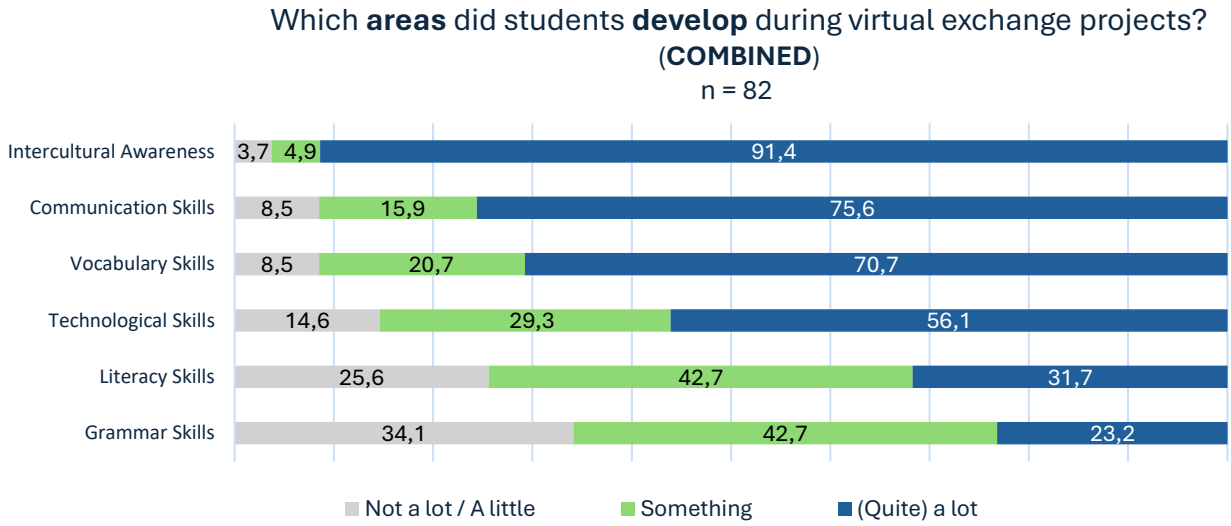

Figure 41: Development of students' skills/areas in VE projects (combined)

Teachers find that as a result of VE projects students have made the most progress in the area of intercultural awareness, with 91.4% of teachers surveyed saying that students have improved *(quite) a lot*. The least noticeable impact of VE is reported to be at the level of grammar.

There are statistically significant differences ($\chi^2(F) = 160.516$, df. = 5, $p < .001$) between teachers' ratings of pupils' skill development. Grammar skills are significantly lower scored in all comparisons, similar is true for literacy skills, whilst intercultural awareness is significantly higher in all cases.

RQ5: How do participants perceive their competence regarding VE projects?

We presented participants with 10 competences related to organising and carrying out virtual exchange projects, where they had to assess themselves on a numerical scale ranged from 1 (*not competent at all*) to 10 (*extremely competent*). For those that had no previous experience, instructions clearly stated that they should estimate their ability in spite of their non-existent prior exposure. Apart from being anchored at the extremes, no other label was provided for the scales.

The inquiries started with *How competent do you feel at/in…:*

1. *Finding partners*
2. *Communicating with partners*
3. *Finding content*
4. *Designing tasks that develop students' linguistic competence*
5. *Designing tasks that develop students' sociocritical competence*
6. *Designing tasks that develop students' digital competence*
7. *Designing tasks that develop students' intercultural competence*
8. *Working with students*
9. *Setting up the classroom for online meetings (e.g., setting up ZOOM sessions)*
10. *Preparing students for online meetings*

Our initial premise was that all these items pertain to the domain of *competencies*; nevertheless, in order to simplify this set, we again decided for *exploratory factor analysis*. Seeing that scales were equally coded, there was no need for any alterations. Our data set has again proved to be suitable for EFA, which we justify with:

1) patterned relationship between variables ($r > \pm .30$ was present in 97.8% of bivariate combinations) and no multicollinearity (i.e. correlations above ± .90 were non-existent) with determinant score of .002 being above the required threshold of .00001;

2) KMO value was .895, often interpreted as meritorious and outpacing the cut-off of .50, whilst a significant Bartlett's Test of Sphericity ($\chi^2 = 1513.57$; $p < .001$) indicated that sample's correlation matrix differs from identity matrix.

Based on principal axis extraction method we adopted a one-factor solution, with which we explain 56.99% of total variance, i.e. dispersion. This was corroborated graphically by scree plot, further reinforcing our decision.

Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loading		
Factor	Total	% of Variance	Comulative %	Total	% of Variance	Comulative %
1	**5.699**	**56.986**	**56.986**	**5.248**	**52.479**	**52.479**
2	.969	9.687	66.673			
3	.825	8.254	74.928			

Figure 42: Exploratory factor analysis regarding competencies

Although we initially opted for varimax rotation, the latter could not be performed due to just one extracted factor. However, all ten variables (i.e., competencies) loaded onto that factor with loadings higher than 0.49, making us name this latent variable as **self-assessed VE competence**, whose reliability score of Cronbach α was .909, deeming an interpretation of being excellent. These results are comparable with that of pre-service teachers.

Competence	**Valid N**	**Mode**	**Mean**	**SD**	**Mean Rank**
Finding partners	***249***	***8***	***5.65***	***2.556***	***3.27***
Communicating with partners	250	8	7.82	2.047	6.59
Finding content	250	8	7.41	1.885	5.68
Linguistic competence	250	8	7.17	1.981	5.32
Sociocritical competence	250	8	6.64	2.046	4.13
Digital competence	250	8	6.48	2.199	4.10
Intercultural competence	250	8	7.25	1.879	5.38
Working with students	**250**	**10**	**8.51**	**1.675**	**7.91**
Setting up the classroom	250	9	7.76	2.181	6.55
Preparing students	248	9	7.55	2.125	6.08

Figure 43: Participants' self-assessed VE competence

Since competencies were rated on a numerical scale, we checked for the normal distribution of data, yet the assumption of normality was violated for every item. Friedman's test was used to determine whether participants self-assess themselves differently across scales, which was proven significant ($\chi^2(F) = 565.409$, df. = 9, $p < .001$). Additional tests revealed that participants perceive themselves significantly less competent in *finding partners* than in every other domain save for in *developing students' digital* and *sociocritical competencies* which were

similarly low scored. These two had lower scores with regard to other items, whilst *working with students* and *setting up the classroom* were ranked significantly higher.
It is no surprise that these tasks are prioritized, as working with students—whether in virtual exchange or not—and preparing for classes are fundamental parts of a teacher's daily responsibilities. Additionally, it's difficult to distinguish between *working with students* in general and *working with students in VE*, which may have led to both being seen as one combined task, creating a unified perception. Conversely, *finding partners* would not be something recurrent and analogous to teaching or child-rearing practices where all stakeholders are customarily either self-imposed (i.e., parents and children) or institutionally appointed (i.e., teachers, principals, counsellors), where there is little to no need to seek additional parties.

Having a firm grasp on how participants rated themselves in general, we wanted to check for differences that would stem from independent segments.

Variable	**Mann-Whitney Test: Competences**				
	Mean Ranks	U	Z	p	effect size estimate
Gender	(n = 22) Male: **143.45** (n = 213) Female: 115.37	1783.00	-1.845	**.065**	**r = - 0.12**
Occupation	(n = 38) Pre-school teacher: **120.09** (n = 199) Primary-school teacher: 118.79	3739.50	-.107	**.915**	**r = - 0.006**
Involvement in VE	(n = 67) Involvement: **163.68** (n = 183) No involvement: 111.52	3572.50	-5.053	**< .001**	**r = - 0.32**

Figure 44: Participants' self-assessed VE competence by gender, occupation and involvement in VE

Gender and occupation have no effect on how teachers rated themselves, however, significant differences were found between those who have been involved in VE and those who have not been, with the former having higher average ratings, ergo, those teachers have better self-assessment scores, whilst the difference is not only statistically significant but also has a moderate to large effect, indicating practical importance with possible real-life implications.

Variable	Kruskal-Wallis H				Dunn's post hoc test*	
	Mean ranks	χ^2	df.	p	Sample 1-Sample 2	adj. p
Country	(n = 34) Germany: 115.49 (n = 75) Poland: 112.83 (n = 81) Slovenia: 101.51 (n = 38) Spain: **144.62**	11.129	3	**.011**	Slovenia-Spain	**.005**
Years of service	(n = 45) Novice teachers (1–3 years): 110.22 (n = 36) Proficient teacher (4–6 years): **130.51** (n = 91) Experienced teachers (7–18 years): 114.47 (n = 60) Veteran teachers (19–31 years): 119.73 (n = 2) Master teachers (32–40 years): 118.00	2.100	4	**.717**	/	/
Number of projects	(n = 44) VE Beginner (1–3 projects): 29.53 (n = 12) VE Intermediate (4–8 projects): 40.29 (n = 4) VE Advanced (9–15 projects): 44.00 (n = 7) VE Master (more than 15 projects): **45.57**	7.094	3	**.069**	**/**	**/**

Figure 45: Participants' self-assessed VE competence by country, years of service and number of projects

Years of service and number of projects were found to be not significant in terms of subsample differences. Yet, there are statistically significant differences between nationalities in how they assessed themselves with regard to VE-related competencies. Slovenian teachers have lowest mean ranks, the contrast being most prominent with Spain, whose teachers were ranked highest. However, in such cases we must not rush to conclusions saying that Slovenian teachers are least competent, Polish and German somewhat mediocre and Spanish being the utmost capable. The reason is rather simple: these competencies were not objectively attained but self-attributed, whereas such an individual's perception is bound to several manifested as well as latent influences, including but not limited to sense of self, appreciation of one's profession in society, mentality, personal philosophy etc. Let's present a fabricated example with which we illustrate our premise: person A might be completely devasted if presented with merely a passing grade, as their culture is orientated towards being "the best", whilst person B might be elated with the same assessment because their logic does not revolve around being the absolute victor, but are completely satisfied with being successful as such. We construe similar reasoning can be applied here.

RQ6: What is the participants' training history regarding VE?

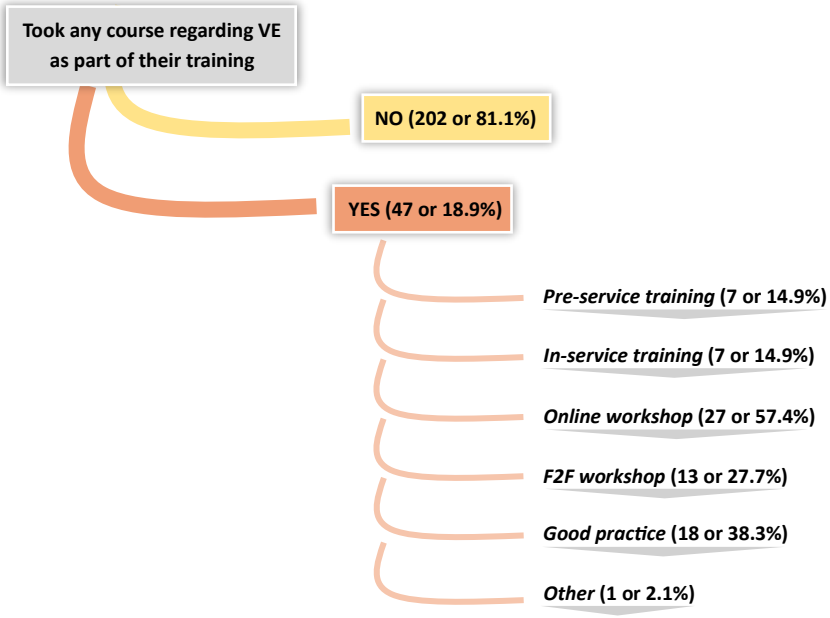

Figure 46: Participants' training history regarding VE

Only 18.9% of teachers had any kind of VE-related training which is why we have decided not to segment these percentages any further. The majority of them received tutelage via online workshops and presentations of good practice.

RQ7: Which areas of VE projects would the participants like to develop professionally?

Next, participants were asked to rank ten areas related to VE about which they would have liked to learn more. Ranking followed the logic of 1 meaning "*the most*", 10 standing for "*the least*"; that is, the lower the mean score, more important is the area in respondents' perception. We hereby provide the table in ascending order mean-wise.

Area	**Valid N**	**Mode**	**Mean**	**SD**	**Mean Rank**
Creating a VE programme	237	1	4.64	3.165	4.61
Finding appropriate tools for VE	234	6	5.29	2.562	5.31
Getting partners	237	2	5.30	3.148	5.29
Designing meaningful tasks for VE	232	9	5.39	3.083	5.42
Improving ICT competences for VE projects	231	3	5.45	2.875	5.49
Finding resources for VE	232	4	5.55	2.413	5.56
Integrating VE projects within the curriculum	230	10	5.56	3.261	5.54
Organising VE online meetings	231	5	5.68	2.613	5.70
Managing online meetings with students	237	8	5.97	2.603	6.05
Finding topics for VE projects	234	7	5.99	2.616	6.03

Figure 47: Participants' interest in VE areas

The hypothesis we set was: Do mean ranks differ between areas that participants would like to learn more about? Or in other words, are certain areas ranked significantly different (i.e., lower, higher) than others.

Friedman Test (χ^2 = 36.778, df. = 9, $p < .001$) left us with the decision to reject the null hypothesis, acknowledging that there are significant differences in participants' desire to develop, with *creating a VE* programme being not only the most sought-after area, but substantially more coveted than *finding resources*, *organising VE meetings*, *finding topics*, *managing online meetings with students* and even than *integrating VE projects into the curriculum*. Other combinations were not tested as significant, but what this tells us is that the very creation of a project takes precedence over everything else, which is logical. Every other area, apart from *getting partners* (which, coincidentally is also relatively high placed) is not exclusive to VE as such or at least not in its entirety. In COVID-19 times teachers had to organise online meetings, manage students' online presence, had to at least somewhat elevate their ICT skills, were these not sufficiently developed before. During their everyday practice they are required to introduce meaningful tasks to students (i.e., relevant, purposeful, intentional, diverse, etc.) and not stray away from the curriculum, ceaselessly interact with children in hopes of transferring any kind of knowledge. Yet, creating a project, that is something peculiar and not ubiquitous, necessitating training and captivating interest.

RQ8: How many participants plan to carry out VE projects in their future?

Apart from either confirming, negating or being ambivalent, teachers had to provide their own justification concerning the introduction of VE into their future lesson plans. As this was open-ended and positioned towards the end of the survey, we anticipated a substantial drop-out, which has indeed happened. Nevertheless, 137 open answers were given which were categorised by several researchers to obtain the least subjective division.

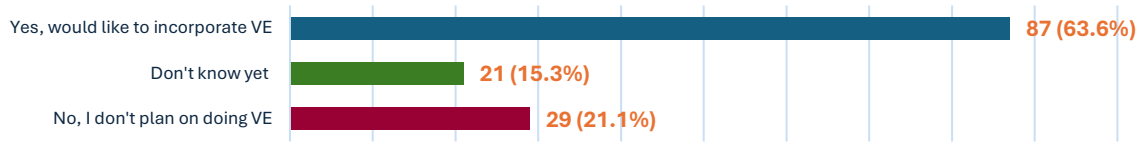

Figure 48: Participants' plans to carry out VE in the future

It clearly shows that most teachers (87 out of valid 137 or 63.6 %) have a positive inclination towards integrating VE into their future pedagogical work. Further clarification is presented below; however, even though the category of "yes" is now split into several more distinct arguments, the general positive response received the majority of counts (38 out of 137 or 27.7%).

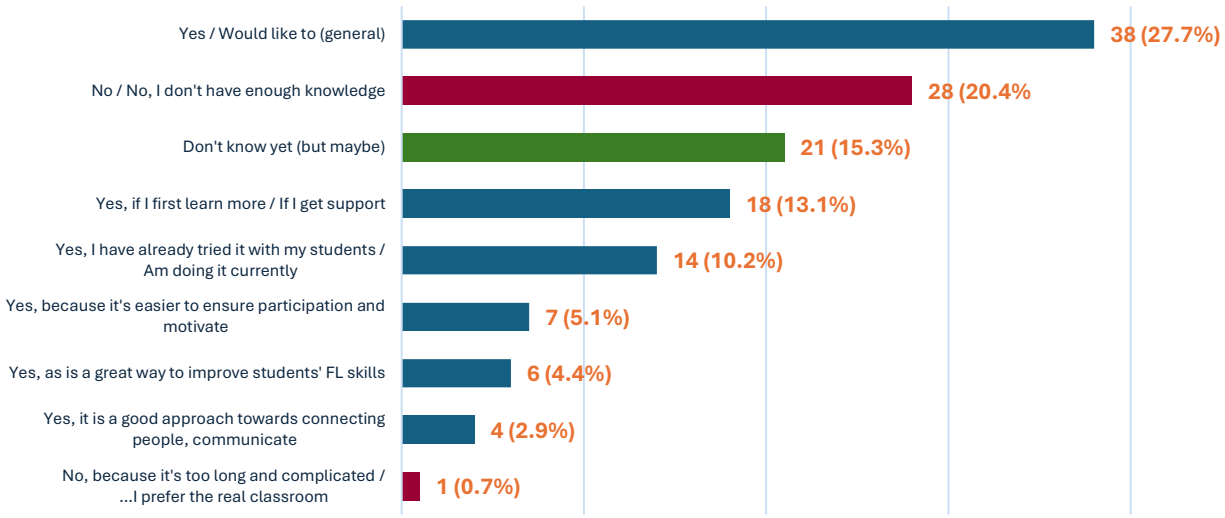

Figure 49: Participants' plans to carry out VE in the future (explained)

To analyse in more detail, inferential testing was conducted to verify which independent variables might affect teachers' willingness to incorporate VE into their lessons. But prior to that, we have recoded the pertinent variable in a way exhibited below.

Variable	*Values (categories)*	*Recoded values that were analysed*
Carrying VE in the Future	1 – No / No, I don't have enough knowledge 2 – No, because it's too long and complicated / Prefer the real classroom 3 – Don't know (but maybe) 4 – Yes, if I first learn more / If I get support 5 – Yes, a great way to improve students' FL skills 6 – Yes, a good approach towards connecting people, communication 7 – Yes, because it's easier to ensure participation / Motivate students 8 – Yes, I have already tried it with my students / Am doing it currently 9 – Yes / Would like to (general)	0 – No, I don't plan on doing VE 1 – I don't know 2 – Yes, I would like to incorporate VE

Figure 50: Participants' willingness to incorporate VE into their lessons

Here we provide a summation of enforced tests and their preliminary interpretation. Differences based on:

gender have proved to be **non-significant** (2î = 2.209, df. = 9, p = .331)

country have proved to be **non-significant** (2î = 18.100, df. = 10, p = .053)

years of service have proved to be **non-significant** (2î = 6.695, df. = 8, p = .570)

occupation have proved to be **non-significant** (2î = 1.562, df. = 2, p = .458)

number of projects have proved to be **non-significant** (2î = 3.721, df. = 6, p = .714)

involvement in VE have proved to be **significant** (χ^2 = 14.472, df. = 2, p < .001).

Since only *involvement* was tested as significant, we checked its effect size with Cramer's V, whose value of .325 (p < .001) signalled moderate to strong association. *Post hoc* testing of residual standardised values confirmed that those who have dabbled in VE were more sympathetic towards incorporating VE in their lessons (in 89.2%) in comparison to those who had no prior experience with VE, since these wanted to feature VE in just 54.0%. The latter had a more expressed percentage of rejections (27.0%) than teachers who have been involved (5.4%).

What followed was checking whether having VE as part of teachers' forthcoming professional activity is dependent on their self-assessed competence. How we have treated *competence* has already been described, so we shall focus only on any additional procedures. Thus, we have taken the combined variable ***self-assessed VE competence*** and recoded it into three groups:

- (Value 1) **Not competent** (6.00 ≥ average score) or [1, 6]
- (Value 2) **Competent** (6,00 < average score < 9.00) or (6, 9)
- (Value 3) **Very competent** (9.00 ≤ average score) or [9, 10]

		Carrying VE in the Future			
		No, I don't plan on doing VE	**I don't know, maybe**	**Yes, I would like to incorporate VE**	Total
Not competent [1, 6]	Count	11	5	16	32
	Expected count	6.8	4.9	20.3	
	%	34.4%	15.6%	50.0%	
Competent (6, 9)	Count	18	15	55	88
	Expected count	18.6	13.5	55.9	
	%	20.5%	17.0%	62.5%	
Very competent [9, 10]	Count	0	1	16	17
	Expected count	3.6	2.6	10.8	
	%	0.0%	5.9%	94.1%	
Total	Count	29	21	87	137
	%	21.2%	15.3%	63.5%	100.0%

Figure 51: Participants' self-assessed VE competence regarding carrying VE in the future

Significant differences were found between groups with differing competence scores with regard to carrying VE in the future (2î = 14.081; df. = 4; p = .007). Both sample data and *post hoc analyses* suggest that those who were categorised as very competent plan to introduce VE into their future lessons to a substantially higher degree (i.e., in 94.1%) than those who were considered competent (62.5%) or not competent (50.0%). Those with lowest average competency are most likely to obviate from VE (in 34.4% as opposed to 20.5% with teachers belonging to the category of *competent* and 0.0% with *very competent*).

Data Presentation per Country

In the following analyses, we have centred around predominantly descriptive group comparisons based on project partners (i.e., Germany, Poland, Slovenia and Spain), which are universally presented in alphabetical order. Causal and inferential statistics have been portrayed in the preceding segments, which is why significant differences might be referred to, yet not calculated anew. Percentages are displayed only if the overall count is high enough.

Added Value of VE

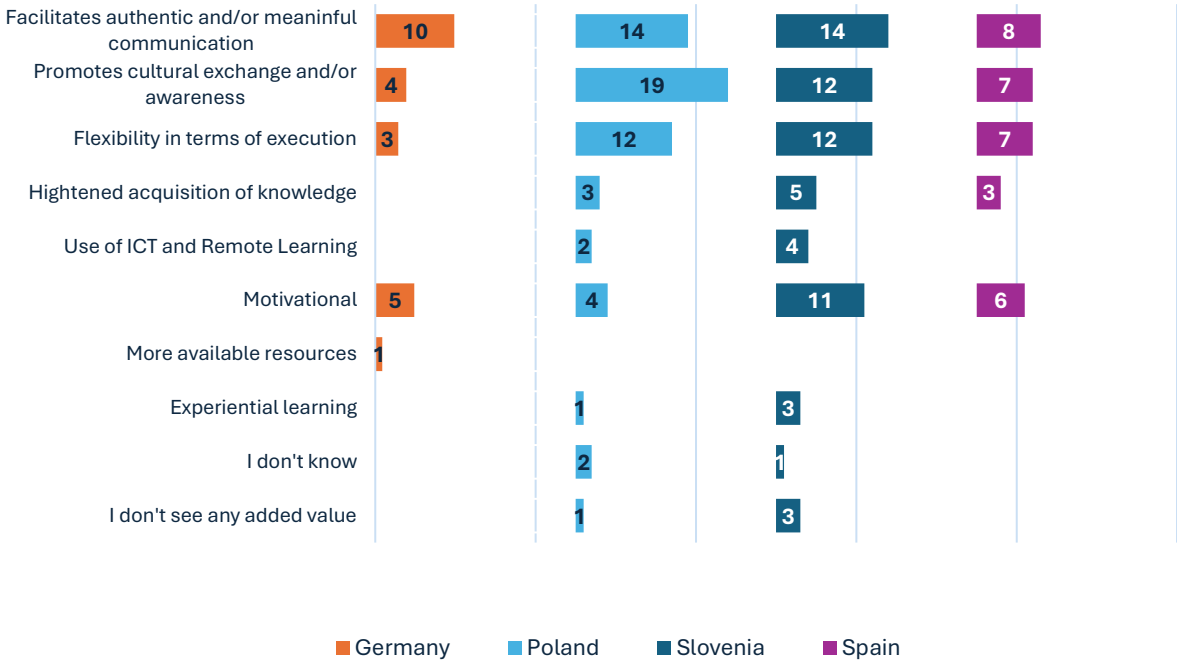

Figure 52: Participants' perceptions of VE's added value (per country)

Germany

German teachers tend to favour *facilitating authentic/meaningful communication* as a factor that establishes a perception of VE's added value, however, this needs to be interpreted with reservation as they were not able to ascertain a high answer count.

Poland

Promoting cultural exchange/awareness seems to be the predominant aspect contributing to the increase in value. *Authentic communication* and *flexibility* should be considered as well.

Slovenia

With minimal differences that can even be construed as marginal, Slovenian teachers place *authentic communication*, *cultural* exchange and *flexibility* as the most important features that elevate virtual exchange worth-

Spain

Although various options received distinct number of answers, these are in practical terms too similar to declare any as the front-runner. The situation is comparable to that of Slovenia with several categories being almost equally addressed.

Most Challenging in Conducting VE

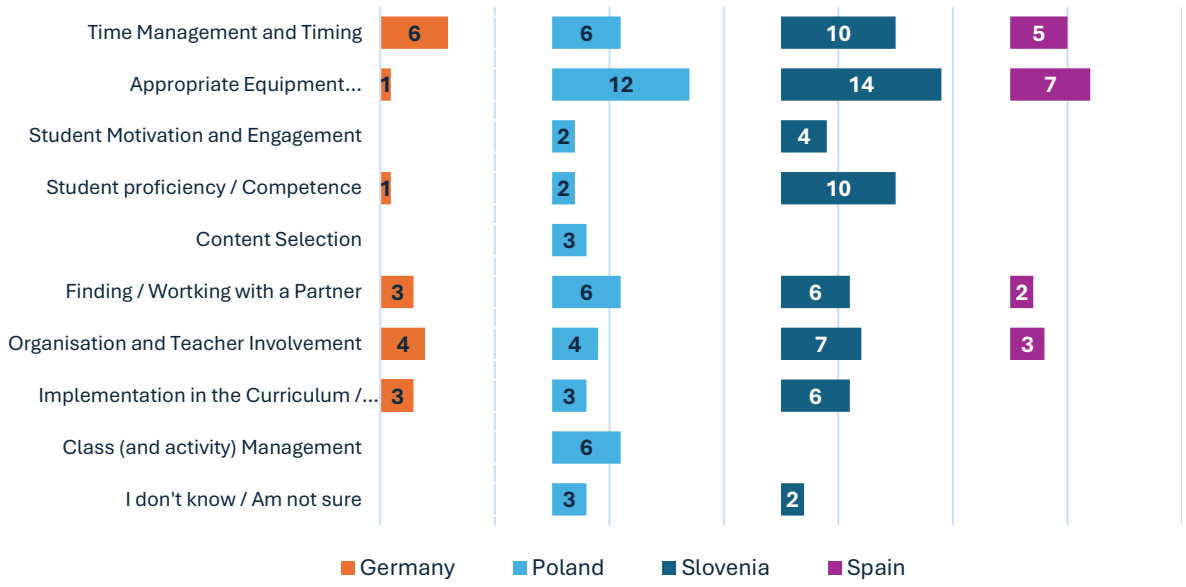

Figure 53: Participants' perceptions of most challenging aspects in conducting VE (per country)

Germany

Due to the scarcity of teachers' responses it would not be prudent to analyse any further than what is illustrated.

Poland

Polish teachers report *equipment, technology and maintaining connection* as most challenging in conducting VE (12 out of 47 or 25.5%), followed by *time management*, *working with partner* and *class management* with 12.8% each.

Slovenia

Appropriate equipment, technology and maintaining connection is regarded by the majority of Slovenian teachers as most challenging (14 out of 59 or 23.7%) with *time management* and *student competence* as tied for a close second (16.9% each).

Spain

Due to the scarcity of teachers' responses it would not be prudent to analyse any further than what is illustrated.

Involvement in VE projects, their averages and descriptions

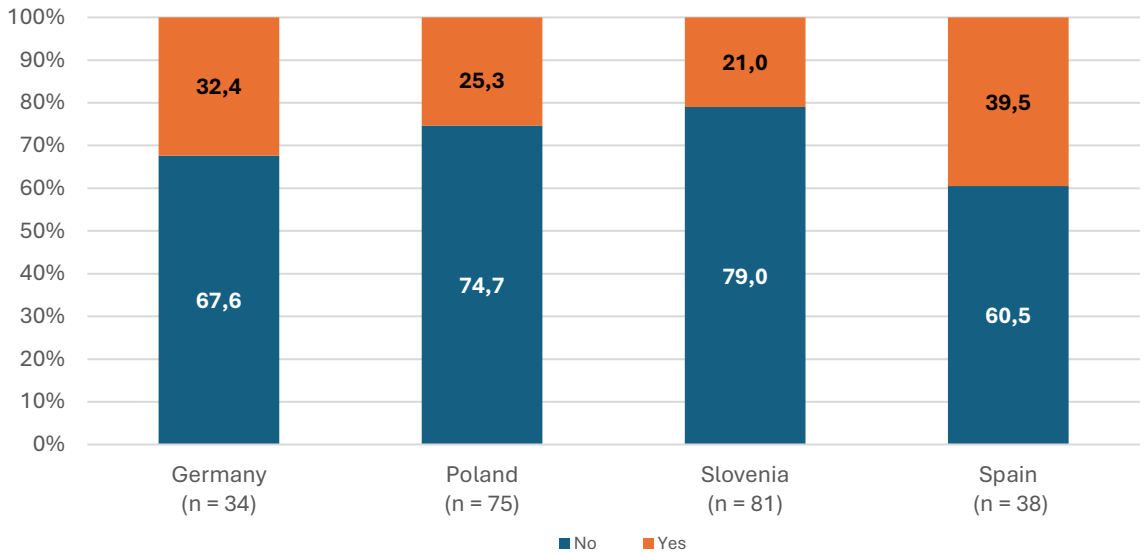

Figure 54: Involvement in VE projects per country

Germany

32.4% (11 out of 34 participants) have previous experience with VE and on average had **3.73 projects** (95% CI between 1.85 and 5.61). Upon describing said projects 3 teachers reported them as eTwinning, another 3 participants as general VE projects, whereas reciprocal presentations, videocalls, ERASMUS+, etc. received one mark each.

Poland

25.3% (19 out of 75 participants) have previous experience with VE and on average had **8.53 projects** (95% CI between 4.22 and 12.83). These were most commonly described as eTwinning (7 out of 19 respondents), VE in general (6 answers) or as projects that focused on content (4 statements).

Slovenia

21.0% (17 out of 81 participants) have previous experience with VE and on average had **3.47 projects** (95% CI between .96 and 5.98). Teachers further on labelled them as eTwinning (7 out

of 16 replies), ERASMUS+ with eTwinning (additional 2 explanations) or as VE in general (another 2). Other details were represented by 1 selection or less.

Spain

39.5% (15 out of 38 participants) have previous experience with VE and on average had **15.87 projects** (CI not reported due to negative lower bound). Focussing on content (6 out of 14) was apparently the most regular strategy with Spanish teachers, with eTwinning and VE in general following with 3 representatives each.

Activities during VE projects

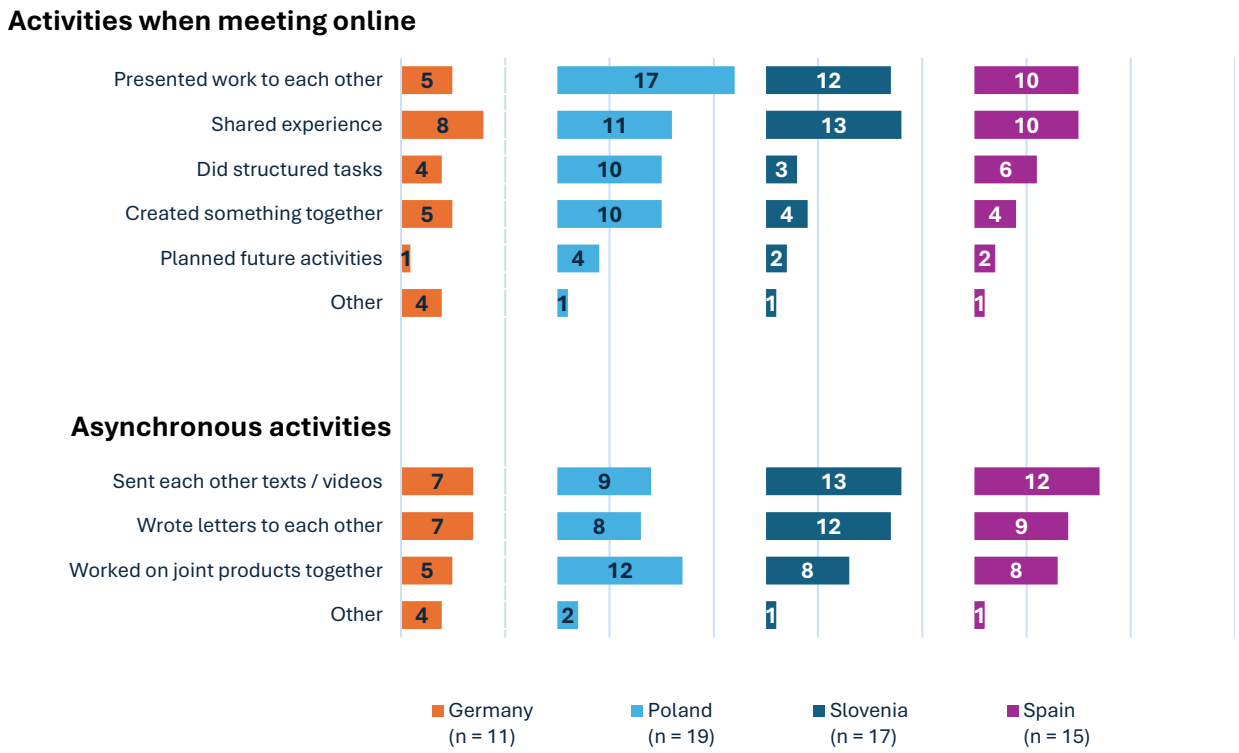

Figure 55: Activities during VE projects per country

Germany

Not only are the activities comparable in terms of the number of responses, but the total number of selections does not allow for any meaningful highlighting.

Poland

Although we are not working with large samples here either, we can still say that the most frequently chosen types of activity are presenting works and reporting or sharing experiences. We have not converted this into percentages because it would simply be nonsensical.

Slovenia

Although we are not working with large samples here either, we can still say that the most frequently chosen types of activity are presenting works and reporting or sharing experiences. We have not converted this into percentages because it would simply be nonsensical.

Spain

Although we are not working with large samples here either, we can still say that the most frequently chosen types of activity are presenting works and reporting or sharing experiences. We have not converted this into percentages because it would simply be nonsensical.

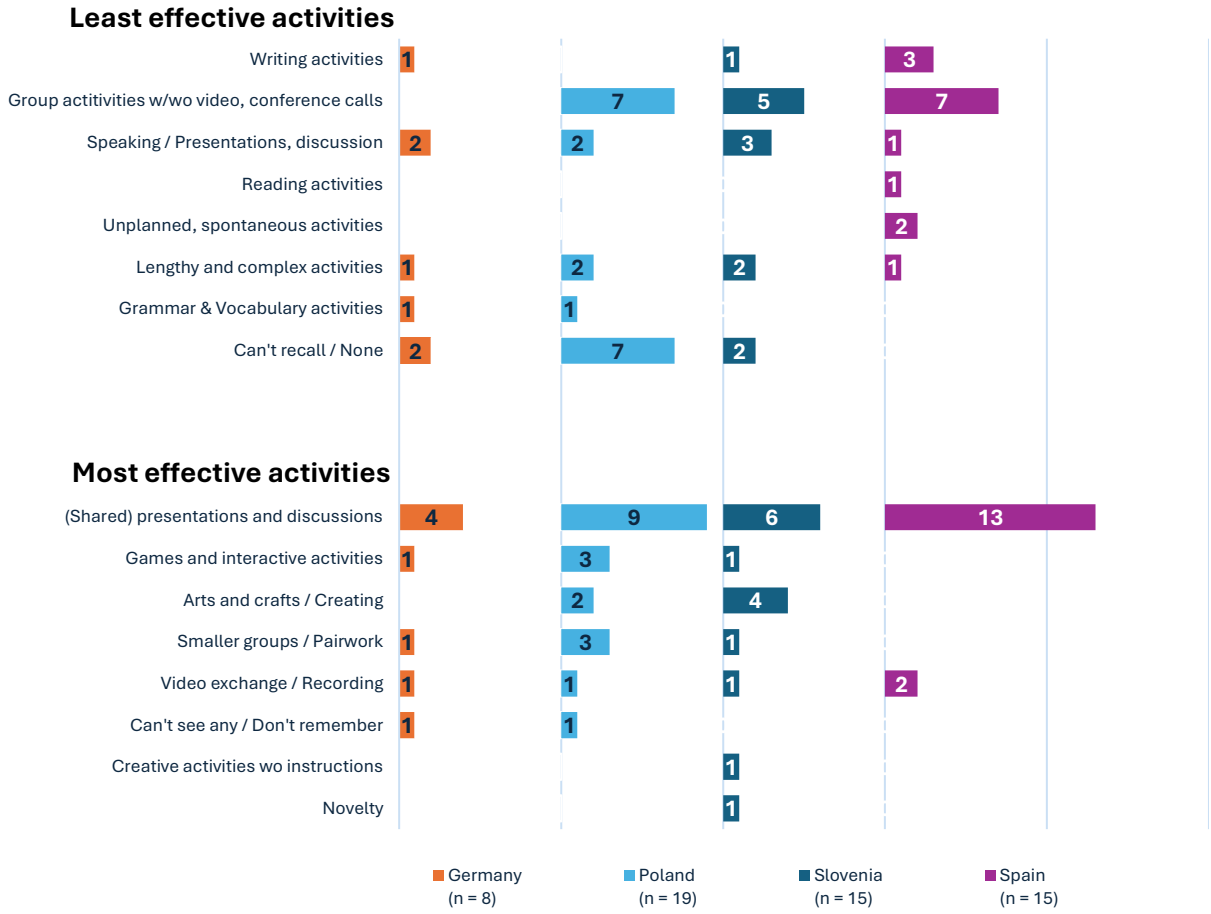

Figure 56: Most and least effective activities during VE projects per country

Given that most options have fewer than 5 representatives each (or none at all), we do not present country by country and refer to the overall presentation given in the previous chapters. The only item that can be singled out is that presentations and discussions prevail in all countries as the most effective activities in VE.

Frequency of Organisation Types

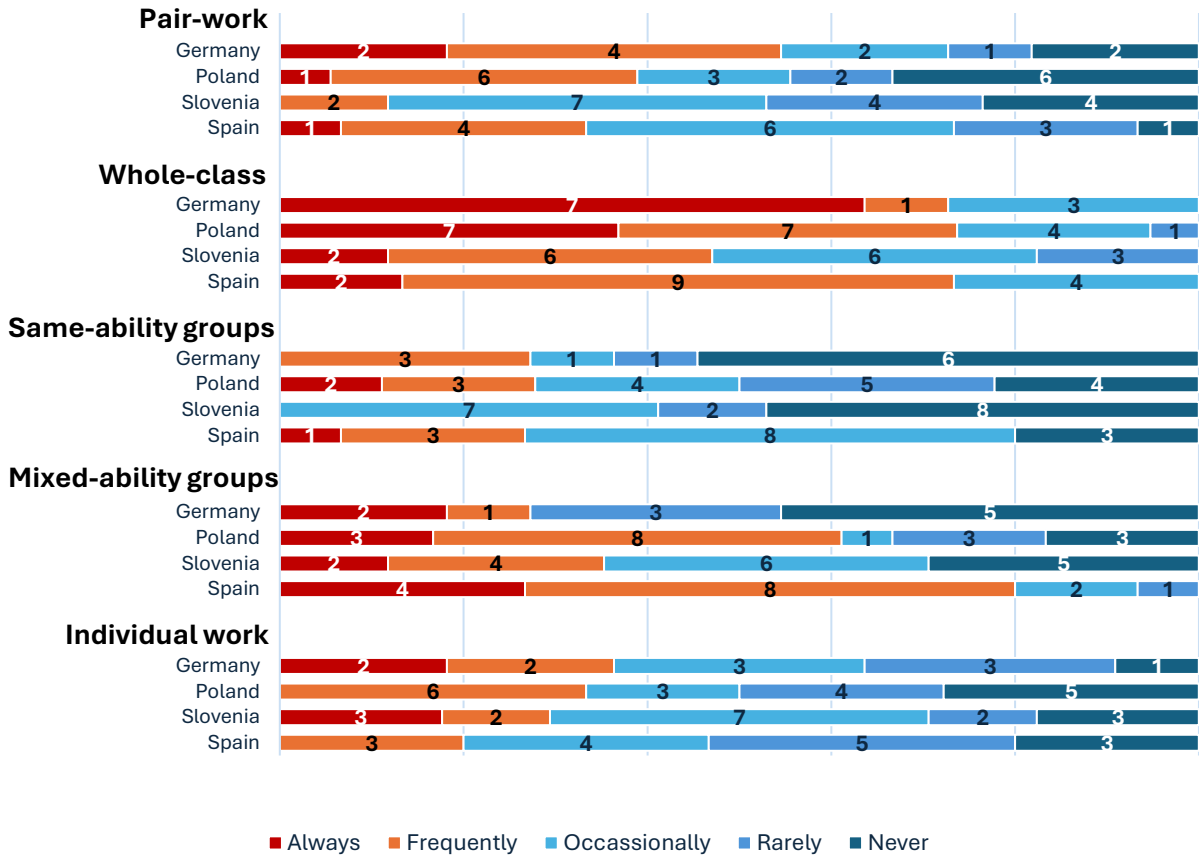

Figure 57: Frequency of organisation types during VE projects per country

Germany

Not enough data for functional quantitative analysis (n = 11) acknowledging dispersion into a 5-point scale.

Poland

Not enough data for functional quantitative analysis (n = 18) acknowledging dispersion into a 5-point scale.

Slovenia

Not enough data for functional quantitative analysis (n = 17) acknowledging dispersion into a 5-point scale.

Spain

Not enough data for functional quantitative analysis (n = 15) acknowledging dispersion into a 5-point scale.

Preparation for VE sessions

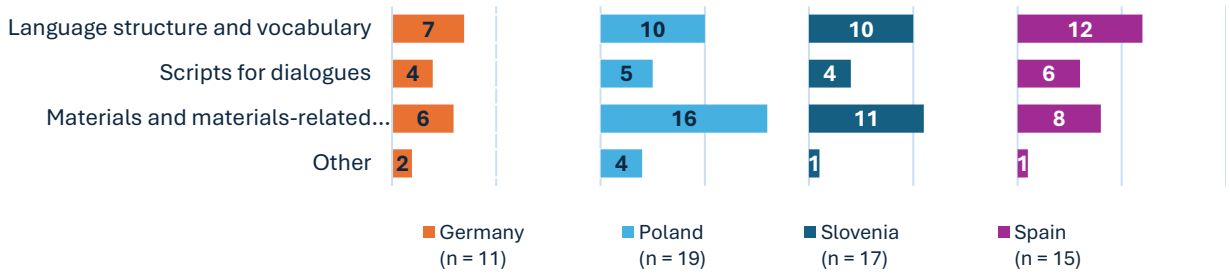

Figure 58: Preparation for VE sessions per country

Training about the development of VE

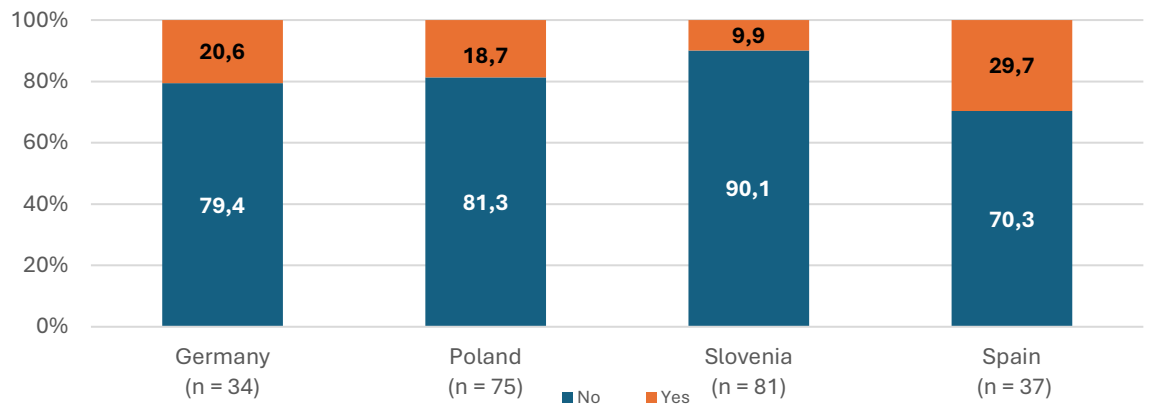

Figure 58: Participants' training about the development of VE per country

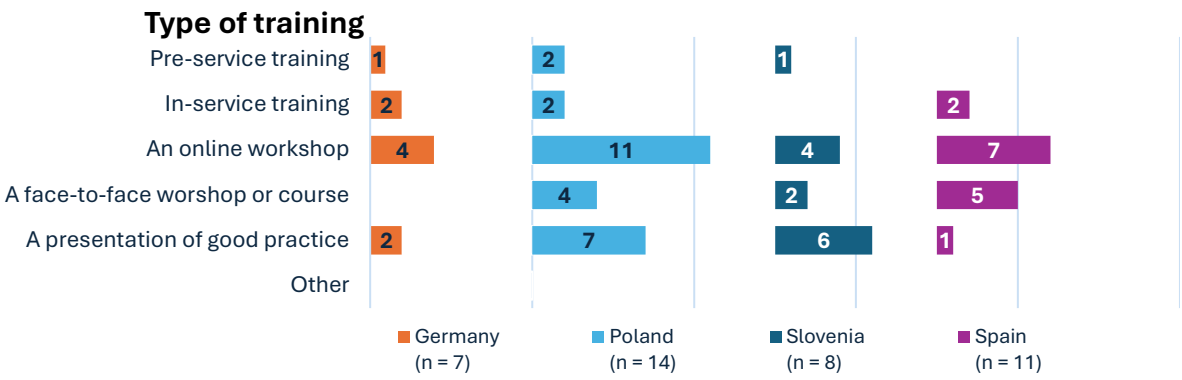

Figure 59: Participants' type of training about the development of VE per country

Germany

20.6% (7 out of 34) received some kind of training, the majority of that being executed through online workshops.

Poland

18.7% (14 out of 75) received some kind of training, the majority of that being executed through online workshops (11 out of 14) and presentations of good practice (7 out of 14).

Slovenia

9.9% (8 out of 81) received some kind of training, the majority of that being executed through presentations of good practice (6 out of 8).

Spain

29.7% (11 out of 37) received some kind of training, the majority of that being executed through online workshops (7 out of 11) and face-to-face courses (5 out of 11).

Carrying out virtual exchange projects with students

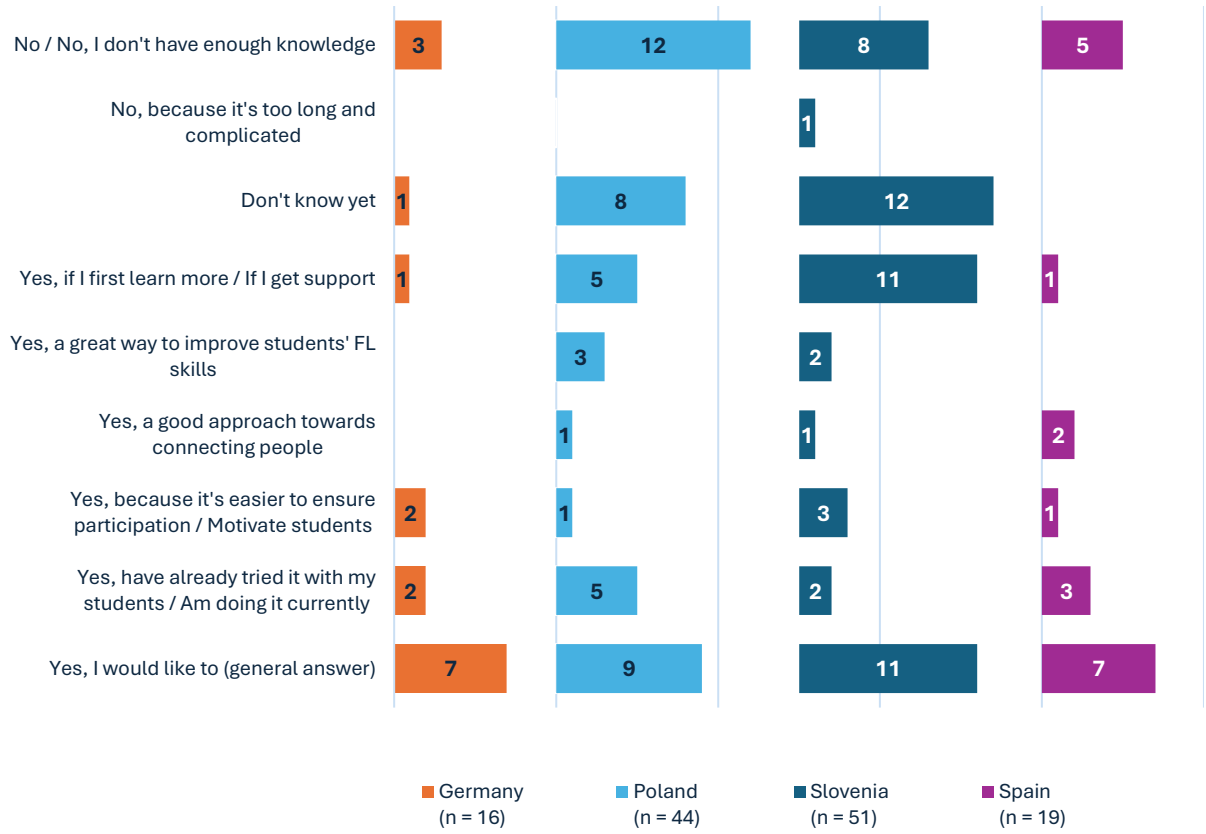

Figure 60: Carrying out VE projects with students per country

Germany

The only response that obtained more than 5 answers was the general *Yes, I would like to* which does not reveal any details or concrete prompts.

Poland

Affirmative answers outweigh the negative when viewed as a whole, however, if analysed individually those who would decline (12 out of 44 or 27.3%) have a bigger proportion than any other decisions yet are closely followed by the general positive response (9 out of 44 or 20.5%).

Slovenia

Top three ranks in the descending order would occupy those who *don't know yet* (23.5%), the general albeit vague *yes* and the conditional confirmation issued by those who would first like to either learn more or get support (21.6% each).

Spain

The only response that obtained more than 5 answers was the general *Yes, I would like to* which does not reveal any details or concrete prompts.

References

Dooly, M., & Vinagre, M. (2021). Research into practice: Virtual exchange in language teaching and learning. *Language Teaching 55*(3), 392–406. https://doi.org/10.1017/S0261444821000069

EVOLVE Project Team. (2020). *The Impact of Virtual Exchange on Student Learning in Higher Education.* EVOLVE Project Report. University of Groningen.

Lenc, A., Žagar Pečjak, M., Šeraj, U. and M. Abramič. (2016). *Study of the Impact of the eTwinning Programme on School Education in Slovenia.* CMEPIUS: Ljubljana.

O'Dowd, Robert. (2023). *Internationalising Higher Education and the Role of Virtual Exchange.* Routledge: Oxon and New York.

Pennock-Speck, B. & Clavel-Arroitia, B. (2022). Virtual exchanges among primary-education pupils. In Potolia, A. & Derivry-Plard, M. (eds.), *Virtual Exchange for Intercultural Language Learning and Teaching: Fostering Communication for the Digital Age*. Routledge. https://doi.org/10.4324/9781003024620

Pettenati, M. C., Nucci, D., Tosi, A., Ranieri, M., Fabbro, F. and E. Imbimbo. (2021). The Impact of eTwinning on the In-service Professional Development of Teachers in Italy. INDIRE: Florence. Nemiña, R. E., Gillanders, C., Leone, V. & Trigo, C. (2023). Expanding learning environments in intitial teacher education. *Pedagogies: An International Journal, 18*(3), 519-533. https://doi.org/10.1080/1554480X.2022.2065995

Renard, O. Y. A. & Milt, K. (2023). Language policy. https://www.europarl.europa.eu/factsheets/en/sheet/142/language-policy