

## Student exchange in primary and secondary education and its effect on language gains, intercultural competence and language learning motivation: a systematic review of research in the European context

Sybille Heinzmann, Catherine Ferris, Thomas Roderer & Kristina Ehram

To cite this article: Sybille Heinzmann, Catherine Ferris, Thomas Roderer & Kristina Ehram (05 Apr 2024): Student exchange in primary and secondary education and its effect on language gains, intercultural competence and language learning motivation: a systematic review of research in the European context, International Journal of Multilingualism, DOI: [10.1080/14790718.2024.2331609](https://doi.org/10.1080/14790718.2024.2331609)

To link to this article: <https://doi.org/10.1080/14790718.2024.2331609>



© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 05 Apr 2024.



[Submit your article to this journal](#)



Article views: 1982



[View related articles](#)



[View Crossmark data](#)



Citing articles: 1 [View citing articles](#)

# Student exchange in primary and secondary education and its effect on language gains, intercultural competence and language learning motivation: a systematic review of research in the European context

Sybille Heinzmann <sup>a,b</sup>, Catherine Ferris <sup>a</sup>, Thomas Roderer<sup>a</sup> and Kristina Ehrsam <sup>a</sup>

<sup>a</sup>Institute for Language Teacher Education, St.Gallen University of Teacher Education, St. Gallen, Switzerland;

<sup>b</sup>School of Education, University of Applied Sciences and Arts Northwestern Switzerland, Muttenz, Switzerland

## ABSTRACT

While internationally there is long-standing research activity on the effectiveness of language contacts (especially in the form of mobility stays) at tertiary level and a number of literature syntheses on this field have been published, research projects at the primary and secondary level are still comparatively rare and the results are often not published in peer-reviewed journals. This literature review compares and systematises these works, some of which have received little or no international attention. By analysing and systematizing the findings, the literature review primarily aims to identify what is known about the effectiveness of student exchanges at these lower levels of education and where there are research gaps and a need for future research.

## ARTICLE HISTORY

Received 25 January 2023

Accepted 12 March 2024



## KEYWORDS

Student exchange; study abroad; language gains; intercultural competence; language learning motivation; systematic review

## Introduction

This article offers a critical review of research on the effectiveness of student exchange during primary and secondary school, with a focus on the European context.<sup>1</sup> Unlike other existing reviews in the language field, it focuses on several target variables, more specifically linguistic gains, intercultural competence and language learning motivation.

Complementary to regular foreign language instruction, student exchange programmes (both long and short) offer an opportunity for immersion in the target language (TL) and culture. At the tertiary level, mobility programmes are often an integral part of the curriculum. This may be related to the fact that students at this educational stage are more willing to leave home for a (possibly longer) period of time, stay with a host family or at a student residence, and attend courses at a host university. At primary

**CONTACT** Sybille Heinzmann  sybille.heinzmann@fhnw.ch  Institute for Language Teacher Education, St.Gallen University of Teacher Education, Notkerstrasse 27, 9000, St. Gallen, Switzerland

© 2024 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

and secondary level, they are less common and less institutionalised but still a regular phenomenon, at least in Europe.

Meanwhile, there are a number of state-of-the-art articles, literature reviews and meta-analyses dealing with student mobility (Borràs & Llanes, 2021; Hirai, 2018; Isabelli-García et al., 2018; Llanes, 2011; Roy et al., 2019; Tseng et al., 2024; Tullock & Ortega, 2017; Xiao, 2015; Xu, 2019; Yang, 2016). What is common to all these reviews is that most of the treated literature focuses on tertiary level education, thus involving university students. Younger learners are only marginally considered. To date, only the review by Borràs and Llanes (2021) and the meta-analyses of Tseng et al. (2024) and Xu (2019) include studies involving participants at primary or secondary school level and these studies clearly represent a small portion of the covered literature in these reviews. If we are, however, interested in the effects of study abroad (SA) more holistically and long-term or for different age groups, we must also investigate SA programmes at lower educational levels.

Further, Llanes (2011) mentions the lack of research on the effectiveness of SA from a European perspective, a niche in which her own research is located (cf. Llanes, 2012; Muñoz & Llanes, 2014). Since then, research efforts in this area have increased and the SAREP (e.g. Howard, 2019; 2020; Mitchell & Tyne, 2021), LANGSNAP (e.g. Mitchell et al., 2017), SALA (e.g. Pérez-Vidal, 2014) and PEERLANG (e.g. Paradowski et al., 2021; Paradowski et al., 2022) projects provide us with important insights into adult SA in Europe. What is still lacking, though, is a synthesis of research on younger SA participants in the European context.

The current review on European SA specifically targets the under-researched population of primary and secondary students. Unlike previous reviews, it includes literature in languages other than English (German and French in addition to English) as well as grey/fugitive literature in order to prevent a publication bias.

## **Participants**

The disentanglement of students in primary and secondary education from older learners (university students/adults) can provide new insights into the potential of student exchange for developing language competences of young learners with low proficiency levels. Admittedly, learners at primary and secondary school level differ from adult students in their cognitive and emotional maturity (Kray & Schaefer, 2018; Mahatmya et al., 2012; Weichold & Silbereisen, 2018; Zimmer-Gembeck & Collins, 2006). We may not only assume higher self-awareness and language learning awareness amongst secondary school students, but also more independent engagement with their language development. Further, they may be expected to consider mobility more frequently and for longer periods of time. Learners at primary level, on the other hand, are neither expected to engage in long exchanges involving (far) travel, nor to reflect on their language development to the same extent. They are expected to require more coaching and scaffolding from educators. While we acknowledge these differences in target groups, an inclusion of both in the same review seems justified in light of the fact that the meta-study by Tseng et al. (2024) found that age is not a significant variable in predicting learners' linguistic gains in SA programmes. There is no a priori reason to think that different age groups may not also equally profit interculturally or motivationally provided programmes are well-tailored to their age-specific needs. Thus, including both

primary and secondary school levels should tell us more about SA at earlier stages of education more generally.

### ***Notion of student exchange***

The review covers studies on a broad range of exchange programmes that bring students from different linguistic and cultural backgrounds together, ranging from typical two-way class exchanges or individual exchanges, to school camps, culture trips, youth encounters, but also digital formats of exchange. We will use the umbrella term student exchange to refer to this broad spectrum of programmes. Exchange in this sense includes all L2-related contact activities where students are brought into contact with people from another linguistic and cultural background with the aim to foster personal, cultural and/or linguistic development irrespective of whether they actually travel and whether there are reciprocal visits (where one student comes and another one goes). In the article we will also use the terms SA, student mobility, and other derivations depending on what term best describes the programme referred to.

### ***Definition of constructs under investigation***

As pointed out above, unlike most existing reviews, this literature synthesis focuses on several target variables. In the following, we will briefly explain how we define the constructs under investigation and elaborate on our strategy for the inclusion/exclusion of studies. In general, we were inclusive in terms of the theoretical background and elaboration of the included studies. We decided to include studies which themselves claim to investigate language competence, language learning motivation or intercultural competence irrespective of whether their conceptualisation matches our own understanding of these constructs in order to be maximally comprehensive in terms of the available empirical basis and to include grey literature which tends to be under-theorised. The aim of this strategy being in particular to not exclude potentially interesting studies that may lack detailed theoretical elaboration.

Intercultural competence (IC) is a complex and multi-layered construct. A comprehensive definition that does justice to this complexity and will be used as a working definition in this paper is the one by Deardorff (2006, p. 247) which defines intercultural competence as ‘the ability to communicate effectively and appropriately in intercultural situations based on one’s intercultural knowledge, skills and attitudes’. As Deardorff points out, this definition is derived from Byram’s (1997) work on intercultural communicative competence, a model that is used in a number of studies included in the review. However, the review also includes studies based on developmental models of IC (e.g. Bennett’s (1993) developmental model of intercultural sensitivity) or studies focusing on particular personality traits that have been claimed to be part of IC, such as respect, empathy, flexibility, curiosity, or tolerance of ambiguity (Ruben, 1976), as well as literature that does not outline what model or concept of IC they are based on.

While there is no universally accepted definition of motivation, most motivation researchers agree that motivation has to do with the direction and magnitude of human behaviour. A variety of factors, such as interest, expectations, self-efficacy beliefs or perseverance can influence this directionality and quality. Different motivation theories and models focus on different factors, some putting more emphasis on cognitive aspects,

others on affective factors and yet others on behavioural factors. Once again, the current review is comprehensive. It includes studies based on social psychological perspectives on motivation, such as Gardner's (1985) socio-educational model of motivation, educational psychological perspectives, such as self-determination theory (Ryan & Deci, 2000) or integrative perspectives, such as Dörnyei's (2009) L2 motivational self-system. However, just as in the case of IC, the review also includes studies that do not elaborate on the motivational models underlying the study. Hence, the review covers all empirical studies that deal with the effect of an L2 contact activity on students' language learning motivation with or without reference to specific theoretical models. What is not included in the review are studies that do not focus on the motivation to learn the TL but on students' motivation to participate in an exchange or achieve motivation more generally.

As far as language competence is concerned, we follow the conceptualisation of the Common European Framework of Reference for Languages (CEFR) (Council of Europe, 2020). The CEFR views language learners as language users and social agents. Language use, including language learning, comprises the actions performed by persons who as individuals and as social agents develop a range of both general (e.g. knowledge of the world, socio-cultural competence, IC) and communicative language competences (linguistic, sociolinguistic, pragmatic and plurilingual competences) by drawing on the competences acquired in various contexts under various conditions and activating those strategies which seem most appropriate for carrying out the tasks to be accomplished. The language activities that users engage in involve reception, production, interaction and mediation. In line with this conceptualisation the current review includes studies dealing with a broad spectrum of linguistic skills / competences at different granular levels, such as speaking skills (including pronunciation, fluency, accuracy, complexity), writing skills (including fluency, accuracy, complexity), interaction skills, listening skills, and pragmatic competence.

The questions we aim to answer in this review are:

- (1) What empirical research has been undertaken on the impact of student exchange on European primary and secondary school students' language gains, language learning motivation and intercultural competence?
- (2) How effective is student exchange in primary and secondary schools with respect to the development of
  - a. language competence
  - b. language learning motivation
  - c. intercultural competence

## **Method**

### ***Study selection criteria***

As pointed out above, the present review focuses on student exchange in primary and secondary school education in Europe. Studies dealing with tertiary level students have explicitly been excluded as the state of research for this level of education is well documented in previous literature reviews (e.g. Isabelli-García et al., 2018; Roy et al., 2019). To further limit the scope of the review, studies published before 1990 have been excluded. This cut-off has been chosen as (a) there has been little systematic research

on student exchange before 1990 (b) existing literature before 1990 mainly consists of field reports (see Ferguson in Freed, 1995) and (c) the 1990s marked the beginning of European educational cooperation initiatives fostering transnational collaboration and exchange (e.g. Socrates, Leonardo, Comenius). To limit publication bias, the current review includes fugitive literature (e.g. research reports) as suggested by Norris and Ortega (2000). This 'fugitive' literature is expected to be important, as many evaluations of exchange projects culminated in unpublished research reports written for third-party funders rather than being published in scientific journals. Furthermore, there is a tendency not to publish project findings that produce non-significant or counterintuitive results, the 'file-drawer' problem (Rosenthal, 1979 in Norris & Ortega, 2000). In addition, the present review includes literature in languages other than English, specifically German and French. The decision to review studies written in German and French, in addition to English (the preferred language for scientific publication), was based on the researchers' mastery thereof and the status of German and French as target languages in the field. English, German, and French are not only the chosen languages of publication, but they also reflect the most represented target languages amongst the identified literature. We see this as a first step in including languages other than English and encourage other researchers to expand the scope and review literature in further languages that we were unable to include. The review only includes empirical studies with primary data; accordingly, meta-studies and review articles are not considered.

The inclusion criteria for the selection of publications in this review were the following:

### ***Inclusion criteria***

#### ***School level.***

- Compulsory schooling (primary and lower secondary)
- Upper secondary level (vocational and general education)

#### ***Geographical.***

- Europe (incl. participants from Europe studying abroad in non-European countries)<sup>2</sup>

#### ***Timeframe.***

- Publications from 1990 onwards

#### ***Type of literature.***

- Journal articles (with and without peer review)
- Monographs
- Book chapters
- Grey/fugitive literature (e.g. research reports, master theses, unpublished dissertations)

#### ***Target variables.***

- Language competence
- Intercultural competence
- Language learning motivation

Studies were excluded if they met one or more of the following exclusion criteria:

### ***Exclusion criteria***

#### ***School level/ school.***

- Tertiary level (college and university)
- Studies focusing solely on teachers

#### ***Geographical.***

- Studies with non-European participants (incl. those with Europe as target country)

#### ***Type of literature.***

- Studies from before 1990
- Meta-studies, review articles

#### ***Target variables.***

- Studies in which language learning is not an issue (youth encounters without language learning, e.g. studies on international youth (social) work)
- Studies on the effects of SA on personality traits (e.g. self-confidence, self-efficacy)

#### ***Languages.***

- Publications in languages other than English, German and French

### ***Search process***

In order to accumulate insights based on all relevant research that has been undertaken, the review includes qualitative, quantitative as well as mixed-method studies. Further, as our topic of interest is likely not to be confined to a specific academic discipline or a single database, we followed the recommendation to look for studies in various sources (Gough et al., 2017). The following five databases were searched: Bibliographie Linguistischer Literatur Datenbank (BLLDB), Education Resources Information Center (ERIC), JSTOR, Modern Language Association Bibliography (MLA) and Web of Science. In addition, we searched for literature on Google Scholar and conducted a general web search in order not to miss informal or grey literature. For both these latter searches, sighting of the results was limited to the first four pages of hits (approx. 40-60 hits per search term) as the relevance of the results strongly decreased (e.g. advertisements instead of articles) and there was a strong overlap of results due to the thematically related search prompts.

Relevant studies were identified by searching for the following predefined keywords:

German: Sprachaustausch, Mobilität, Sprachbegegnungen, Jugendbegegnungen, Jugendaustausch, Schüleraustausch, Sprachkontakte

French: rencontres linguistiques, rencontres internationales, rencontres interculturelles, échange linguistique, échange interculturel, mobilité

English: study abroad, mobility, linguistic exchange, language exchange, exchange activities, student exchange, intercultural encounters

The search process was refined by adding boolean operators if a search term yielded more than 300 hits (e.g. mobility AND language learning), by using combinations of terms like 'study abroad' or 'foreign language' in inverted commas or linking terms with the logical connection 'and' by searching both with and without truncation, (e.g. foreign\* language\*).

During the search process a first screening of the identified items was carried out on the basis of the title and abstract and only studies which fit the above-mentioned inclusion criteria were taken up in the preliminary corpus. Thus, in total 4996 hits (3699 in English, 742 in French, 555 in German) were reduced to 159 relevant research papers (113 in English, 18 in French, 28 in German) which met our inclusion criteria. Of these, 20 were duplicates, so that 139 publications were entered in an Endnote-database for further inspection and classification.

All 139 studies were retrieved from accessible libraries on- and offline, via web searches, orders from libraries or direct contact with the authors to be read in their entirety. The detailed reading of the texts revealed that another 79 texts did not meet the inclusion criteria. Most often, the sample criteria were not met because participants in the programmes were either too old or from outside Europe. In addition, identified publications were often not empirical studies with their own data, but e.g. project reports without data or meta-analyses and reviews. In some cases, it only became clear during reading the text that the research focus was on a type of exchange not considered in this review since it is not related to language learning in the school context (e.g. youth work). This process resulted in 60 publications which were subjected to an in-depth content analysis using the MAXQDA software (Figure 1).

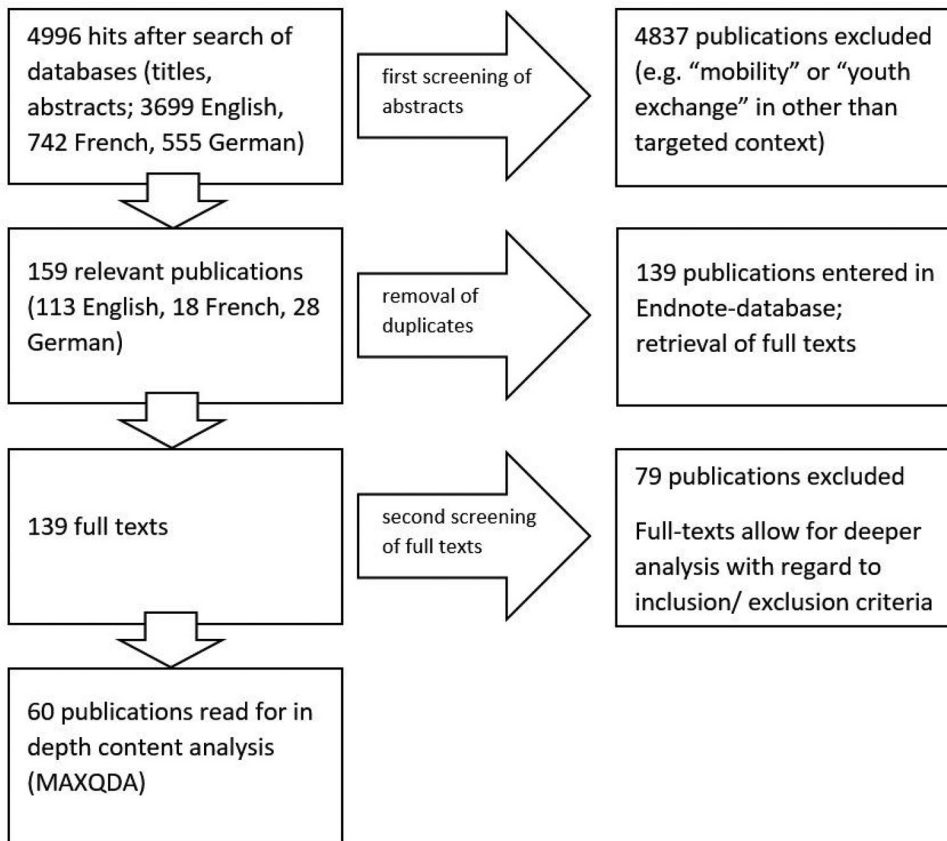
### ***Coding characteristics***

After the identification of relevant publications, these were coded and classified with reference to the following four dimensions using MAXQDA.

- (1) characteristics of the studies (e.g. publication type, language of publication, peer-review status)
- (2) characteristics of participants and programmes (age, first language(s), target language, educational level, exchange format, duration, etc.)
- (3) the methodology of the research in order to assess the quality and relevance of the studies (e.g. quantitative, qualitative, mixed-method, target variables, sample size, sampling technique, randomisation)
- (4) the findings in order to integrate and synthesise the results.

In a first step, four project collaborators coded two publications each to assess the feasibility and comprehensiveness of the coding scheme. Based on this trial stage, some codes were added, and some codes were refined. The full coding scheme can be





**Figure 1.** Flow diagram of review process.

found in [Table A1](#) in the appendix. While some variables (e.g. publication type, educational level) were entered into a variable matrix in MAXQDA (using values as in statistical programmes), most variables consisted of free text coding (e.g. summaries of the findings, procedure, etc.).

## Results

### ***Mapping of empirical research on the impact of student exchange on European primary and secondary school students' language gains, language learning motivation and intercultural competence? (RQ 1)***

The sections below will outline relevant characteristics of the included studies with reference to the participants and types of programmes involved.

#### ***Participants***

The reported age of the participants was very broad. In 35 studies the reported age of the participants ranged from 9 to 20 years. In 12 studies different age groups including adults

(max. age 54) were compared. In an additional 13 studies, information concerning the participants' age was missing. As studies in tertiary education (college and university) were excluded, academic status ranged from primary (21 studies) to lower secondary (23 studies) to upper secondary in either general (21) and/or vocational (7) education.

A variety of participants' first languages were mentioned across the studies. German was the most frequently mentioned first language (36), followed by French (24), Spanish (15), and Catalan (10). English, on the other hand, was much more often the target language than the students' first language and was only mentioned as an L1 eight times. Both Italian and Dutch were mentioned three times as participants' first language.

With regard to the target language examined in the studies, English was mentioned most frequently (31), followed by German (27) and French (26). Italian and Spanish were both mentioned three times. Other languages were mentioned in six studies, including Danish, Dutch, Flemish, Polish, and Portuguese. Additionally, in four publications target languages were defined as 'various', especially in studies targeted at the evaluation of large exchange programmes including different participants learning different target languages. More than one target language could be mentioned per study (Table 1).

Given the focus of the review on Europe, it is not overly surprising that most of the target countries are also in Europe: Eighteen studies refer to Switzerland, 18 to Germany, 15 to France and 12 to England as the target country. Other European countries were mentioned either specifically or as various in 26 studies. The USA and Canada are mentioned in 10 studies. Here, too, multiple mentions are possible. It is important to note that not all studies provide a full list of target countries.

### **Exchange settings**

Regarding the exchange setting, most studies investigated a face-to-face setting (42), in nine studies it was a virtual exchange and in another nine studies it was mixed combining face-to-face and virtual elements. Most studies (24) dealt with individual programmes, followed by studies investigating whole class programmes (22) or rotational programmes (5). In eight studies, the exchange format was identified as 'other/ diverse' and in one study as missing.

The duration of the exchange also varied greatly across the publications, from half a day to 18 months. However, most exchanges varied between 1 week and three

**Table 1.** Frequency of target languages mentioned in the publications.

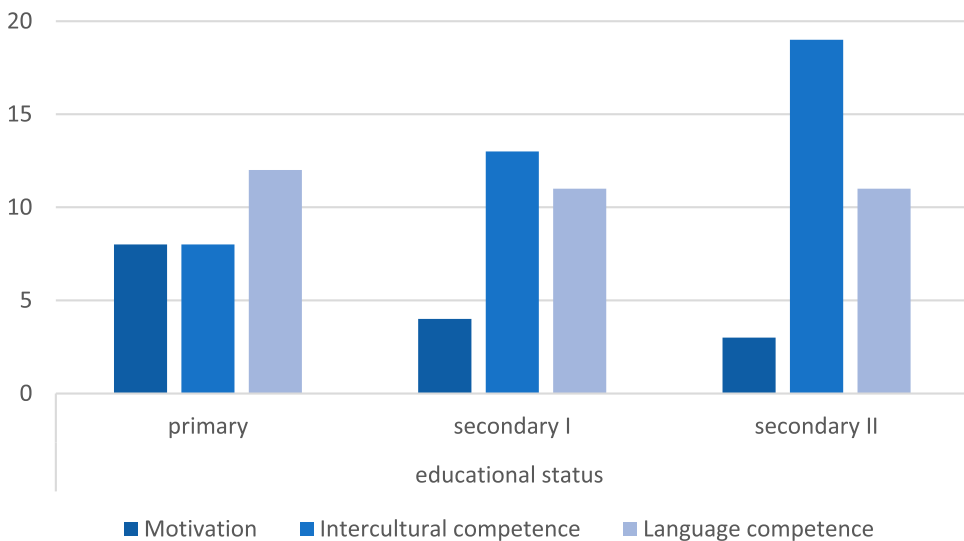
Target language	Frequency
English	31
German	27
French	26
Italian	3
Spanish	3
Polish	2
Danish	1
Dutch	1
Flemish	1
Portuguese	1
various	4

months, with a tendency for longer durations for older participants. This was especially true for durations of half a year or longer. In three overview studies, the duration was hard to classify, as it encompassed exchanges from one week to several months. The information on the duration was missing in the remaining seven studies.

In 32 studies, the sojourn was embedded in the curriculum, e.g. including scheduled preparation and follow-up activities. In these cases, the L2 contact activity usually took place in a school setting and was accompanied by teachers. In four studies it could be concluded from the description that no such embedding took place. As the description of the embedding into curriculum was not mentioned in 24 cases, it is not possible to compare the effectiveness of settings with and without curricular embedding.

### Target variables

The present review covers studies looking at the development of language competences, IC or language learning motivation in relation with student exchange. Of the 60 publications identified, 31 studies look at language competence, but only in 19 of these was there specific evidence of its measurement. Motivation was looked at in 13 studies and IC in 35 studies. It is important to note that the type of assessment of these variables ranged from individual items in a questionnaire or individual interview questions to complete and refined measurement instruments. [Figure 2](#) provides an overview of the number of studies looking at each of the three target variables at the three educational levels considered.



**Figure 2.** Frequency of target variables in the publications by educational status.

### Quality criteria

The methodological rigour of the studies was assessed on the basis of three features of the study design which are relevant for studies investigating the effectiveness of programmes: randomised sampling, pretest/posttest design and the inclusion of a control group. It should be noted that these are mainly quality features for quantitative

studies. None of the included studies fulfilled all of these criteria. However, 33 studies had a pretest/posttest design, of which 16 studies also included a control group. In 4 publications a control group, but no pretest/posttest design was identifiable. 23 studies could not be classified according to these characteristics, of which four studies were quantitative, eight qualitative, one was not defined and 10 were of a mixed-methods design. Finally, 26 publications were identified featuring another quality criterion: purposeful sampling, where some thought and effort was invested in the recruiting and make-up of the sample (two qualitative, eleven mixed methods, and 13 quantitative publications).

### ***Methodology***

All research papers in this review involve the collection and analysis of data. Twenty-one of the studies considered include a quantitative design, 10 are qualitative but most studies use a mixed-methods approach (26). In three studies, the type of data collection or evaluation could not be clearly identified. This may be partly related to the categorisation of information density according to which 4 publications were classified as 'deficient / not intersubjectively verifiable' and 16 studies lacked at least some information necessary for intersubjective verification.

### ***Effectiveness of student exchange for the development of language competences (RQ 2a)***

The following section focuses on the various aspects of language competence which were researched in the identified studies. For the present discussion, these studies are grouped according to the following competence areas: global proficiency, oral proficiency, and written language skills. Many studies focused on more than one competence area and are thus included in different sections. In line with earlier findings, oral competences received the most attention, as this is seen as the area in which participants make the most progress during a stay abroad (cf. Llanes, 2011). Also, improving one's oral competences seems to be a main motivation for students to go abroad in the first place (cf. Allen, 2010). The studies included on L2 oral production commonly focus on such aspects as fluency, accuracy (e.g. grammatical accuracy) and (syntactic) complexity, as well as specific pragmatic skills such as correction and self-correction, requesting and mitigating strategies. Amongst the pool of studies, only Llanes (2016) and Muñoz and Llanes (2014) examine foreign language pronunciation and the impact of learning context and age on learners' degree of foreign accent. The results confirm the superiority of an SA context over an at-home context regarding L2 pronunciation development with medium<sup>3</sup> effect sizes.

### ***Global proficiency***

There are a number of studies that examine students' linguistic progress by means of self-reports (cf. Bachner & Zeutschel, 1994; Braun & Höchle, 2006; Höchle Meier, 2014; Ilg & Dubiski, 2015; Vatter et al., 2011; Weis et al., 2013). While these studies do not provide systematic, quantitative analyses of participants' linguistic progress, they shed light on students' and often hosts' or organisers' (e.g. Ilg & Dubiski, 2015) personal perception of the participants' language competence development. In the above-mentioned studies, most

students reported on the benefits of a language stay for the development of their TL competences (for an exception see Schulze (2010)).

Amongst the studies in the database, there is only one examining global proficiency using language tests. It is the longitudinal intervention study ( $N = 392$ ) by Heinzmann, Paul, et al. (2022) which compared general language competence, speaking and writing skills (see below) of a group of primary school children participating in a student exchange with those of a control group. The intervention included two physical reunions with the partner class and preparatory and follow-up activities. General language proficiency was measured by means of a C-Test before and after the intervention. Results show that the intervention had a positive impact on the pupils' general language competence with a medium effect size even though the physical reunions were very short (2 days each).

### *Oral competences: speaking & listening*

As mentioned above, oral competences are the most frequently investigated language skills in the SA literature. Several studies by Llanes and colleagues investigate the development of oral production during SA, including such aspects as fluency, lexical richness, grammatical complexity and grammatical accuracy (cf. e.g. Llanes, 2012; Llanes et al., 2018; Llanes & Munoz, 2009; Serrano & Llanes, 2015). The studies collectively show a benefit of a stay abroad for L2 oral development with effects ranging from small to large. Moreover, they consider further factors such as age of participants and differing proficiency levels, language learning aptitude and amount of classroom instruction, or length of the stay abroad, which may potentially have an influence on the language learning experience.

In their 2009 study, Llanes and Muñoz studied the effect of a short stay abroad (3–4 weeks) on lower and upper secondary students' ( $N = 24$ ) language competence gains, specifically oral fluency and accuracy, and listening comprehension. They were interested in the effect of the participants' age, pre-departure proficiency level, length of stay, and length of previous stays abroad. Their results showed that participants who stayed for four weeks benefited more than those who were abroad for three weeks, supporting the argument that the length of the stay abroad plays a role with regard to language improvement. Also, they found positive effects of the stay abroad on oral fluency and listening comprehension. This study did not include a control group, thus, preventing conclusions about the superiority of a SA context over a classroom context.

To address this pitfall, a number of studies explore the impact of instruction, for example by including an 'at-home' control group which may receive regular or intensive classroom instruction in the target language but lacks the immersion experience (cf. Evans & Fisher, 2005; Heinzmann, Paul, et al., 2022; Llanes & Munoz, 2013; Llanes & Serrano, 2017; Serrano et al., 2016). These studies suggest that the SA context is more beneficial for the improvement of learners' oral competences than the at-home setting, even though an intensive course at home can be almost as effective as a SA context as the study by Serrano et al. (2016) demonstrates. In their comparative study of explicit grammar knowledge, knowledge of formulaic sequences, and oral production skills of SA learners on the one hand and at-home learners in an intensive course on the other hand, they only found a small effect of SA vis à vis the intensive course.

Llanes and Munoz (2013) provide a comparative analysis of the SA vs. at-home context across two age groups (children vs. adults) with regards to oral and written competences in English ( $N = 139$ ). The results show the SA children's faster development in oral skills, which according to the study authors could be explained by amount, type and quality of L2 contact, as well as amount of instruction received. Another possible explanation for the faster development of the children's oral proficiencies, not mentioned by the authors, is the lower starting proficiency of these. Moreover, results show that SA participants surpass at-home learners with the effect of the SA context being medium. This effect too may also be explained by exposure intensity and amount of input received. SA participants were exposed to the foreign language more, and more intensively, than their at-home peers (Llanes & Munoz, 2013).

That intensity and amount of exposure and interaction importantly affect the effectiveness of a stay abroad is underscored by the study by Sauer and Ellis (2019) which provides a quantitative and qualitative account of two German-speaking adolescent learners' experiences during a 5.5 month stay abroad in New Zealand. Findings showed that opportunities for L2 engagement varied considerably with context – some supporting and others restricting interaction. Especially the presence of co-nationals impeded L2 interaction and required the students themselves to seek out opportunities to use the target language. Also, the two students' L2 development was not linear, but rather measures tended to fluctuate throughout the stay, a finding which points to the limitations of pretest/posttest methodologies and argues for dense data collection procedures.

The 2012 study by Llanes includes the same children's cohort as Llanes and Munoz (2013) and is rather unique as it presents one of the few studies investigating long-term effects of stay abroad experiences on children, in this case 11-year-old bilingual Catalan / Spanish learners of English ( $N = 16$ ). The study design included a pretest (one week prior), posttest (one week after) and delayed posttest (one year after the students' stay abroad in Ireland). The language measure included oral fluency, lexical and syntactic complexity, accuracy and written fluency. The findings showed large effects of the stay abroad on all oral measures. Regarding the long-term effect of the stay abroad, findings suggest that the L2 gains obtained during a SA experience are indeed long lasting with effects still being large for some measures and moderate or small for others. Due to the rather small sample size (9 participants abroad, 7 participants at home), the results are not generalisable. On the other hand, the fact that significant differences and large effect sizes were found in such a small sample suggests that SA has a sizeable impact on oral skills even long-term.

Among the studies including a comparison group, the intervention study by Heinzmann, Paul, et al. (2022) comprises the largest sample ( $N = 392$ ). It suggests that even very short student encounters at primary school level with beginners have the potential to spur language development in terms of speaking skills if they are didactically well embedded as it found a significant effect of the exchange on speaking skills. Given the very short duration of the face-to-face encounters it is not surprising that the effect was small though.

Among the identified studies there are only three studies targeting listening all of which testify to positive effects (Evans & Fisher, 2005; Hodel et al., 2006; Llanes & Munoz, 2009). One of those three studies (Evans & Fisher, 2005) also reported effect sizes and reports large effects of a short stay on listening skills.

### **Pragmatics**

The studies focusing on the development of pragmatic competence examine frequent communication strategies such as making requests, using direct vs. indirect speech acts, and mitigation (cf. Montero et al., 2017; Sell et al., 2019). Montero et al. (2017) for example, researched study abroad students' ( $N=95$ ) development of communication strategies in the foreign language. Their dataset included four groups of participants, two groups of Spanish children aged 10–11 and two groups of Spanish adults aged 19–33. The study included a control group of children and adults staying at home, attending their usual school or university in Spain. The findings suggest that the SA context has a significant and large impact on the use of communication strategies of the children. SA children developed significantly more effective communicative strategies than their stay-at-home counterparts after their three-month programme. They also considerably reduced their use of L1-based strategies. The results also suggest that SA is particularly advantageous for children as they made more progress than adults. At the pretest, significant differences were found between children's and adults' use of communication strategies. Adults had higher effectiveness scores and lower L1-based communication strategies. At the posttest, SA children's scores were close to those of the adults. Contrary to the outcome for children, the results for adults do not show a clear advantage for the SA context. While not finding changes with regards to functional categories of SA students in terms of realisation of requests, Sell et al. (2019) found a significant overall improvement with a large effect size regarding lexical choices of SA students with lexical choices becoming more similar to those of the target community.

### **Writing skills**

There are only a few studies examining the effect of student exchange on writing competencies of younger learners. While some studies on L2 writing skills report on a significant improvement of written competences after a stay abroad (cf. Evans & Fisher, 2005; Llanes et al., 2018) or after an E-Mail exchange (Massler, 2008; Schenker, 2012) with effects ranging from small to large, other studies show no superior effect of a sojourn over the at-home experience on writing skills (cf. Llanes, 2012; Llanes & Munoz, 2013; Llanes & Serrano, 2017). One study reports mixed findings with no effect on task completion but a small positive effect on lexical diversity in writing (Heinzmann, Paul, et al., 2022).

Llanes and Munoz (2013) show that the SA environment appears to particularly enhance children's oral skills, while the at-home context appears to promote the advancement of writing skills in adults, even though effect sizes of the age-context interaction are small. According to the authors, the length of the stay abroad, two months, may not have been sufficient for gains in L2 writing development to occur, whereas they seem to be enough for oral gains.

### **Effectiveness of student exchange for the development of language learning motivation (RQ 2b)**

There are far fewer studies investigating the effects of student exchange on language learners' motivation than those dealing with its effects on IC and language competence. The identified publications reporting on the development of language learning motivation used either purely quantitative methods or mixed methods. The comparability of

the study results proved difficult as most researchers do not specify which theoretical concepts they draw on and how they operationalise the construct 'language learning motivation' (for exceptions, see Heinzmann et al., 2014; Heinzmann, Hilbe, et al., 2022; Paul-Frischknecht & Aguirre, 2014; Salzmann, 2022; Schmid, 2008).

### *Primary school*

At primary school level, two comparable Swiss intervention studies (Heinzmann, Hilbe, et al., 2022; Paul-Frischknecht & Aguirre, 2014) have shown that short-term exchange programmes consisting of short face-to-face encounters (1–2 days) embedded in preparatory and follow-up lessons containing remote, asynchronous contact activities (such as writing letters) do not significantly increase pupils' motivation to learn the target language. In both these intervention studies conducted with 5th and 6th graders in German- and French-speaking Switzerland (four classes in Paul-Frischknecht and Aguirre (2014) and 392 pupils in Heinzmann, Hilbe, et al. (2022)), the pupils were motivated to learn the foreign language already at the beginning of the exchange with no significant impact on motivation as measured after the exchange.

Salzmann (2022) reports more encouraging results in another intervention study ( $N = 277$ ) in Switzerland where the motivational development of the intervention group was slightly more positive than that of the control group. Since no significance tests were run, we do not know whether the modest differences found in the development of the two groups are significant, however. Results from Pelz's (1999) evaluation study conducted with 600 classes in Germany and France showed that according to the teachers, multiple whole-class personal encounters do neither lastingly increase the motivation of the pupils to learn the foreign language nor their interest in the neighbouring country. Analogous to the other studies at primary level, the encounters in Pelz' study were very short (1/2 or 1 day), with only some teachers organising one-week exchanges or several encounters during the year. In contrast to Paul-Frischknecht and Aguirre's (2014) and Heinzmann, Hilbe, et al.'s (2022) studies, Pelz' (1999) study did not involve a control group. Despite this methodological limitation, her results largely confirm those of the quasi-experimental studies.

Two studies focusing on the impact of online exchange formats found that no groundbreaking improvement can be expected through the use of technologies alone. Jauregi and Melchor-Couto (2018) investigated a two-month-long telecollaboration project between secondary school learners of Spanish in the Netherlands and primary school learners of English in Spain ( $N = 39$ ). The results showed that the Spanish learners of English perceived the exchange to be much more useful and their attitudes towards communicating in English were more positive. The authors report that the two teachers involved in the project differed in the way they implemented the project and the amount of guidance they offered with the Dutch teacher offering less guidance or scaffolding. On the other hand, they stress that the age difference between the groups was not an issue according to self-reports of the participants. The authors concluded that the Dutch learners might have needed further guidance and monitoring to fully benefit from the exchange, which points to the crucial role of the teacher for exchange activities at lower levels of education. The authors still conclude that their study shows the value of telecollaboration projects across school levels. In Botturi et al.'s (2018) study conducted in three Swiss language regions, the interviewed teachers reported that a one-year-programme



(whole-class virtual and face-to-face exchange) motivated their 4th to 6th grade pupils ( $N = 113$ ) to learn the foreign language. In contrast, the quantitative student survey data showed significant variation across classes in terms of motivational development. While the motivation to learn the foreign language increased in some classes, it remained stable in others and even decreased in most of them. Overall, girls, pupils in the French- and Italian-speaking regions of Switzerland, pupils with multilingual backgrounds and those who had already spent a prolonged stay in the L2 region were significantly more motivated. The authors conclude that negative motivational development seems to be triggered by a generally negative attitude towards foreign languages, the perceived complexity of the foreign language or a preference for English over other foreign languages, a preference that was evident in the German-speaking classes but not the French- or Italian-speaking classes. Overall, their data suggest that the value of technology-enhanced exchange formats depends on a variety of contextual elements, e.g. attitude towards foreign languages, previous experiences in a different linguistic region as well as the teacher.

### *Lower secondary school*

At lower secondary school level, Schmid (2008) investigated two classes ( $N = 18$ ) that participated in a one-week face-to-face exchange. Similar to the results reported for primary school level, the short-term exchange did not necessarily lead to an increase in general language learning motivation. The only motivational aspect that profited from the exchange was the motivation to speak the language. Baenziger (2012) conducted a comparable study with two lower secondary classes ( $N = 22$ ) from the French-speaking and German-speaking regions of Switzerland, respectively. The student questionnaire yielded mixed results regarding the impact of the one-week face-to-face exchange on language learning motivation. Answers to one item with regards to motivation<sup>4</sup> suggested no significant impact on the French-speaking students' motivation to learn German while another question<sup>5</sup> pointed to a slight increase in German language learning motivation as two thirds of the students felt they were more motivated after the exchange. The large-scale cross-sectional Swiss survey study by Braun and Höchle (2006) also yielded mixed findings as somewhat less than 50% of the students ( $N = 2,100$ ) reported an increased interest in the TL after the exchange. Given the small sample in two of the three studies, the general lack of a control group and the purely descriptive statistics at item-level, the results of these studies need to be interpreted with caution, however.

### *Upper secondary school*

At upper secondary school level exchange activities seem to have a greater potential to positively affect learners' language learning motivation. Heinzmann et al.'s (2014) quasi-experimental study ( $N = 540$ ) investigated a range of L2 contact activities, including language stays, student exchanges, internships etc., at Swiss upper secondary schools. The results showed that these immersion activities have a positive impact on adolescents' language learning motivation both short term and longer term (three months after return). In Hodel (2004), students ( $N = 50$ ) report that their motivation to learn French after a stay abroad was higher than before. Ilg and Dubiski (2015) conducted a large-scale survey among secondary and upper secondary school students living in Germany,

Poland, France and other non-specified European countries ( $N = 5,206$ ) who participated in face-to-face youth exchanges lasting between four days and three weeks. Questionnaire answers showed that 77% of the participants were more motivated to learn the foreign language after the exchange.

### ***Effectiveness of student exchange for the development of intercultural competence (RQ 2c)***

Another frequently investigated area in studies dealing with the impact of student exchange on primary and secondary school students is the development of IC. The identified studies employ a range of different constructs or operationalizations of IC or do not elaborate on the construct employed at all, which makes it difficult to compare findings across studies.

#### ***Primary school***

At primary school level only three quantitative studies could be identified, two of them with a pretest/posttest design but lacking a control group (Botturi et al., 2018; Krok et al., 2010) and one study with a control group but lacking a pretest (Alcaraz-Mármol, 2020). Despite lacking a methodological design allowing robust conclusions to be drawn regarding the effectiveness of student exchange, these studies can provide us with tentative evidence.

Of these three studies, the study by Alcaraz-Mármol (2020) which compared the self-reported IC of primary school children participating in a virtual telecollaboration through E-twinning with that of a control group ( $N = 38$ ) is the only study pointing to a positive effect of the exchange experience. In their study, the experimental group had a statistically higher self-reported IC than the control group. Given the fact that no pretest of the intercultural competences of the two groups was administered before the exchange, these results can hardly be considered generalisable. Botturi et al.'s (2018) study ( $N = 113$ ) deals with the effects of a hybrid exchange format between partner classes from different linguistic regions of Switzerland. The study mainly focuses on the motivational and linguistic effects of this programme but also touches on IC. The results pertaining to IC are presented in terms of two items only,<sup>6</sup> which makes it difficult to draw reliable conclusions. The findings indicate that there is considerable variation among classes in terms of IC, and, overall, there is no clear trend of students becoming more proficient in this aspect.

The rather disappointing or at best variable outcomes in terms of intercultural learning in online exchange settings and the importance of the teacher in this respect also emerge in a qualitative study at primary school level. Piipponen and Karlsson (2019), investigating a seven-week virtual storycrafting exchange between Finnish and Scottish primary school pupils, identified both sensitive (framed as interculturally competent) and defensive responses to stories of the partner class, while the former prevailed. Their analysis also revealed that the role the teacher took on during the storycrafting sessions had an impact on how the intercultural encounters unfolded. They conclude that the exchange setting has potential for intercultural learning but can only unfold this potential if the teacher creates a participatory space enticing non-judgmental intercultural encounters.

That face-to-face exchange formats are not necessarily more effective for intercultural learning at primary school level is suggested by Krok et al.'s (2010) longitudinal study ( $N = 75$ ) comprising a pretest, posttest and delayed posttest but no control group. The study investigated face-to-face encounters lasting between 6-28 days of German children with children from other countries without the explicit aim of language learning with rather discouraging findings. Forty one percent of the children reported not wanting any contact with children from abroad at the end of the encounter and only one fifth of all respondents were still in contact with at least one child from another country after returning to their everyday lives. The authors conclude that intercultural learning is a long process and that not all learning experiences are sustainable. Importantly, Krok et al.'s (2010) study identified different learning trajectories. The data suggests that children who have more resources (broader range of experiences and foreign language skills) because of their age, experiences and attitudes, interact and learn more successfully in intercultural encounters. The findings by Krok et al. (2010), thus, once again point to the importance of careful planning of intercultural encounters and careful guidance and scaffolding of the learning process.

Pelz' (1999) large-scale mixed-method evaluation of an exchange programme between primary school children from neighbouring areas in Germany and France (600 classes) also found it to be rather ineffective for intercultural learning. Despite personal encounters, the programme did not manage to establish a stronger connection to the other language region beyond the encounters. Even though this study boasts a large sample and a triangulation of methods, a lot of the data is not systematically analysed. Furthermore, the study lacks a pretest or a control group for which reason its results have to be interpreted with caution.

### *Lower secondary school*

At lower secondary school level there are only a few studies looking at intercultural learning in a student exchange setting and the results of these few studies, none of which are (quasi-) experimental studies with a control group, are inconclusive.

Ilg and Dubiski's (2015), Massler's (2008), Schulze's (2010) and Vatter et al.'s (2011) studies all report rather disappointing results in terms of intercultural development. Vatter et al. (2011) for example, examined organisers' and participants' views and opinions on a 2-4-week long interregional exchange between individual students from Germany and individual exchange partners in France, Luxembourg or Belgium by means of a survey after the exchange. The survey results demonstrate that the goal of intercultural learning only seems to be attained by part of the students as there were different developmental trajectories in terms of IC. While certain students mentioned experiencing heightened tolerance and openness, reflecting on and partially dismantling their stereotypes, other students' comments indicate that existing stereotypes were reinforced or new ones developed, leading to an increased sense of intolerance. The authors indicate the age of the students, difficult relationships with exchange partners and lack of preparation as well as of support during the exchange and follow-up reflections as possible reasons why some students did not benefit.

Unlike Vatter et al. (2011), Schulze's (2010) panel study ( $N = 79$ ) employed a pretest/posttest design and, hence, was able to analyse actual change. Given that no control group was involved, it cannot be ascertained whether identified changes are attributable

to the exchange programme, however. As the exchange programme in Schulze's study was very short, consisting of a 7-day excursion of German secondary students to Great Britain with accommodation in host families, it is at least likely that changes identified in such a short time span are attributable to the exchange. This study differs from most of the other studies cited here in that it examines the effect of a school trip rather than a classic student exchange and that the students involved were not in contact with students their age from the host country (except possibly in the host family). The results of Schulze's study regarding intercultural learning are also mixed. Overall, both students and accompanying teachers felt that the trip was successful with regard to broadened cultural knowledge, stronger linguistic motivation and rising interest. But the data also showed that not all students profited equally. Depending on the age of the students, the excursion had a differential effect on stereotyping. While it was able to at least partially correct stereotypes in the older students, these were built up in the younger ones, a finding that ties in with those of Vatter et al. (2011). Similar to Vatter et al., Schulze argues that intensive preparation and follow-up work on school stays abroad is essential for successful intercultural learning. Indeed, the participants in his study reported not having felt optimally prepared for the trip. Furthermore, they lamented the lack of contact with locals their age. Schulze (2010) concludes that accommodation in host families is not necessarily sufficient for immersion in the local culture. Rather, he argues, to give the students a better insight into the everyday life of the target culture, the classic student exchange, in which the students also come into contact with their peers and their living environment, is best suited.

Two more small-scale studies reporting on the impact of a student exchange at lower secondary school level with more encouraging findings are the studies by Gautschi (2019) and Vasbø (2013). Gautschi (2019) reports on the effect of a 10-day exchange between students from the German- and French-speaking part of Switzerland ( $N=12$ ). The results of participant questionnaires and close interviews with 4 students show that the language exchange had a positive effect on the intercultural competences of all students even though the amount of change is dependent on the duration of the exchange. Vasbø's (2013) small-scale ethnographic study ( $n=6$ ) using interviews and participant observation does not provide sufficient information on the exchange programme and the data collection to make it intersubjectively verifiable. Nonetheless, it is worth mentioning as it deals with a very particular and interesting student population, namely potential school dropouts. Six potential Norwegian school dropouts participated in a face-to-face exchange with Brazilian youth. The findings suggest that these youth encounters provided a positive learning experience for these potential school dropouts and an effective learning context for the development of social, communicative and cultural knowledge and skills needed to live together in a multicultural and complex world.

In addition to the above-mentioned studies examining face-to-face encounters, the studies by Jauregi (2015) and Kohn and Hoffstaedter (2017), which are both based on the TILA project (Telecollaboration for Intercultural Language Acquisition) as well as the studies by Massler (2008) and Preiser (2015) focus on the display of intercultural communicative competence/effectiveness in telecollaboration projects at the lower secondary school level.

Massler (2008) and Preiser (2015) both reported rather disappointing results in terms of intercultural development in their qualitative studies of a virtual exchange. On the basis of website posts, lesson observations, student questionnaires and interviews, Preiser (2015)

found the online exchange to offer little opportunity for substantial intercultural learning, even though students were friendly with each other and interested in one another. Intercultural learning remained at a superficial level with students mainly discovering visible aspects of culture but not dealing with invisible aspects. In the study by Massler (2008) most students did not establish meaningful and lasting social rapport. None of the students kept in touch after the project, and the students reported that they did not perceive the exchange as a meaningful encounter. Among the implications for further such projects, both Massler (2008) and Preiser (2015) highlight the importance of the teacher for more profound learning to happen. Teachers should be familiar with project-oriented work and familiarise students with it. They should actively encourage and monitor the creation and maintenance of rapport, they should identify dysfunctional communication and intervene early on, and sensitise students about the pitfalls of online interaction (such as lack of social presence). Furthermore, both Massler (2008) and Preiser (2015) advocate including a face-to-face encounter in virtual exchange projects. In addition, Massler (2008) suggests carrying out such projects with more advanced and older learners instead.

Due to the lack of (a) an experimental design and (b) a longitudinal analysis, the study by Kohn and Hoffstaedter (2017) cannot inform us about the effectiveness of such virtual projects for the development of intercultural communicative competence. What it does show, however, is that students exhibit IC in their virtual conversations through expressions of empathy, encouragement and face-saving conversational moves. As the authors point out, most conversation partners succeeded in creating a comforting atmosphere of high rapport quite in contrast to the findings by Massler (2008). With regards to interactional partners, students reported feelings of insecurity and anxiety when communicating with native speakers. In contrast, they felt more at ease when communicating with other non-native speakers, a finding which demonstrates the potential of lingua franca telecollaboration for improving functional communicative competence, text and media competence.

Jauregi (2015) also stresses that, overall, a cooperative communication attitude prevailed in the exchanges analysed. While it does not become clear from the publication whether the study employed a pretest/posttest design, this seems to have been the case as Jauregi (2015) states that the telecollaboration exchanges had a positive impact on pupils' intercultural communicative competence and awareness. Jauregi's study, thus, is the only study at the primary and lower secondary school level claiming positive effects of a virtual exchange on intercultural learning.

### *Upper secondary school*

At upper secondary level there is more empirical evidence on the impact of student exchange on IC, due to a higher volume of studies at this level. The majority are quantitative or mixed-method studies even though some of these operate with very small samples (e.g. Mapes, 2020; Schenker, 2012). Most of the identified studies are cross-sectional retrospective survey studies (Bachner & Zeuschel, 1994; 2009; Genkova & Schubert, 2020; Gözpınar, 2018; Meyers et al., 2020; Thomas et al., 2007; Weichbrodt, 2014; Weis et al., 2013), a few are longitudinal studies tracking development with pretest and posttest measurements (Mapes, 2020; Maréchal, 2007; Ogay, 2000; Schenker, 2012), and only two are longitudinal experimental studies comprising a control group as well as a

pretest and a posttest, in one case also a delayed posttest (Hammer, 2005; Heinzmann et al., 2014; 2020). Both experimental studies report positive effects of a student exchange on the development of IC.

The study by Hammer is a large-scale quasi-experimental study ( $N = 2,000$ ) examining the impact of an exchange year organised in the context of the American Field Service network (AFS) on intercultural learning, operationalised as progress on the Intercultural Development Inventory scale (see Hammer et al., 2003). The findings suggest that the exchange year is effective for intercultural learning. In general, the students with the AFS experience advance to higher levels of intercultural sensitivity at nearly twice the rate of the control group. AFS returnees have a much greater intercultural awareness than their friends of similar backgrounds who did not go abroad.

Heinzmann et al. (2014, 2020) investigated the impact of different types of short-term stay abroad programmes (language stay,<sup>7</sup> student exchange, internship, au-pair work) with differing lengths on IC basing their instrument on Byram's (1997) model of intercultural communicative competence. Most of the programmes lasted one week to two months. Even though the programme duration in this study was significantly shorter than in the study by Hammer, the analyses across the intervention and control group consistently indicated that a short-term stay abroad has a positive influence on the development of the participants' intercultural attitudes and skills. However, the analyses also showed that the development of the group of participants with the shortest programme duration (1-2 weeks) did not significantly differ from that of the control group. A stay of 1-2 weeks, consequently, does not seem to be particularly useful for the development of IC. Further predictors identified for students' learning trajectory were their level of IC before the stay, the amount of contact with target language speakers and the amount of target language use. Intercultural learning was greater among students who already exhibited higher levels of IC at the outset and among students who frequently interacted with TL speakers, and, simultaneously, frequently used the TL.

The findings from these experimental studies are supported by a series of cross-sectional retrospective studies investigating SA programmes of different design and length during which the participants report having become more interculturally competent, aware or sensitive as a result of the SA experience (Brougère et al., 2006; Genkova & Schubert, 2020; Gözpinar, 2018; Weis et al., 2013). Genkova and Schubert's (2020) cross-sectional study also points to the importance of the length of stay for intercultural learning as the level of IC was higher, the longer participants were abroad.

The cross-sectional studies by Bachner and Zeitschel (1994, 2009), Thomas et al. (2007) and Weichbrodt (2014), examined the long-term impact of SA, surveying former SA participants years after their return. Findings from these studies lend support to the notion that the sojourn experience contributes to positive and long-lasting attitudinal, behavioural and cognitive changes in the majority of individual participants.

In contrast to the study by Thomas et al. (2007), where the long-term impact of short-term group exchanges was found to be comparable to that of longer individual programmes, duration turned out to be a pivotal variable in Bachner and Zeitschel (1994, 2009). The longer the sojourn of participants, the greater the depth and quality of impact reported by them. In addition, participation in exchange programmes also seems to lead to the pursuit of follow-up exchange activities and can thus act as a motivator for increased mobility, as suggested also by Weichbrodt (2014). As the authors point

out, it is important to bear in mind that exchange is a variable phenomenon with outcomes that can vary due to the complexity of activities and circumstances. Not every exchange is the same and has the same effect. Important variables that influence impact are programme duration, type of programme, programme goals, nationality & gender of participants. Naturally, long-term impact studies are fraught with methodological challenges (difficulty of tracking individuals, impossibility to control for confounding variables, memory loss of participants) and if they are not longitudinal and do not comprise a control group they cannot establish clear cause-and-effect relationships.

Unlike the cross-sectional studies discussed so far, the large-scale study by Meyers et al. (2020) ( $N = 16,997$ ) on Erasmus+ and Youth in Action programmes was not primarily interested in how effective SA is for the development of IC but in how effective it is for socially disadvantaged students in comparison to other students in terms of different learning domains,<sup>8</sup> one of them being intercultural interaction. Participants were between 18 and 30 years old, hence also including post-secondary students. The study did not assess intercultural skills in a systematic way based on a theoretical framework of intercultural communicative competence, however, and the scale employed contained two items only. The findings suggest that social inequalities have – if any – only a very small effect on self-assessed learning in terms of intercultural interaction. What turned out to be important instead are: the age of the participants, the duration of the activity, and the type of programme. With increasing age (adulthood), lower outcomes are reported. The longer the duration of the activity is, the higher the outcomes in terms of intercultural interaction. Finally, the project level had the highest explanatory power for learning outcomes, which means that the programme type is more important than individual differences for learning trajectories.

Apart from these cross-sectional studies, there are a few longitudinal small-scale studies that deal with the intercultural development of upper secondary students as a result of student exchange. Maréchal (2007) reporting on a short (3 days in total) face-to-face exchange between French- and Dutch-speaking students in Belgium and Mapes (2020) examining a four-week face-to-face exchange (2×13 days) between 29 Danish and U.S. upper secondary students, report cautiously optimistic outcomes. Schenker's (2012) dissertation is the only study examining a virtual exchange at upper secondary level ( $N = 50$ ) involving a semester-long telecollaborative project combining the use of one-on-one, group-on-group, as well as reflective computer-mediated communication tools. It is the only study at this level that does not report positive effects.

Finally, Ogay's (2000) qualitative, hermeneutic study employing interviews, questionnaires, photos and intercultural sensitivity cards to gauge the intercultural dynamics experienced by 8 French-speaking Swiss students spending four weeks in German-speaking Switzerland points to a great diversity of experiences of participants which does not allow for generalisations. Based on the interview data she hypothesises, however, that a strong attachment to one's own cultural group favours an orientation towards cultural differences while a weak attachment to one's own cultural group favours a minimisation of cultural differences. Furthermore, a good locus of control and pronounced self-awareness seem to be essential for successful intercultural learning. Ogay (2000) stresses that it is difference, or rather the conception of difference which is at the core of the intercultural dynamic. Recognising the other as different, conceiving of their otherness, defines the possibility of discovery of the other. She criticises that in an effort to be 'interculturally

correct', harmonious intercultural communication is overvalued. This hinders reflection on difference and intercultural learning, which can only happen if we are involved in the relationship with the 'Other', including the destabilising elements. This view is shared by Maréchal (2007) and Perrefort (in Brougère et al., 2006) who also argue for a less naïve intercultural pedagogy that is not too focused on harmony and the demise of stereotypes but acknowledges the value of conflict for learning.

## **Discussion and desiderata for future research**

In light of the importance that is attributed to student exchange at the level of educational policy the available empirical evidence on the effectiveness of student exchange in primary and secondary education can be considered scarce. Firstly, the number of studies dealing with certain outcome variables (IC, motivation, language gains) and skills areas (e.g. listening, speaking, writing) on a given educational level tend to be very limited, so that it is difficult to draw reliable conclusions for the effectiveness of student exchange for a certain age group and learning area. Secondly, the studies included exhibit various methodological limitations, including small sample sizes, a lack of longitudinal and (quasi-)experimental designs, self-selection bias, dearth of multivariate analyses, lack of reporting of effect sizes, especially in the areas of motivation and IC, a scarcity of language testing in studies focusing on linguistic gains, and inadequate documentation of methods and procedures in some cases, as well as an over-reliance on self-reports. Considering all this, it must be concluded that our understanding of the effectiveness of student exchange in primary and secondary education is still sketchy and fragmentary.

This is similar to the situation lamented by Kinginger (2009) more than a decade ago in relation to SA research primarily looking at tertiary level students which, according to her comprehensive review largely consisted of a patchwork of small-scale studies with convenience samples of students from the researchers' own institutions. Since then, this strand of research has seen a dramatic increase in systematic research and has profited from several large-scale collaborative efforts as Kinginger (2019) points out. This is exactly what we also need in terms of research on student exchange at primary and secondary school level. After summarising what we know based on the evidence available today, we will outline what seem to be important avenues for future research to us. Since many of the original studies operated with small samples and did not provide effect sizes or confidence intervals and given that these measures were not calculated for this review either, the summarised evidence must be considered tentative.

### ***Language gains***

With regard to the development of different linguistic skills somewhat robust evidence for the effectiveness of student mobility can at present only be reported for the development of oral proficiency at different levels of education as this is the area where most research has been conducted and positive effects of a student exchange have been demonstrated by a series of experimental, methodologically sound studies. These studies suggest that student exchange has a small to large effect on students' speaking skills and a medium



effect on their pronunciation. Effects tend to be small if the stay is very short (less than a week) and medium or large if it is longer (several weeks).

The few studies identified looking at general proficiency, listening skills and pragmatic competence all demonstrate positive effects of a student exchange, but clearly, more research replicating these findings is needed to arrive at reliable conclusions. While large effects have been reported by studies on listening skills and pragmatic competence, medium effects have been reported for general proficiency. Writing skills, too, only marginally feature in research on student exchange and results are contradictory, so that here, too, clearly more research is needed. If effects were found, they range from small to large.

Generally speaking, the reported effect sizes are encouraging in particular considering the short nature of most exchange programmes in which this student population is engaged suggesting that student exchange can be an important and effective learning context for young learners. The sizeable effects identified may be related to the lower initial proficiency level of the participants compared to that of higher education students which have been the focus of most SA research to date as progress tends to be slower and harder to achieve at high levels of proficiency (Baker-Smemoe et al., 2014; Köylü, 2023; Vande Berg et al., 2009).

Our discussion of studies investigating linguistic gains also highlights several factors which may influence gains in L2 development: duration of the stay abroad, (amount) of foreign language instruction and amount of exposure to the target language, participants' age, starting proficiency, and aptitude in language learning.

### ***Language learning motivation***

In summary, the findings across studies suggest that especially at primary and lower secondary school level (except for Ilg and Dubiski's (2015) study), language exchange programmes do not necessarily lead to increased language learning motivation, irrespective of the exchange format. In contrast, at upper secondary level, different exchange activities positively impacted students' language learning motivation.

However, studies dealing with the impact of exchange activities on learners' language learning motivation are still scarce with many of the identified studies only mentioning this issue in passing. Furthermore, there are only a limited number of studies comprising a control group as well as a pretest and posttest, which is a prerequisite for drawing reliable conclusions about the effectiveness of a particular programme. Considering this, clearly, more empirically sound research on the potential of student exchange programmes to boost language learning motivation is needed.

### ***Intercultural competence***

The identified studies employ a range of different constructs of IC which makes comparisons of results across studies difficult. They also span a range of very diverse programmes, which once again are difficult to compare. As Meyers et al.'s (2020) study illustrates programme characteristics (such as duration, intensity of encounter) are important. Indeed, programme type turned out to be more important than individual differences for learning outcomes in terms of intercultural interaction.

Nonetheless, some valuable trends/indications can be gleaned from the synthesis. At primary school level only three quantitative studies could be identified, none of which have a methodological design allowing robust conclusions to be drawn regarding the effectiveness of exchange programmes. The available empirical evidence at this level suggests that exchange programmes are not very effective for intercultural learning.

At lower secondary school level there are only a few studies looking at intercultural learning in a student exchange setting and the results of these few studies, none of which are (quasi-) experimental studies with a control group, are inconclusive with some studies finding stays abroad to be effective and others not.

Overall, empirical evidence at upper secondary school level is more robust due to a higher number of studies carried out. The available evidence suggests that face-to-face exchange formats seem to be effective for intercultural learning even in the case of rather short programmes and for different kinds of learners. Furthermore, findings from the few available long-term studies suggest that the positive effects of such exchange programmes are long-lasting. Importantly studies at this level also provide evidence on a number of key variables affecting the effectiveness of programmes, the most important ones being duration of stay, type of programme, age of participants, and starting conditions of participants.

For the most part virtual settings do not seem to be very profitable for intercultural learning either at the primary school level or the lower and upper secondary school level. Numerous studies at different educational levels point to the importance of a careful organisation and coaching of the students for learning to happen (e.g. Krüger-Potratz & Wagner, 2018; Preiser, 2015; Schulze, 2010). At primary level the role of the teacher is particularly eminent. He/she needs to be able to create a space for intercultural learning. This argues for the necessity to introduce or strengthen mobility didactics/methodology into teacher training programmes.

## **Directions for future research**

### ***Methodology:***

Further research with larger samples and more robust methodological designs comprising a pretest and posttest, a control group and randomisation is needed to deepen our understanding of the potential gains (linguistic, motivational, intercultural) resulting from a student exchange. For example, controlled, experimental studies could help elucidate if the ineffectiveness of exchange programmes identified at the primary school (and partly secondary school) level regarding the development of intercultural competence and language learning motivation may be attributed to the age of the participants and the demands that this poses for pedagogic guidance and scaffolding or the duration of the exchange, which tends to be very short at these levels of education. In fact, several authors suggest that the investigated exchanges may have been too short to record an impact and previous research at upper secondary school level underscores the importance of the duration of the programme. Studies, such as those of Llanes and colleagues, which compare the effectiveness of SA programmes of equal length for different age groups (primary, secondary, and tertiary level students), provide a fruitful methodological design to disentangle effects of age and duration.

In view of the over-reliance on self-reports even in studies dealing with language gains, the field would also profit from the use of alternative ways of assessment, in particular but not only for intercultural competence, such as participant observation, analysis of text productions, role-plays, etc.

Once a more comprehensive body of methodologically sound studies on primary and secondary school students is available for the target variables under examination, a meta-analysis of these primary studies to examine the effectiveness of study abroad for different learners (e.g. different age groups, different pre-departure proficiency level), different programmes (regarding length, type of accommodation, etc.) and outcome measures would be highly desirable.

### ***Long-term & cumulative effect of student exchange***

There are comparatively few studies employing a delayed posttest several months after the return of participants to investigate the retention of learning gains or cross-sectional studies looking at the long-term impact of a student exchange, so that more research on retention and long-term effects is warranted. As part of long-term effects, future research should also try to probe into potential cumulative effects of such programmes, an area that, hitherto, has not been studied at all in primary and secondary education. But there is some preliminary coincidental evidence from studies looking at long term effects that a cumulative effect of SA activities could exist. Weichbrodt (2014) showed that exchange activities seem to act as a motivator for increased mobility in the future as participants in exchange programmes tend to participate in more exchange activities at a later stage. Similarly, Heinzmann et al. (2020) found that three months after their return exchange students who had never been abroad, or only once or twice demonstrated significantly lower willingness to engage with people from another cultural background than students who had already been abroad more often. If such as yet unsystematic evidence on a potential cumulative effect of exchange activities is corroborated by future research this has important implications for policies as it suggests that student exchange should start early in learners' education and be repeated throughout their educational biography. It also means that short exchange activities at early levels of education that may fail to demonstrate an immediate impact may, nonetheless, exert an influence by motivating participation in future exchange programmes and facilitating learning in subsequent activities.

### ***Additional language skills: listening, writing, pronunciation***

In line with a general trend (cf. Llanes, 2012), the majority of the studies in our database examine oral competences, such as fluency first and foremost, as well as lexical accuracy and complexity. In comparison, there is no study that focuses on L2 pronunciation gains. There is a need for future studies on the development of pronunciation, especially if one considers that students' foreign language pronunciation is often what is first assessed by others and a clear reflection of the amount of one's immersion.

The number of studies which look at the effect of a stay abroad on students' writing skills is small. Especially in light of the growing trend to 'transfer' student exchange to a virtual setting, writing may gain significance as a mode of communication.

In comparison to productive skills, the effect of student exchange on receptive skills, such as listening and reading is poorly studied – this may be partially due to the difficulty in investigating the relationship (even correlation) between the immersive setting and the said skills. Still, additional studies on students' development of listening comprehension and reading skills would shed light on the benefits of student exchange on language learning more holistically.

### ***English as a Lingua Franca***

Given the spread of English across the globe, most interactions in English are English as a Lingua Franca (ELF) interactions. While study abroad research at tertiary level has developed an interest in the potential of ELF contexts for English language learning in recent years, this context with its constraints and affordances has not been studied to any great extent for student exchange at lower levels of education (for an exception see Kohn and Hoffstaedter (2017)). In light of the encouraging findings relating to EFL contexts produced by SA research on tertiary level learners, it seems worthwhile that future research investigates the learning potential of this context for younger learners.

### ***Individual differences***

With the exception of Vasbø (2013) and Meyers et al. (2020), studies up to date do not focus on the effectiveness of student exchange for particular student populations. Therefore, we still do not know enough about which kinds of exchange programmes are effective for which kinds of learners and which kinds of learning goals. Future studies could investigate if specific groups of students (e.g. potential school dropouts, special needs children, socially disadvantaged children, multilingual children) profit more from study abroad activities than the average student or what types of programmes are effective for what type of learners. This necessitates large-scale studies involving different types of programmes and/or different kinds of learners. Only large-scale studies can disentangle the respective importance of different individual and contextual variables for learning outcomes. In order to realise such large-scale studies national and international research collaborations and funding opportunities are key.

### **Notes**

1. By Europe we mean the EU and a limited number of associated countries, e.g. Switzerland, Norway, and the UK.
2. We excluded studies dealing with non-European students studying abroad in Europe as our review focuses on the European educational context and, hence, students enrolled in European schools.
3. Llanes (2016) reported a large effect size (Cohen's  $d = 0.94$ ). According to the field-specific recommendations by Plonsky and Oswald (2014) this corresponds to a medium effect. Where field-specific recommendations are available we will report these.
4. I am not motivated at all / rather unmotivated / rather motivated / very motivated for learning German.
5. Since the exchange, do have the impression of being more motivated to learn German than before?

6. 'It is enough to know only my own language.' / 'has your attitude towards the foreign language region changed positively?'
7. During a language stay, students typically travel to another linguacultural area where they attend a language course and stay with a host family.
8. (1) active participation, (2) learning and personal development, (3) intercultural interaction.

## Disclosure statement

No potential conflict of interest was reported by the author(s).

## ORCID

Sybille Heinzmann  <http://orcid.org/0000-0002-2842-7466>

Catherine Ferris  <http://orcid.org/0000-0001-6496-9214>

Kristina Ehram  <http://orcid.org/0000-0002-3888-0176>

## References

- Alcaraz-Mármol, G. (2020). Developing intercultural communication in the EFL primary education classroom: Internationalization through virtual team collaboration with eTwinning [Desarrollo de la comunicación intercultural en inglés como lengua extranjera en educación primaria: Internacionalización a través de la colaboración virtual en equipo con eTwinning]. *TEJUELO. Didáctica De La Lengua Y La Literatura. Educación* [TEJUELO. Didactics of Language and Literature. Education], 32, 147–170. <https://doi.org/10.17398/1988-8430.32.147>
- Allen, L. Q. (2010). The impact of study abroad on the professional lives of world language teachers. *Foreign Language Annals*, 43(1), 93–104. <https://doi.org/10.1111/j.1944-9720.2010.01062.x>
- Bachner, D. J., & Zeutschel, U. (1994). Utilizing the effects of youth exchange: A study of the subsequent lives of German and American high school exchange participants. *Occasional Papers on International Educational Exchange* 31. Research Series.
- Bachner, D. J., & Zeutschel, U. (2009). Long-term effects of international educational youth exchange. *Intercultural education (London, England)*, 20(sup1), S45–S58. <https://doi.org/10.1080/14675980903370862>
- Baenziger, M. (2012). *Un échange linguistique peut-il influencer le rapport au savoir et la motivation des élèves?* <https://books.google.ch/books?id=wSJtswEACAAJ>
- Baker-Smemoe, W., Dewey, D. P., Bown, J., & Martinsen, R. A. (2014). Variables affecting L2 gains during study abroad. *Foreign Language Annals*, 47(3), 464–486. <https://doi.org/10.1111/flan.12093>
- Bennett, M. J. (1993). Towards ethnorelativism: A developmental model of intercultural sensitivity. In R. M. Paige (Ed.), *Education for the intercultural experience* (2nd ed., pp. 21–71). Intercultural Press.
- Borràs, J., & Llanes, À. (2021). Re-examining the impact of study abroad on L2 development: A critical overview. *The Language Learning Journal*, 49(5), 527–540. <https://doi.org/10.1080/09571736.2019.1642941>
- Botturi, L., Kappler, D., & Negrini, L. (2018). Digitally-supported language exchanges in primary school: The AlpConnectar project. *Studies in Second Language Learning and Teaching*, 8(4), 795–843. <https://doi.org/10.14746/sslT.2018.8.4.5>
- Braun, C., & Höchle, K. (2006). *Wege aufeinander zu. Persönliche Erfahrungen mit Austauschbegegnungen und Konsequenzen für die zukünftige Förderung von Austausch.* ch Stiftung für eidgenössische Zusammenarbeit.
- Brougère, G., Colin, L., Merckens, H., Kaufmann, K., Nicklas, H., Perrefort, M., & Saupe, V. (2006). *L'immersion dans la culture et la langue de l'autre : Une recherche évaluative du programme Voltaire.* [Rapport de recherche] Office Franco-Allemand pour la Jeunesse. (hal-03634511).

- Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Multilingual Matters. <https://doi.org/10.21832/9781800410251>
- Council of Europe. (2020). *Common European framework of reference for languages: Learning, teaching, assessment – companion volume*. Council of Europe Publishing. <https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4>
- Deardorff, D. K. (2006). Identification and assessment of intercultural competence as a student outcome of internationalization. *Journal of Studies in International Education*, 10(3), 241–266. <https://doi.org/10.1177/1028315306287002>
- Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9–42). Multilingual Matters. <https://doi.org/10.21832/9781847691293>
- Evans, M., & Fisher, L. (2005). Measuring gains in pupils' foreign language competence as a result of participation in a school exchange visit: The case of Y9 pupils at three comprehensive schools in the UK. *Language Teaching Research*, 9(2), 173–192. <https://doi.org/10.1191/1362168805lr162oa>
- Freed, B. F. (1995). *Second language acquisition in a study abroad context* (Vol. 9). John Benjamins Publishing.
- Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation*. E. Arnold.
- Gautschi, L. (2019). *Interkulturelle Kompetenzen im Schüleraustausch. Bedeutung und Entwicklung der interkulturellen Kompetenzen von Deutschschweizer Schülerinnen und Schüler im Austausch mit gleichaltrigen Westschweizern*. PHSG.
- Genkova, P., & Schubert, C. E. (2020). Die Wechselwirkungen zwischen Auslandsaufenthalt, interkultureller Kompetenz und soziokultureller Anpassung für Auslandsentsendungen. *Gruppe. Interaktion. Organisation. Zeitschrift für Angewandte Organisationspsychologie (GIO)*, 51(1), 103–116. <https://doi.org/10.1007/s11612-020-00498-x>
- Gough, D., Oliver, S., & Thomas, J. (2017). *An introduction to systematic reviews*. Sage.
- Gözpinar, H. (2018). School education projects experiences: Fostering lifelong learning as a person and a professional. *Folia Linguistica et Litteraria (Online)*, 21(21), 155–175. <https://folia.ac.me/image/21/Gozpinar.pdf>
- Hammer, M. R. (2005). *The educational results study: Assessment of the impact of the AFS study abroad experience*. A. I. Programs.
- Hammer, M. R., Bennett, M. J., & Wiseman, R. (2003). Measuring intercultural sensitivity: The intercultural development inventory. *International Journal of Intercultural Relations*, 27(4), 421–443. [https://doi.org/10.1016/S0147-1767\(03\)00032-4](https://doi.org/10.1016/S0147-1767(03)00032-4)
- Heinzmann, S., Hilbe, R., Paul, S., & Schallhart, N. (2022). Sprachaustausch in der Grundschule als Motivator für das Erlernen der Landessprachen Französisch und Deutsch? Befunde einer Interventionsstudie in der Schweiz. In A. Bergmann, C. O. Mayer, & J. Plikat (Eds.), *Welche Zielsetzungen sind für Französisch, Spanisch, Russisch & Co. (noch) zeitgemäß?* (pp. 169–191). Peter Lang.
- Heinzmann, S., Künzle, R., Schallhart, N., & Müller, M. (2020). The effect of study abroad on intercultural competence: Results from a longitudinal quasi-experimental study. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 26(1), 187–208. <https://doi.org/10.36366/frontiers.v26i1.366>
- Heinzmann, S., Paul, S., Hilbe, R., Schallhart, N., & Egli Cuenat, M. (2022). Development of productive language skills through language exchange in primary schools in Switzerland – an exploratory intervention study. *European Journal of Applied Linguistics*, 11(1), 98–131 <https://doi.org/10.1515/eujal-2021-0024>
- Heinzmann, S., Schallhart, N., Müller, M., Künzle, R., & Wicki, W. (2014). *Sprachliche Austauschaktivitäten und deren Auswirkungen auf interkulturelle Kompetenzen und Sprachlernmotivation*. Forschungsbericht Nr. 44. Pädagogische Hochschule Luzern.
- Hirai, A. (2018). The effects of study abroad duration and predeparture proficiency on the L2 proficiency of Japanese university students: A meta-analysis approach. *JLTA Journal*, 21(0), 102–123. [https://doi.org/10.20622/jltajournal.21.0\\_102](https://doi.org/10.20622/jltajournal.21.0_102)
- Höchle Meier, K. (2014). *Construction discursive des représentations de stages professionnels dans des entreprises de la région du Rhin supérieur : une étude de cas*. A. Francke.

- Höchle Meier, K., & Yanaprasart, P. (2012). La dimension formative de la mobilité professionnelle : les échanges/stages comme instruments de la gestion de la diversité et/ou du développement du personnel. *Bulletin Suisse de linguistique appliquée*, 95, 35–56.
- Hodel, H.-P. (2004). Sprachaufenthalte beschreiben und evaluieren. *Schweizerische Zeitschrift für Bildungswissenschaften*, 26(3), 397–418.
- Hodel, H.-P., Lüdi, G., & Saudan, V. (2006). *Sprachaufenthalte : Perspektiven und Untersuchungen im außerschulischen Sprachlernen*. A. Francke.
- Hoffstaedter, P., Kohn, K., Jauregi, K., Morató, T., Van den Bergh, H., Melchor-Couto, S., Pennock-Speck, B., Clavel-Arroitia, B., Derivry, M., Stratilaki, S., Cribaillet, S., Potolia, A., & Boughnim, A. (2015). *TILA Research Results on Telecollaboration*. European Commission.
- Howard, M. (2019). *Study abroad, second language acquisition and interculturality* (M. Howard, Ed.). Multilingual Matters.
- Howard, M. (2020). *Study abroad and the second language learner: Expectations, experiences and development* (M. Howard, Ed.; 1st ed.). Multilingual Matters.
- Ilg, W., & Dubiski, J. (2015). *Begegnung schafft Perspektiven: Empirische Einblicke in internationale Jugendbegegnungen*.
- Isabelli-García, C., Bown, J., Plews, J. L., & Dewey, D. P. (2018). Language learning and study abroad. *Language Teaching*, 51(4), 439–484. <https://doi.org/10.1017/S026144481800023X>
- Jauregi, K. (2015). Integrating telecollaboration for intercultural language acquisition at secondary education: Lessons learned. In F. Helm, L. Bradley, M. Guarda, & S. Thouesny (Eds.), *Critical CALL-proceedings of the 2015 EUROCALL Conference, Padova, Italy* (pp. 268–273). Research-publishing.net.
- Jauregi, K., & Melchor-Couto, S. (2018). *Successful telecollaboration exchanges in primary and secondary education: What are the challenges?* <https://doi.org/10.14705/rpnet.2018.26.822>
- Kinginger, C. (2009). *Culture and identity in study abroad contexts: After Australia, French without France*. Taylor & Francis.
- Kinginger, C. (2019). Four questions for the next generation of study abroad researchers. In M. Howard (Ed.), *Study abroad, second language acquisition and interculturality* (pp. 263–278). Multilingual Matters. <https://doi.org/10.21832/9781788924153-013>
- Kohn, K., & Hoffstaedter, P. (2017). Learner agency and non-native speaker identity in pedagogical lingua franca conversations: Insights from intercultural telecollaboration in foreign language education. *Computer Assisted Language Learning*, 30(5), 351–367. <https://doi.org/10.1080/09588221.2017.1304966>
- Köylü, Z. (2023). The ERASMUS sojourn: Does the destination country or pre-departure proficiency impact oral proficiency gains? *The Language Learning Journal*, 51(1), 48–60. <https://doi.org/10.1080/09571736.2021.1930112>
- Kray, J., & Schaefer, S. (2018). Mittlere und späte Kindheit (6-11 Jahre). In W. Schneider & U. Lindenberger (Hrsg.), *Entwicklungspsychologie*, 5 (pp. 215–238). Beltz.
- Krok, I., Bruhns, K., & Rink, B. (2010). *Interkulturelle Kompetenz durch internationale Kinderbegegnung: Hauptstudie: die Kinder im Mittelpunkt*. Deutsches Jugendinstitut.
- Krüger-Potratz, M., & Wagner, B. (2018). *Deutsch-französischer Grundschulaustausch – informelles und interkulturelles Lernen : Eine videoethnographische Studie*. Waxmann.
- Llanes, Á. (2011). The many faces of study abroad: An update on the research on L2 gains emerged during a study abroad experience. *International Journal of Multilingualism*, 8(3), 189–215. <https://doi.org/10.1080/14790718.2010.550297>
- Llanes, Á. (2012). The short- and long-term effects of a short study abroad experience: The case of children. *System*, 40(2), 179–190. <https://doi.org/10.1016/j.system.2012.05.003>
- Llanes, Á. (2016). The influence of a short study abroad experience on perceived foreign accent: An exploratory study beyond the immediate effects. *Study Abroad Research in Second Language Acquisition and International Education*, 1(1), 88–106. <https://doi.org/10.1075/sar.1.1.04lla>
- Llanes, Á., & Munoz, C. (2009). A short stay abroad: Does it make a difference? *System (Linköping)*, 37(3), 353–365. <https://doi.org/10.1016/j.system.2009.03.001>
- Llanes, Á., & Munoz, C. (2013). Age effects in a study abroad context: Children and adults studying abroad and at home. *Language Learning*, 63(1), 63–90. <https://doi.org/10.1177/1474022213483716>

- Llanes, À., & Serrano, R. (2017). The effectiveness of classroom instruction 'at home' versus study abroad for learners of English as a foreign language attending primary school, secondary school and university. *The Language Learning Journal*, 45(4), 434–446. <https://doi.org/10.1080/09571736.2014.923033>
- Llanes, À., Tragant, E., & Serrano, R. (2018). Examining the role of learning context and individual differences in gains in L2 writing performance: The case of teenagers on an intensive study-abroad programme. *The Language Learning Journal*, 46(2), 201–216. <https://doi.org/10.1080/09571736.2015.1020332>
- Mahatmya, D., Lohman, B. J., Matjasko, J. L., & Farb, A. F. (2012). Engagement across developmental periods. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of research on student engagement* (pp. 45–63). Springer Science + Business Media. [https://doi.org/10.1007/978-1-4614-2018-7\\_3](https://doi.org/10.1007/978-1-4614-2018-7_3)
- Mapes, J. (2020). *Development of intercultural competence and cultural empathy in high school students participating in a short-term study abroad exchange*. CORE.
- Maréchal, D. (2007). Quand des élèves flamands et francophones se rencontrent: les stéréotypes communautaires en Belgique à l'épreuve de la pédagogie interculturelle. *Le langage et l'homme*. - Louvain-la-Neuve : EME Éditions, 42(2), 5–12.
- Massler, U. (2008). Ausdruck, Analyse und Förderung der schriftlichen Kommunikations- und Kooperationsfähigkeit in E-Mail-Projekten der Sekundarstufe I. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 13(1), 20S.
- Meyers, C., Mayerl, M., & Fennes, H. (2020). *Exploring inclusion in Erasmus+ youth in action. Effects of social inequalities on learning outcomes*. G. a. E. S. Institute.
- Mitchell, R., Tracy-Ventura, N., & McManus, K. (2017). *Anglophone students abroad: Identity, social relationships and language learning*. Routledge. <https://doi.org/10.4324/9781315194851>
- Mitchell, R., & Tyne, H. (2021). *Language, mobility and study abroad in the contemporary European context* (R. Mitchell & H. Tyne, Eds.). Routledge.
- Montero, L., Serrano, R., & Llanes, À. (2017). The influence of learning context and age on the use of L2 communication strategies.
- Muñoz, C., & Llanes, À. (2014). Study abroad and changes in degree of foreign accent in children and adults. *The Modern Language Journal*, 98(1), 432–449. <https://doi.org/10.1111/j.1540-4781.2014.12059.x>
- Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50(3), 417–528. <https://doi.org/10.1111/0023-8333.00136>
- Ogay, T. (2000). *De la compétence à la dynamique interculturelles : des théories de la communication interculturelle à l'épreuve d'un échange de jeunes entre Suisse romande et alémanique*. Lang.
- Paradowski, M. B., Cierpich-Kozziel, A., Chen, C., & Ochab, J. K. (2022). How output outweighs input and interlocutors matter for study-abroad SLA: Computational social network analysis of learner interactions. *The Modern Language Journal*, 106(4), 694–725. <https://doi.org/10.1111/modl.12811>
- Paradowski, M. B., Jarynowski, A., Czopek, K., & Jelińska, M. (2021). Peer interactions and second language learning: The contributions of social network analysis in study abroad versus at-home environments. In *Language, mobility and study abroad in the contemporary European context* (Vol. 5, pp. 99–116). Taylor & Francis. <https://doi.org/10.4324/9781003087953-10>
- Paul-Frischknecht, S., & Aguirre, M. L. (2014). *Veränderung der Sprachlernmotivation durch Sprach Austausch in der Primarschule: eine Interventionsstudie*. éditeur non identifié. <https://books.google.ch/books?id=e-LZuQEACAAJ>
- Pelz, M. (1999). *Sprachbegegnung und Begegnungssprache : zur experimentellen Überprüfung des Programms "Lerne die Sprache des Nachbarn"*. Peter Lang.
- Pérez-Vidal, C. (Ed.). (2014). *Language acquisition in study abroad and formal instruction contexts*. John Benjamins. <https://doi.org/10.1075/aals.13>
- Piipponen, O., & Karlsson, L. (2019). Children encountering each other through storytelling: Promoting intercultural learning in schools. *The Journal of Educational Research* (Washington, D.C.), 112(5), 590–603. <https://doi.org/10.1080/00220671.2019.1614514>
- Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912. <https://doi.org/10.1111/lang.12079>



- Preiser, G. (2015). Overcoming barriers: Engaging younger students in an online intercultural exchange. *Intercultural Education*, 26(5), 361–376. <https://doi.org/10.1080/14675986.2015.1091238>
- Rosenthal, R. (1979). The “file drawer problem” and tolerance for null results. *Psychological Bulletin*, 86, 638–641.
- Roy, A., Newman, A., Ellenberger, T., & Pyman, A. (2019). Outcomes of international student mobility programs: A systematic review and agenda for future research. *Studies in Higher Education*, 44(9), 1630–1644. <https://doi.org/10.1080/03075079.2018.1458222>
- Ruben, B. D. (1976). Assessing communication competency for intercultural adaptation. *Group & Organization Studies*, 1(3), 334–354. <https://doi.org/10.1177/105960117600100308>
- Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *The American Psychologist*, 55(1), 68–78. <https://doi.org/10.1037/0003-066X.55.1.68>
- Salzmann, S. (2022). *Échange de classes dans l'école primaire située près de la frontière linguistique dans le canton du Valais*. Université de Fribourg (CH).
- Sauer, L., & Ellis, R. (2019). The social lives of adolescent study abroad learners and their L2 development. *The Modern Language Journal*, 103(4), 739–762. <https://doi.org/10.1111/modl.12589>
- Schenker, T. (2012). *The effects of a virtual exchange on language skills and intercultural competence*. ERIC. <https://doi.org/10.25335/M5VF1P>
- Schmid, R. (2008). *Motivationssteigerung für die Zweitsprache durch Sprachtausch*. Pädagogische Hochschule Wallis. <https://books.google.ch/books?id=XcpRYAAACAAJ>
- Schulze, S. (2010). Zwischen Spaßreise und interkultureller Begegnung: Der Einfluss eines schulischen Aufenthaltes in Großbritannien auf das Landesbild der Teilnehmer. *Zeitschrift für Interkulturellen Fremdsprachenunterricht*, 15(1), 98–115.
- Sell, F., Renkowitz, K., Sickinger, P., & Schneider, K. P. (2019). Measuring pragmatic competence on the functional and lexical level: The development of German high-school students' requests during a stay abroad in Canada. *Journal of Pragmatics*, 146, 106–120. <https://doi.org/10.1016/j.pragma.2018.11.005>
- Serrano, R., & Llanes, À. (2015). An explanatory study of the role of age and language learning aptitude in a short stay abroad. *Vigo International Journal of Applied Linguistics*, (12), 107–126.
- Serrano, R., Llanes, À., & Tragant, E. (2016). Examining L2 development in two short-term intensive programs: Study abroad vs. “at home”. *System*, 57, 43–54. <https://doi.org/10.1016/j.system.2016.01.003>
- Thomas, A., Chang, C., & Abt, H. (2007). *Erlebnisse, die verändern : Langzeitwirkungen der Teilnahme an internationalen Jugendbegegnungen*. Vandenhoeck & Ruprecht.
- Tseng, W.-T., Liu, Y.-T., Hsu, Y.-T., & Chu, H.-C. (2024). Revisiting the effectiveness of study abroad language programs: A multi-level meta-analysis. *Language Teaching Research*, 28(1). <https://doi.org/10.1177/1362168820988423>
- Tullock, B., & Ortega, L. (2017). Fluency and multilingualism in study abroad: Lessons from a scoping review. *System*, 71, 7–21. <https://doi.org/10.1016/j.system.2017.09.019>
- Vande Berg, M., Connor-Linton, J., & Paige, R. M. (2009). The Georgetown consortium project: Interventions for student learning abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 18(1), 1–75. <https://doi.org/10.36366/frontiers.v18i1.251>
- Vasbø, K. B. (2013). Intercultural learning across contexts. *International Electronic Journal of Elementary Education*, 6(1), 83–94.
- Vatter, C., Lüsebrink, H.-J. R., & Mohr, J. (2011). *Interkulturelles Lernen im interregionalen Schüleraustausch zwischen Deutschland und Frankreich : Evaluationsergebnisse und didaktische Materialien des COMENIUS-Regio-Projekts ILIS*. Röhrig Universitätsverlag.
- Weichbrodt, M. (2014). *Ein Leben lang mobil? : langfristige Schüleraustauschprogramme und die spätere Mobilität der Teilnehmer als Element gesellschaftlicher Transnationalisierung*. Wissenschaftliche Schriften der WWU.
- Weichold, K., & Silbereisen, R. K. (2018). Jugend (10–20 Jahre). In W. Schneider & U. Lindenberger (Hrsg.), *Entwicklungspsychologie*, 5 (pp. 239–264). Beltz.

- Weis, D., Meyers, C., & Willems, H. (2013). *Lernprozesse und Lernerfahrungen in Jugendprojekten. Eine Studie zu Aspekten non-formaler Bildung in den Projekten von "Jugend in Aktion" in Luxemburg*. Integrative Research Unit on Social and Individual Development (INSIDE), Ministère de la Famille et de l'Intégration (MFI), Service National de la Jeunesse (SNJ), Universität Luxemburg (UL). <https://www.jugend-in-luxemburg.lu/lernprozesse-und-lernerfahrungen-in-jugendprojekten>
- Xiao, F. (2015). Adult second language learners' pragmatic development in the study-abroad context: A review. *Frontiers: The Interdisciplinary Journal of Study Abroad*, 25(1), 132–149. <https://doi.org/10.36366/frontiers.v25i1.349>
- Xu, Y. (2019). Changes in interlanguage complexity during study abroad: A meta-analysis. *System*, 80, 199–211. <https://doi.org/10.1016/j.system.2018.11.008>
- Yang, J.-S. (2016). The effectiveness of study-abroad on second language learning: A meta-analysis. *The Canadian Modern Language Review*, 72(1), 66–94. <https://doi.org/10.3138/cmlr.2344>
- Zimmer-Gembeck, M. J., & Collins, W. A. (2006). Autonomy development during adolescence. In G. R. Adams & M. D. Berzonsky (Eds.), *Blackwell handbook of adolescence* (pp. 174–204). Blackwell Publishing. <https://doi.org/10.1002/9780470756607.ch9>

## Appendix

**Table A1.** Coding scheme.

Criteria /Codes	Answer categories
<b>(A) Publication</b>	
Authors	free text
Year of publication	free text
Title	free text
Publication Type	Article Book chapter Report Dissertation grey, unpublished literature
Language of publication	German English French
Peer-review	Yes No NA (not specified)
Related Items	This code is used to record further publications of the literature synthesis that refer to the same study/sample. The code makes reference to the title of the related publication.
<b>(B) Study participants</b>	
Age	free text; NA where information is lacking
Gender	free text (percentage); NA where information is lacking
L1 of participants	free text; NA where information is lacking
Educational level	Primary (yes/no) lower secondary (yes/no) upper secondary, general education (yes/no) upper secondary, vocational education (yes/no)
Target language	English (yes/no) German (yes/no) French (yes/no) Italian (yes/no) other (yes/no)
Exchange format	whole class half-class individual rotational school trip / encounter day other, namely NA

(Continued)

**Table A1.** Continued.

Criteria /Codes	Answer categories
Countries / linguistic areas involved	free text Countries/linguistic areas involved in the exchange (e.g. Germany & U.S.A)
Type of contact	face-to-face virtual face-to-face & virtual NA
Duration of programme	free text
Curricular Embedding	Yes No NA free text for description
<b>(D) Research Design and Methodology</b>	
Theoretical Approach	free text
Research Design	free text
Methodology	quantitative qualitative mixed-method
Target Variable Motivation	Yes No free text for summary of relevant passages
Target Variable Intercultural Competence	Yes No free text for summary of relevant passages
Target Variable Language Competence	Yes No free text for summary of relevant passages
Sample	free text description of the sample
Sample Size	free text (number)
Research Questions	free text
Hypotheses	free text
Aims	free text
Data/Instruments	free text
Analysis procedure	free text
<b>(E) Results</b>	
Results	free text
Conclusions	free text
Implications	free text
Limitations	free text
<b>(F) Quality Criteria</b>	
Pretest – Posttest	Yes No NA
Control Group	Yes No NA
Language testing (not self-reports)	Yes No NA
Randomised sample	Yes No NA
Purposeful sampling	Yes No NA
Transparency/Information density	Complete / intersubjectively verifiable some information missing / mostly/partly intersubjectively verifiable deficient / not intersubjectively verifiable

**Table A2.** Overview of included studies.

Reference	Publication type	Exchange format	Exchange settings	Duration	Method	Sample size	Pretest-posttest design	Control group	Educational status	Target language	Target variables
Bachner and Zeutschel (1994)	report	individual	face-to-face	broad spectrum	mixed methods	>1000	no	yes	upper secondary	English, German	language competence & intercultural competence
Bachner and Zeutschel (2009)	article	individual	face-to-face	NA	mixed methods	>1000	no	yes	upper secondary	English, German	intercultural competence
Baenziger (2012)	grey, fugitive literature	whole class	face-to-face & virtual	2–3 weeks	Quantitative	21–50	yes	no	lower secondary	German	motivation
Botturi et al. (2018)	article	whole class	face-to-face & virtual	10 months +	mixed methods	101–200	yes	no	primary	French, German, Italian	language competence, motivation & intercultural competence
Braun and Höchle (2006)	report	various	face-to-face & virtual	NA	mixed methods	>1000	no	no	primary & lower secondary	French, German, Italian	intercultural competence
Brougère et al. (2006)	report	individual	face-to-face	5–6 months	Mixed methods	101–200	NA	no	upper secondary	German, French	Intercultural competence
Evans and Fisher (2005)	article	whole class	face-to-face	2–3 weeks	mixed methods	51–100	yes	yes	primary	French	language competence
Gautschi (2019)	grey, fugitive literature	rotational	face-to-face & virtual	1 week	mixed methods	11–20	yes	no	lower secondary	French	intercultural competence
Genkova and Schubert (2020)	article	individual	face-to-face	10 months +	Quantitative	201–300	no	no	upper secondary	NA	intercultural competence
Gözpinar (2018)	article	individual	face-to-face	2–3 weeks	Qualitative	21–50	no	no	upper secondary	English	intercultural competence
Hammer (2005)	report	individual	face-to-face	10 months +	not defined	>1000	yes	yes	upper secondary	English, various	language competence & intercultural competence

*(Continued)*

**Table A2.** Continued.

Reference	Publication type	Exchange format	Exchange settings	Duration	Method	Sample size	Pretest-posttest design	Control group	Educational status	Target language	Target variables
Heinzmann, Hilbe, et al. (2022)	book chapter	whole class	face-to-face & virtual	1–3 days	Quantitative	301–400	yes	yes	primary	French, German	motivation
Heinzmann et al. (2014)	article	Individual	face-to-face	2–3 weeks	quantitative	401–500	yes	yes	upper secondary	English, French, German	intercultural competence
Heinzmann et al. (2022)	article	whole class	face-to-face & virtual	1–3 days	NA	301–400	yes	yes	primary	French, German	language competence & motivation
	report	whole class	face-to-face & virtual	1–3 days	mixed methods	301–400	yes	yes	primary	French, German	language competence & motivation
Heinzmann et al. (2014)	report	Individual	face-to-face	2–3 weeks	mixed methods	401–500	yes	yes	upper secondary	English, French, German	motivation & intercultural competence
Höchle Meier and Yanaprasart (2012)	article	individual	face-to-face	NA	qualitative	21–50	no	no	upper secondary	French, German	language competence & intercultural competence
Höchle Meier (2014)	dissertation	individual	face-to-face	NA	qualitative	21–50	no	no	upper secondary	French, German	language competence
Hodel (2004)	article	whole class	face-to-face	1–3 months	mixed methods	21–50	yes	no	upper secondary	French	language competence
Hodel et al. (2006)	dissertation	whole class	face-to-face	1–3 months	mixed methods	51–100	yes	no	upper secondary	French	language competence
Ilg and Dubiski (2015)	report	individual	face-to-face	1–3 weeks	quantitative	>1000	no	no	lower & upper secondary	French, German, Polish, various	language competence, motivation & intercultural competence

*(Continued)*

**Table A2.** Continued.

Reference	Publication type	Exchange format	Exchange settings	Duration	Method	Sample size	Pretest-posttest design	Control group	Educational status	Target language	Target variables
Jauregi (2015)	article	various	virtual	NA	not defined	501–1000	NA	no	lower secondary	French, Spanish	intercultural competence
Jauregi and Melchor-Couto (2018)	article	whole class	virtual	1–3 months	mixed methods	21–50	no	no	primary & lower secondary	English, Spanish	motivation
Kohn and Hoffstaedter (2017)	article	individual	virtual	NA	mixed methods	11–20	no	no	lower secondary	English, German	language competence & intercultural competence
Krok et al. (2010)	report	individual	face-to-face	2–3 weeks	quantitative	51–100	yes	no	primary	various	intercultural competence
Krüger-Potratz and Wagner (2018)	report	whole class	face-to-face	1 week	qualitative	NA	no	no	primary	German	language competence & intercultural competence
Llanes and Munoz (2009)	article	individual	face-to-face	3–4 weeks	quantitative	21–50	yes	no	lower & upper secondary	English	language competence
Llanes and Munoz (2013)	article	individual	face-to-face	1–3 months	quantitative	101–200	yes	yes	primary	English	language competence
Llanes and Serrano (2017)	article	individual	face-to-face	1–3 months	quantitative	101–200	yes	yes	primary & lower secondary	English	language competence
Llanes (2012)	article	individual	face-to-face	1–3 months	quantitative	11–20	yes	yes	primary	English	language competence
Llanes (2016)	article	whole class	face-to-face	1–3 months	quantitative	11–20	yes	yes	primary	English	language competence
Llanes et al. (2018)	article	individual	face-to-face	2–3 weeks	quantitative	51–100	yes	no	lower secondary	English	language competence
Mapes (2020)	dissertation	various	face-to-face	2–3 weeks	mixed methods	21–50	yes	no	upper secondary	English, Danish	intercultural competence
Maréchal (2007)	article	whole class	face-to-face	1–3 days	qualitative	51–100	yes	no	upper secondary	French, Flemish	intercultural competence
Massler (2008)	article	whole class	virtual	4–9 months	qualitative	51–100	yes	no	lower secondary	English	language competence &

(Continued)

**Table A2.** Continued.

Reference	Publication type	Exchange format	Exchange settings	Duration	Method	Sample size	Pretest-posttest design	Control group	Educational status	Target language	Target variables
Meyers et al. (2020)	report	various	face-to-face	broad spectrum	quantitative	>1000	no	no	upper secondary	various	intercultural competence
Montero et al. (2017)	article	whole class	face-to-face	1–3 months	mixed methods	51–100	yes	yes	primary	English	intercultural competence
Muñoz and Llanes (2014)	article	whole class	face-to-face	1–3 months	quantitative	51–100	yes	yes	primary	English	language competence
Ogay (2000)	dissertation	individual	face-to-face	1–3 months	mixed methods	≤10	yes	no	upper secondary	German	intercultural competence
Paul-Frischknecht and Aguirre (2014)	grey, fugitive literature	whole class	face-to-face	2–3 weeks	quantitative	21–50	yes	yes	primary	French, German	motivation
Pelz (1999)	book	whole class	face-to-face	1–3 days	mixed methods	501–1000	no	no	primary	French, German	language competence, motivation & intercultural competence
Piipponen and Karlsson (2019)	article	whole class	virtual	1–3 months	qualitative	51–100	no	no	primary	English, German	intercultural competence
Preiser (2015)	article	whole class	virtual	4–9 months	qualitative	21–50	no	no	lower secondary	English, German	intercultural competence
Salzmann (2022)	grey, fugitive literature	whole class	face-to-face & virtual	1–3 months	quantitative	201–300	yes	yes	primary	French, German	motivation
Sauer and Ellis (2019)	article	individual	face-to-face	4–9 months	mixed methods	≤10	no	no	upper secondary	English	language competence
Schenker (2012)	dissertation	whole class	virtual	1–3 months	mixed methods	21–50	yes	no	upper secondary	English, German	language competence & intercultural competence
Schmid (2008)	grey, fugitive literature	half-class	face-to-face	1 week	quantitative	11–20	yes	no	lower secondary	French, German	motivation
Schulze (2010)	article	whole class	face-to-face	1 week	mixed methods	51–100	yes	no	lower secondary	English	intercultural competence

(Continued)

**Table A2.** Continued.

Reference	Publication type	Exchange format	Exchange settings	Duration	Method	Sample size	Pretest-posttest design	Control group	Educational status	Target language	Target variables
Sell et al. (2019)	article	individual	face-to-face	10 months +	mixed methods	101–200	yes	yes	lower secondary	English	language competence
Serrano and Llanes (2015)	article	individual	face-to-face	2–3 weeks	quantitative	21–50	yes	no	lower secondary	English	language competence
Serrano et al. (2016)	article	individual	face-to-face	2–3 weeks	quantitative	101–200	yes	yes	lower secondary	English	language competence
Thomas et al. (2007)	book	various	face-to-face	2–3 weeks	mixed methods	501–1000	no	no	primary, lower & upper secondary	English, French, Polish	intercultural competence
TILA; Hoffstaedter et al. (2015)	report	whole class	virtual	4 sessions spread across 1–3 months	mixed methods	201–300	no	no	lower secondary	Dutch, English, French, German, Spanish	intercultural competence
Vasbø (2013)	article	NA	face-to-face	2–3 weeks	qualitative	≤10	no	no	lower secondary	Portuguese, various	intercultural competence
Vatter et al. (2011)	book	individual	face-to-face	2–3 weeks	quantitative	301–400	no	no	lower secondary	French, German	language competence & intercultural competence
Weichbrodt (2014)	dissertation	various	face-to-face	10 months +	mixed methods	>1000	NA	no	lower & upper secondary	English, French, Italian, various	intercultural competence
Weis et al. (2013)	report	various	face-to-face & virtual	broad spectrum	mixed methods	101–200	no	no	upper secondary	NA	language competence & intercultural competence