
What Makes Study Abroad Transformative? Comparing Linguistic and Cultural Contacts and Learning Outcomes in Virtual vs In-Person Contexts

DAN E. DAVIDSON

American Councils for International Education

Email: Davidson@americancouncils.org

NADRA GARAS

American Councils for International Education

Email: ngaras@americancouncils.org

To what extent can the transformative power and language learning affordances provided by the study abroad experience be virtualized? The large-scale shift away from on-site study abroad to online learning in 2020-2021, caused by the COVID pandemic, has made it possible to compare data for in-person immersion learning versus digitally mediated forms of direct instruction and second language (L2) community engagement of otherwise similarly prepared L2 learner cohorts. The present study compares measured proficiency outcomes, as well as other performance-based data of late-adolescent and young adult U.S. participants in a group of federally sponsored overseas intensive immersion study programs for Arabic, Chinese, and Russian operating abroad for the period 2017-2019 and as virtual programs in 2020-2021 (in-person N=1388 and virtual N=770). Program data for early-stage learners, mid-level learners, and advanced learners are analyzed separately by target language, initial proficiency levels, and program durations. As a result of the COVID pandemic, all programs shifted from in-person format to virtual instruction, provided by the existing overseas partner faculties, who adapted core curricula, group instruction, direct enrollment courses and tutorials for use in both synchronous and asynchronous forms of online instruction. Similarly, organization of homestay visits, internship programs and cultural programming was also shifted to virtual format. Given the importance of language contact and L2 interactions, this study provides comparisons of both aggregate and activity-specific L2 time-on-task levels for both cohorts, documenting comparable levels of instructional time available within the two formats (18 hours/week), but less than half the available hours to the virtual student for informal L2 contact and community engagement activities. Comparisons of the speaking proficiency outcomes by language and training levels revealed a relatively consistent gap in mean gain levels across languages, with participants in the virtual groups attaining on average one proficiency sub-level lower in speaking gains than their respective face-to-face counterparts. Comparisons of reading, listening, and writing outcomes, however, showed smaller differences, sometimes none at all. Intercultural Development Index (IDI) and other data related to participants' cultural competencies, including cultural referencing, self-presentation (Identity Competence), socio-pragmatic strategy selection, lexical and collocational choices, and other markers were observed generally to fall short of the standards set in previous years by students enrolled in the in-person versions of the same programs. Finally, the study takes note of curricular and technological interventions introduced during the COVID period which have been recommended for adoption in the new cycle of in-person programs in 2022-2023 and beyond.

INTRODUCTION

Academic mobility has been central to the concept of the western university from its earliest days, viewed as essential for broadening intellectual perspectives and ensuring the well-rounded social and moral growth of the young scholar. With the massification of higher education of the past century, international student mobility has also expanded greatly, with 5.6 million students worldwide as of 2018 pursuing all or part of their university studies beyond the borders of their home country (OECD, 2020).

As Gordon Allport (1954) theorized in his foundational work and presentation of the *contact hypothesis*, extended intergroup contact *under favorable conditions* can reduce if not eliminate cultural stereotyping and prejudice, while enabling deeper understanding of one's own culture and oneself (Allport, 1954, p. 487). That position has been interrogated and supported by successive generations of researchers in social psychology and widely credited with influencing major anti-discrimination and desegregation policies in the U.S., Northern Ireland, South Africa, and other nations (Cohen, 1975; Dovidio et al., 2001). Of particular interest to the present discussion are the findings of Pettigrew and Tropp (2006), whose meta-analysis of 515 separate studies representing responses from 250,493 individuals across 38 countries confirmed the positive relationship between intergroup contact and lower levels of prejudice, predicted by Allport (1954): 94% of the samples showed an inverse relationship between intergroup contact and prejudice and intergroup stereotyping. Moreover, the effect showed a sharp rise in statistical power in the case of the more rigorously conducted experimental studies included in the analysis.

Allport (1954) specifically addresses cross-national and inter-ethnic stereotyping in his study, analyzing their underlying cognitive and motivational components as natural results of human evolution and subject to remediation (Allport, 1954, pp. 45-47). His writing has provided subsequent conceptual support for the international education field for understanding the powerful effects that study abroad can produce, especially on late adolescents and young adults of all backgrounds (Davidson & Lehman, 2005; Kinginger, 2008; Salisbury, 2011). Brewer and Gaertner (2001) refine the concept of intergroup bias in terms of in-group favoritism rather than out-group derogation, describing how extended contacts across groups can stimulate cognitive re-categorization processes affecting changes to an individual's attitude, signaled by a representational shift away from *us/them* in the direction of *we*.

Re-categorization of the *other* is a process that invariably challenges established views of self, of one's home group or culture, of the way things work, of what is acceptable behavior and what is not, experiences that may, and often do, force the sojourner out of their comfort zone and into a less familiar space accompanied by varying degrees of cognitive dissonance and emotional dis-orientation. That process is central to the concept of transformative learning, as developed by Mezirow and Dirkx over the past three decades and now frequently associated with study abroad, second language and intercultural learning (Mezirow, 2012; Dirkx et al., 2006; Leaver et al., 2021).

“Transformative learning refers to the process by which we transform our taken-for-granted frames of reference (meaning perspectives, habits of mind, mind-sets) to make them more inclusive, discriminating, open, emotionally capable of change, and reflective so that they may generate beliefs and opinions that will prove more true or justified to guide action.” (Mezirow, 2012, p. 76)

Within the study abroad context, language immersion students undergo a process of continued socialization into the host country culture, during which they may seek to manage or maintain their first language and culture (L1/C1) identities, while learning to interact with people of different or divergent affective, cognitive, and behavioral orientations to the world. Long before the introduction of the concept of *Transformative Learning* in 1978 by Mezirow (1978), study abroad alumni have often referred to their time abroad as *transformative* and *life-changing* in the popular literature. Viewed *in situ* and in real time, however, the transformative process required for a sojourner to function effectively and appropriately in a different academic, residential, and workplace environment abroad can be challenging, unrelenting, and by no means assured of success, as *Jennifer*, a year-long study abroad participant, points out in her weekly journal entry:

“My domestic program made me aware of the cultural issues that I would run into, but I went into the [year-long overseas] program with the mentality that these would just be one-off issues. I was not prepared for the pervasiveness of these issues and the fact that these were scenarios that I would be faced with almost all day, every day. Taken individually, these situations weren’t all that problematic. It was just the sheer overload of them combined that was often overwhelming.”

Transformative language learning positions these disorienting experiences at the center of the educational enterprise, supporting the learner’s reflections and reconsideration of familiar, established frames of reference, while encouraging students to remain attentive for vestiges of ethnocentrism which may be coloring their own apprehension and judgments of the practices, viewpoints, and cultural narratives they encounter in the host country.

In-person study abroad learning that includes well-monitored opportunities for students to reflect on their own language and cultural learning can be truly transformative. For example, over the past 20 years, students in the year-long Russian Flagship program complete weekly online journals (see Language Utilization Reports LUR section below) which include prompts to assist them in reflecting on their linguistic and cultural learning (challenges as well as successes). All journals are monitored and responded to by their academic director. In addition, monthly in-person *cultural roundtables*, which include moderated group discussions, role plays, and case studies based on actual but anonymized intercultural scenarios of past participants, provide students with a safe, unthreatening, off-site English-language venue for sharing views and experiences as well as possible strategies for navigating cultural dilemmas they have encountered during the preceding week (Davidson, et al., 2021).

Critical reflection and strategic goal setting are principal components of transformative language learning. Participants are encouraged to confront dis-orienting dilemmas first through serious introspection and examination of possible internal sources of the cognitive and emotional distress they are experiencing as a result of a behavior they have encountered (premise reflection). For example, a student reflects on her experience studying in Kazakhstan: “Why does it trouble me when locals keep asking about my nationality, or when I plan to get married? At home, we would consider these questions inappropriate in similar circumstances.” Students then reflect critically on the range of possible intentions of those who pose such questions. Could these inquiries, in fact, be well-wishing, rather than an attempt to label or stereotype them in some way? Within a host country culture which places the highest value on family and children, these questions, which can be uncomfortable or excessively personal for many Americans, are, in fact, entirely acceptable, indeed well-intentioned, and fairly standard forms of local speech behavior. They are intended to convey hope for the brightest possible future for a young person, or to share their own pride of cultural identity they experience as citizens of a multinational, multiethnic young

state like Kazakhstan. Critical reflection and discussion with peers and cultural mentors can aid students in reframing issues in a way that broadens their understanding of the history, value systems and belief structures of the host country and position them better for responding appropriately and with less discomfort when similar cultural issues arise in the future.

Like language proficiency, intercultural competencies appear to develop along a continuum and over a lifetime of learning and intercultural experience. (Bennett, 1993; Byram, 1997). To be sure, intercultural growth is an iterative process, non-linear and cumulative in nature, characterized by Mezirow (1994, pp. 229-230) as “a set of progressive transformations in related meaning schemes,” alternating with occasional setbacks. Longer sojourns abroad combined with advanced levels of linguistic and cultural knowledge, therefore, tend to benefit intercultural development by multiplying numbers and ranges of opportunities for personal and social interactions in the target language and allowing the learner to strengthen their skills at anticipating, recognizing, and managing potentially challenging situations that may arise. Data from the Language Flagship programs demonstrate that sojourners with advanced to superior levels of language proficiency and commensurate access to cultural mentoring and reflection have the potential and often do attain developmental orientation skills consistent with the *acceptance* level on the widely used Intercultural Development Index (IDI) scale with significant growth in the areas of behavioral recoding and cognitive frame-shifting abilities (Davidson et al., 2016; Davidson et al., 2021; Hammer, 2012).

The question posed by the present study and this special issue concerns virtual immersion. To what extent can the documented experiences of linguistic, cultural, and intercultural growth gained through immersion in an overseas linguistic and cultural environment be replicated with the help of digital media and online instruction by similar instructional teams and teaching materials for comparable program durations (e.g., eight-week summer or two academic semesters of study)? Stated in another way, to what extent can the transformative experience of study abroad be replicated in a virtual setting and through digital means?

IDENTITY COMPETENCE

Pellegrino Aveni (2005) observed and studied the linguistic and cultural behaviors of U.S. students participating in an intensive overseas language immersion program over the course of a year, documenting her subjects' affective responses to learner anxiety arising from the need to protect the sense of self, while also creating a self-identity in the target language. Her L2 learner case studies illustrate the tensions as well as the successful strategies used by advanced learners for balancing the social-environmental factors and the learner-internal factors, as the students develop increasing levels of comfort with the language, while learning to maintain a *stable identity* in social situations of increasing complexity over the course of a year.

Successful L2 students learn to perform their identity and roles within the surrounding cultural context in ways that take into consideration local value systems and existing social hierarchies. Construction and management of identity roles may be challenged by the stereotypical identities often ascribed to Americans abroad and reflective of the L2 learner's early-stage dependencies on external cues from local interlocutors for security and approval of their speech production (Quan, 2019). Self-presentation, like other L2 speech activities, can succumb to more general issues of L2 speech anxiety, depending on the gravity of the stakes, risk to saving face, should the speech act fail, and the student's overall level of comfort with the language (Ehrman & Dornyei, 1998; Kinginger, 2008). The study abroad environment can contribute greatly to reducing the anxieties associated with self-presentation in a different

language and culture by providing the context and the motivation for students to maintain a *stable identity* in social situations of increasing complexity as their comfort with the L2 increases (Freed et al., 2004; Pellegrino Aveni, 2005).

LANGUAGE GAINS

In the study abroad context, intercultural growth has been shown to correlate positively with target language proficiency; the higher the level of proficiency, the more nuanced and comprehensive the learner's ability to notice and correctly interpret embedded cultural signals, and, thus, the more likely the student is to be successful in negotiating meaning and navigating cultural differences in an effective and appropriate way (Byrnes, 2013; Davidson et al., 2016, 2021).

Well-structured overseas language immersion programming has been demonstrated frequently over past decades to be a rich and highly motivating environment for language acquisition for young adult learners of all background and proficiency levels (Davidson et al., 2021, 2015; Davidson & Lekic, 2010; Freed et al., 2004; Watson et al., 2013). Many teachers and specialists today consider a substantial period of immersion in the target language and culture essential for ensuring the requisite levels of linguistic, cultural, and regional knowledge for those planning to work or pursue graduate-level study in a field or area, where professional proficiency is required (Commission on Language Learning, 2017).

STUDY POPULATION

The current study presents the measured proficiency outcomes of participants in three federally sponsored overseas immersion study programs for U.S. students of Arabic, Chinese, and Russian, programs administered by the American Councils for International Education in cooperation with domestic and overseas institutional partners. Programs for the periods of 2017-2019 were conducted on site and in-person, while those operating during the 2020-2021 pandemic year were conducted virtually. Included in the study are participants in the National Security Education Program for Youth (NSLI-Y), the Critical Language Scholarship (CLS) programs, and the Language Flagship Programs. NSLI-Y and CLS accept early-stage and intermediate-level learners, and are sponsored by the U.S. Department of State, Bureau of Educational and Cultural Affairs. The Flagship Capstone programs accept advanced-level learners and are sponsored by the Department of Defense, Defense Language and National Security Education Office (See Table 1). NSLI-Y students comprise two categories of learners in the present study: summer (seven week) and academic year (nine month) students; CLS students complete eight-week summer programs in the target language and comprise approximately 70% undergraduate and 30% graduate students with mean ages of 21.1 and 23.3 years, respectively. Flagship Capstone students are undergraduates, primarily graduating seniors, and study overseas for nine to 12 months.

Table 1
Study Population

<i>2017-2021 Program Demographics for Arabic, Chinese and Russian</i>		
	<i>Number of Participants</i>	<i>Average Age</i>
<i>NSLI-Y</i>	972	17.1
<i>CLS</i>	928	20.8
<i>Flagship</i>	258	21.4

The NSLI-Y and CLS programs are fully funded and recruit and select participants on a competitive basis. NSLI-Y students apply to the program from high school, participating in the summer programs prior to their junior or senior year, and in the year-long programs typically as high school seniors or as *gap year* students. Participants in the programs are actively recruited and come from a broad and diverse range of public and private institutions, community colleges, and minority serving institutions, including Historically Black Colleges and Universities (HBCU), Historically Hispanic Institutions (HSI's), and Tribal Colleges (See Table 2 for an example of the race and ethnic diversity of the programs). Flagship students apply from one of 31 domestic Language Flagship university programs throughout the U.S. and must be able to demonstrate the Interagency Language Roundtable (ILR 2) proficiency in speaking and at least one other skill at the time of application and test not lower than 1+ in any skill. The full scholarship provided by the Boren Award for Flagship students carries a one-year government service obligation, while the standard Flagship scholarship awards provide substantial financial assistance to Flagship students for both summer and year-long language study. Approximately 15% of all students across programs are STEM majors, and none of the programs require that the students major in a foreign language in order to participate, although approximately one third of all participants do indeed go on to major or double-major in the target language.

Table 2
Example of Diversity in the 2019 Study Abroad Programs

<i>Breakdown of Race-Ethnicity of Applicants, Semi-Finalists and Participants</i>						
Race/Ethnicity	Applicants		Semifinalists		Participants	
Asian	700	11.40%	179	13.16%	70	10.39%
I decline to state	156	2.54%	38	2.79%	21	3.12%
Other	37	0.60%	13	0.96%	6	0.89%
Native American or Alaska Native	62	1.01%	12	0.88%	8	1.19%
Black or African American	572	9.31%	87	6.40%	52	7.72%
Native Hawaiian or Pacific Islander	34	0.55%	5	0.37%	3	0.45%
Latinx or Hispanic	640	10.42%	101	7.43%	44	6.53%
Middle Eastern or North African	203	3.31%	51	3.75%	21	3.12%
Multi-Racial	468	7.62%	111	8.16%	60	8.90%
South Asian	190	3.09%	69	5.07%	39	5.79%
White	3079	50.14%	694	51.03%	350	51.93%
Total	6141	100.00%	1360	100.00%	674	100.00%

All three of the overseas immersion programs in the present study make use of a generally agreed upon set of curricular and co-curricular components that are adapted by local teachers and program directors to accommodate local conditions and educational resources, as necessary (See Figure 1).

Figure 1
Curricular and co-curricular components

Basic Structural Components of the Overseas Immersion Programs	
<u>Academic Components</u>	
•	Intensive language training in small (3-4 persons) and full groups
•	Direct enrollment courses: 2 – 3 hours (as L2 proficiency allows)
•	Integrated cultural program (incl 3 overnight, accompanied by teachers & language partners)
•	One-on-one tutorials (curricular based, instructional) – 4 hours/week
•	Peer language partners (2 hours per week) – opt, but virtually all participate.
•	Strong Academic Advising Component (Faculty, AD, RD, and individual advisories)
<u>Co-Curricular Components</u>	
•	Integrated homestay program
•	Professional internships
•	Volunteer projects, community service
•	Discussion groups with local students (2-3 times/sem)
•	Language Utilization Reports (time-place, speech, culture prompts)

The emphasis across all programs is on student-centered learning and the importance of ensuring adequate balance across the program of comprehensible input and structured feedback from professional teachers and tutors, as well as maximum access to opportunities for academic and community engagement for all learners. All programs are hosted by overseas universities or centers and supervised by a full-time American resident director, as well as a U.S. based academic advisor. Overseas programs are regularly visited by U.S. faculty members from sending institutions in the U.S. who may also participate in regular on-site and stateside faculty professional development activities.

TO WHAT EXTENT CAN THE SA EXPERIENCE BE VIRTUALIZED?

“Having a distance learning study abroad is kind of like having all of the responsibility and difficulty of a rigorous program, but none of the actual fun parts and cultural experiences of studying abroad.”
 (Virtual student, Sept. 2020)

In the section that follows, comparative data from the three large-scale federally sponsored overseas language training programs will be presented, focusing on the proficiency test results of the NSLI-Y, CLS, and Flagship programs for the pre-COVID years of 2017-2019 versus the corresponding results for the 2020-2021 COVID year, during which all instruction and language support activities were conducted online.

To clarify the differences in the learners’ experience of the virtual and the in-person program venue, the study will briefly identify some of the special practices developed during the COVID year to maximize student learning as well as increase opportunities for student contacts and speech interactions with the remote host country community. Practices judged by the program that have proven effective and worthy of consideration for adoption in the

future will be noted. The study will then take a more detailed look at a program-specific sample of time-on-task data for the two cohorts to provide additional objective information on the differing student experiences of the two intensive programs by reporting how the students actually allocated their time and activities in the virtual program versus the in-person program.

The section will conclude by offering preliminary findings based on the analysis of the speech production of virtual participants and aspects of their L2 performances that appear to distinguish them from the presentational speech of past alumni of those same programs, concluding with the presentation of preliminary data on intercultural development gains, and a brief list of *lessons learned* for the future as a result of the virtual study abroad program year.

VIRTUALIZING STUDY ABROAD L2 PROGRAMS

For the duration of the COVID pandemic, all three programs shifted from in-person to virtual instruction, cooperating closely with their established overseas partner faculties in adapting core curricula, group instruction and tutorials for use in both synchronous and asynchronous forms of online instruction while organizing virtual homestay visits, internships, peer meetings, excursions, and other cultural programming. Best practices in the design of virtual learning environments were taken into consideration by the programs: placing reasonable limitations on the total number of synchronous class hours per week students would be required to attend in order to avoid screen fatigue; recognizing the environmental complexity that an intensive online program can present for students, most of whom were studying from home; providing recordings of classes to students whose participation was disrupted by health, environmental, or connectivity issues, etc. A particular challenge to virtual study abroad students in the present groups was posed by the large time-zone differences over which the virtual instruction was spread, which further constricted the number of synchronous hours available to a teacher and student who might be separated by six to 14 hours on the clock.

Among promising online curricular developments, students and faculty members indicated satisfaction with the results of a number of technological innovations required by the COVID environment, which they recommend be continued in the future:

- *Flipped lectures* and student response presentations.
- *Introduction of Canvas and similar learning management systems* were an innovation for many overseas institutions, but soon deemed to be indispensable in 2020-21; many programs have continued their use in 2021-22.
- *Oral commentaries*: Virtual students were requested to submit oral commentaries related to their written assignments; a practice that was observed to have a positive effect on their speaking production.
- *Feedback on Writing Assignments*: Teachers and students took note of the more timely and focused quality of feedback options for written homework offered by *Google Drive* and indicated an interest in continuing the practice in the future. As a result of the more rigorous approach that emerged, program organizers elected to revise existing rubrics and convene workshops on assessing writing production in the online format.
- *Zoom & Break-Out Rooms*: Counter-intuitively, online classes contributed to a more student-centered classroom environment in a number of the overseas centers by making it easier for teachers to spend more time in the background, as student

groups divided into the breakout room option on Zoom, increasing student-to-student interaction and student-moderated discussions.

Responding to the curricular practices noted above, student evaluations tended to run positive:

“There have been many advantages to the virtual program. For me, the main advantage is having digital copies of lecture material and homework. This is not only better for the environment, but also has made it a lot easier to store, manage, search, and review content.”

“I like that students were expected to prepare and research for presentations and everyone had an assigned role. I also REALLY liked watching the lectures, as this basically doubled our class time, which was amazing, considering how little practice we are actually able to have since we are not immersed.”

“I wish all the classes had prerecorded lectures or the type of homework where we are all assigned a role and have to complete it in class. It is so efficient and it makes sure no one can get distracted during class.”

Apart from online course work and tutorials with overseas faculty, the virtual programs sought to capitalize on a wide range of student interests to maximize opportunities for informal interactions, broadening their vocabularies, cultural learning and simply connecting with other speakers of the language, largely on an optional rather than required format.

- Students met with their homestay family members by WhatsApp or Zoom each week
- Weekly meetings with language partners on WhatsApp
- Virtual Internship meetings and projects
- New and redesigned on-line cultural learning options:
 - Cooking (both traditional ethnic and contemporary)
 - Fitness
 - Gaming
 - Dancing
 - Discussion groups (including target-country social media settings)
 - Talent shows
 - Observation and distance enactment of holiday greetings, practices
 - Virtual site visits to museums, places of worship, theater

To increase contact time for students with the language and culture as well as opportunities for direct interactions with speakers of the language, a number of the above activities were organized as virtual events by U.S.-based specialists and target language peers during the afternoon hours in the U.S. when connections with the overseas centers were not possible due to time differences. The summary comment below offered by one student might be seen as generally indicative of those who took part in the 2020-2021 virtual programs.

“This is a tough year for everyone and adjusting to a year full of uncertainty is difficult. Nevertheless, I am very pleased with my current experience of the program and do not regret my participation.”

COMPARING STUDENT TIME-ON-TASK IN THE VIRTUAL AND IN-PERSON STUDY ABROAD CONTEXTS

The use of online student journals, also known as Language Utilization Reports (LUR), has been a part of the American Councils' advanced-level immersion programming model for well over a decade. Functioning as a tool to assist students' self-management of their time and learning, the LUR (see Davidson & Lekic, 2010) includes a brief time/place chart (Figure 2) and four open-ended prompts related to L2 linguistic and cultural situations encountered in the prior week. LURs are submitted by immersion students on a weekly basis throughout the fall and spring semesters.¹

Figure 2

Language utilization report (LUR) time/place map format

Language Use

How many clock hours during the past seven days did you spend using Russian in the following activities (remember that your individual language classes last 90 minutes each.) Please use decimal-point notation i.e. 1 hr. = "1", 45 min. = "0.75", 30 min. = "0.5", 20 min. = "0.30", 15 min. = "0.25", 10 min = "0.15". Please only record the activity for ONE week.

Activity	Sat	Sun	Mon	Tue	Wed	Thur	Fri
Formal language learning classes							
Language tutorial session							
Host family							
Internship of specialization coursework							
In public transportation or while shopping							
With friends							
Cultural events							
Russian radio or television							
Reading the press							
Professional or academic reading							
Reading for pleasure							
In homework and other preparation for the formal language learning classes							
Other							

Please specify other activities: _____

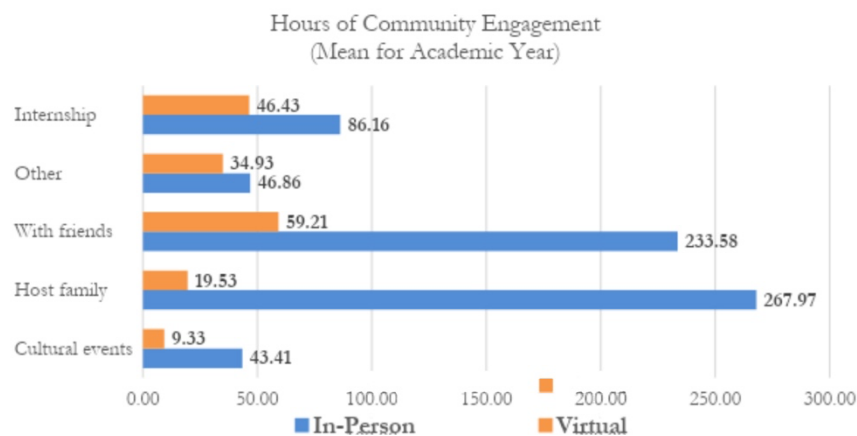
Table 3 demonstrates the significant differences in mean numbers of hours reported by virtual and overseas students in different forms of both structured language use and unstructured forms of community engagement. LUR aggregate reporting demonstrates that in-person students were able to invest, on average, more than four times the number of hours (696 hours) in community engagement activities than their peers in the virtual program (168 hours), whereas time devoted to structured activities such as small group training, individual tutorials, and direct enrollment course work are practically the same for both groups. Virtual students, indeed, devote nearly twice as much time per week in completing their homework assignments as do the in-person students.

Table 3
Time-On-Task Comparison for a Typical Virtual and In-Person Week Abroad

Average Number of Hours/Week by Type of Activity		
Activity	AY 2017-2019 <i>Immersion</i>	AY 2020-2021 <i>Virtual</i>
	October 1, 2019 – October 16, 2019	October 1, 2020 – October 16, 2020
Formal language learning classes	18.18	15.33
Language tutorial sessions	2.75	4.34
Host family	12.93	1.01
Internship or specialization coursework	0.73	0.21
In public transportation or while shopping	1.21	0
With friends	12.18	3
Cultural events	5.24	1
Russian radio or television	3.12	1.53
Reading the press	0.87	0.52
Professional/or academic reading	0.66	0.82
Reading for pleasure	1.57	0.35
Homework & other preparation for formal language learning classes	10.5	20.15
Other	2.22	0.92
Total	72.16	49.18
Total number of reported hours	2165	1130

Now drawing on the LUR database for the full academic year, Figure 3 shows mean per participant hours for the two groups dedicated to self-directed activities related to different forms of community engagement.

Figure 3
Comparison of community engagement levels of virtual and in-person L2 study abroad students



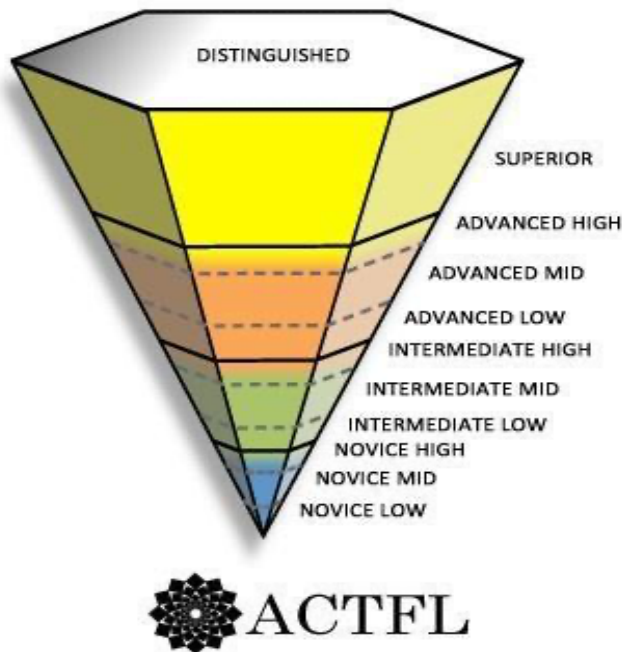
Based on the self-reported time-on-task data presented in Table 3 and Figure 3, the virtual and in-person versions of the program selected for this detailed analysis differ relatively little in terms of their corresponding allocations of structured activities: small group work,

individual tutorials, content-based lectures. Greater utilization of asynchronous activities in the virtual program may have contributed to the higher number of hours devoted to homework by those students in comparison to their in-person counterparts.

Figure 3, however, demonstrates the huge advantage enjoyed by in-person learners with respect to access to different forms of community engagement and the concomitant cultural, intercultural, and linguistic benefits these forms of immersion can provide over time within the scaffolded environment for cultural and linguistic learning described above. The year-long in-person student spends 696 hours on average in different forms of community engagement (with host families, friends, internships, volunteer service), whereas the virtual students devoted 168 hours to those same activities over the same period of time.

How important are these differences in available language utilization time if one is actively acquiring a second language at the advanced to superior level? In the case of Arabic, Chinese, or Russian, pre-/post-testing data collected over many years by American Councils indicate that 1000 to 1500 hours of combined training and immersion time are typically required to move a student of traditional college age and average or above average language learning abilities from ACTFL Advanced to Superior proficiency levels in speaking, reading and listening, and approximately 500 to 750 hours to move from Intermediate to Advanced. (See Figure 4)

Figure 4
Time-on-task estimates for Arabic, Chinese and Russian training by levels



Estimated time-on task for Arabic, Chinese and Russian:

Between Advanced and Superior
1000-1500 hours

Between Intermediate and Advanced
500-700 hours

Between Novice and Intermediate
180-250 hours



A study abroad program with a strong language and cultural acquisition component does well to align its overall instructional design, language contact model, and provisions for

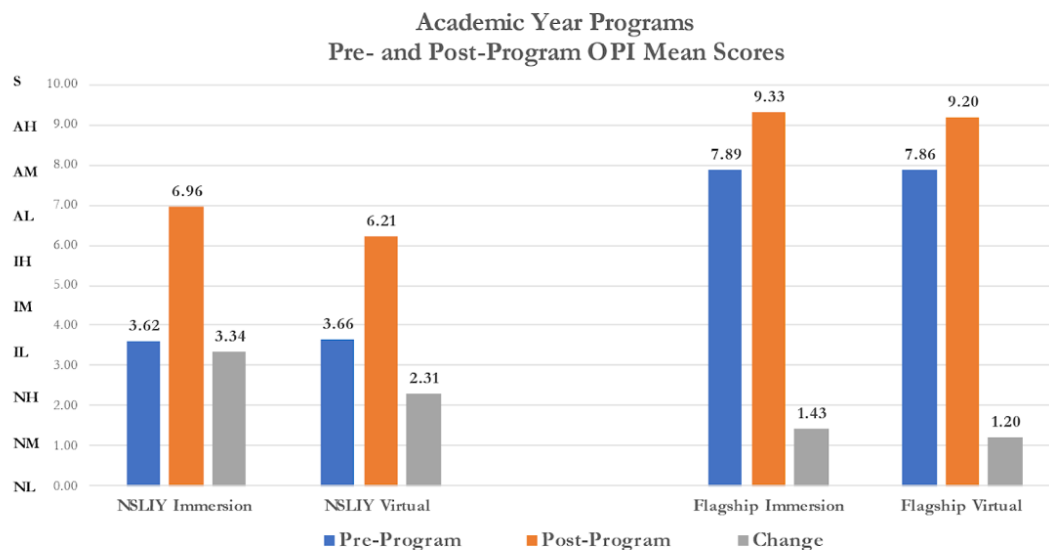
community access to the expectations of its students and stakeholders, ensuring the provision of adequate time-on-task and language support interventions for meeting the program's stated learning objectives.

COMPARING STUDENT PROFICIENCY OUTCOMES IN THE VIRTUAL AND IN-PERSON STUDY ABROAD CONTEXTS

Comparison of the speaking proficiency outcomes by language and training levels of in-person cohorts and virtual cohorts (in person N=1388 and virtual N=770) revealed a relatively small but *consistent gap in mean gain levels across languages*, with participants in the virtual groups attaining on average one proficiency sub-level *lower* in speaking gains than the corresponding face-to-face students.

Figure 5

Pre-post oral proficiency interview (OPI) score comparisons academic year (in-person vs virtual) NSLI-Y and Flagship



Chinese & Russian Flagship Immersion = 189 Chinese & Russian Flagship Virtual = 69 NSLIY Immersion* = 223 NSLIY Virtual* = 50
* p<0.001

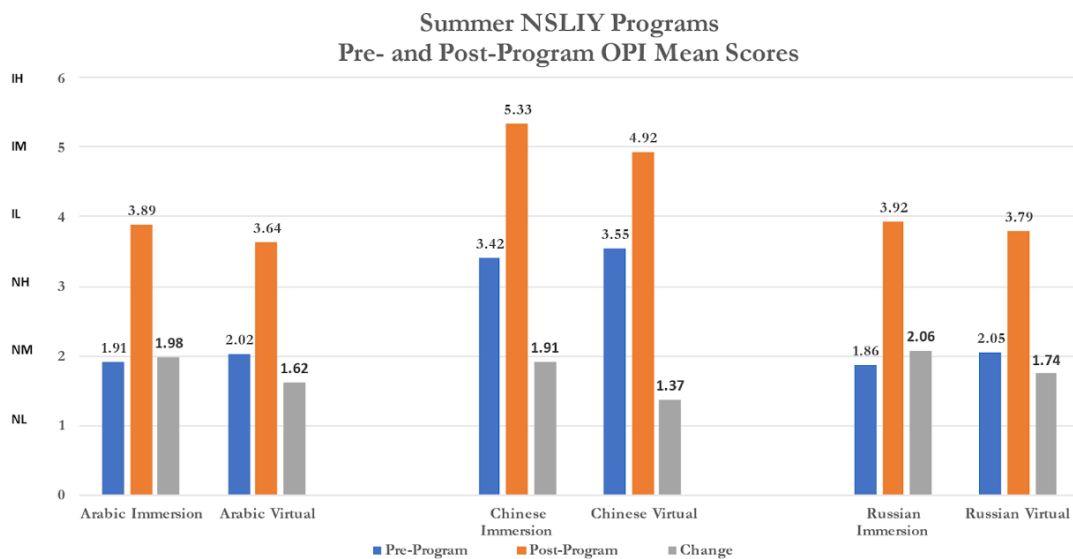
Figure 5 presents summative data for two federal programs providing year-long study abroad in the critical languages: NSLI-Y and Flagship, juxtaposing immersion (in-person) results (N=412) with those of the corresponding virtual versions of each program (N=119). Keeping in mind that NSLI-Y and Flagship operate in multiple locations and across multiple languages, mean pre-program proficiency levels for all tested participants in each program were calculated (blue bar), along with mean post-program proficiency levels (orange bar) and, finally, mean Oral Proficiency Interview (OPI) gain values (post-score minus pre-score) (gray bar), which are then aligned against a vertical axis marked by an intervallic conversation scale that converts ACTFL OPI levels (ordinal values) to numeric values for analytic purposes.²

The pre-program values (blue bars) are helpful in accounting for any variations in the levels of preparation of students entering the program in the COVID year as opposed to the

previous years. In these cases, variation levels are small and not statistically significant. In discussing mean differences in measured gains across programs, it is important to keep in mind the nature of the ILR/ACTFL proficiency measurement artifact (the inverted pyramid, see Figure 4) used to rate the learners L2 proficiency level. Proficiency growth from Novice Mid to Novice High might require 50-100 hours of training in a critical language, whereas movement from Advanced Mid to Advanced High might require 500 hours. In both cases, the “gain” would be represented by the same value (1.0).

Comparison of the OPI test results of the early-stage learners of Arabic, Chinese and Russian participating in the NSLI-Y immersion (N=400) and virtual (N=299) summer programs showed statistically significant differences in post-program mean scores for all groups with small effect size (See Figure 6).³ NSLI-Y groups registered substantial pre- to post-program gains in OPI levels with the virtual groups actually demonstrating higher pre-program OPI scores than their immersion counterparts. The in-person programs nonetheless produced higher mean post-scores and greater net gains across all languages and groups than the corresponding virtual programming.

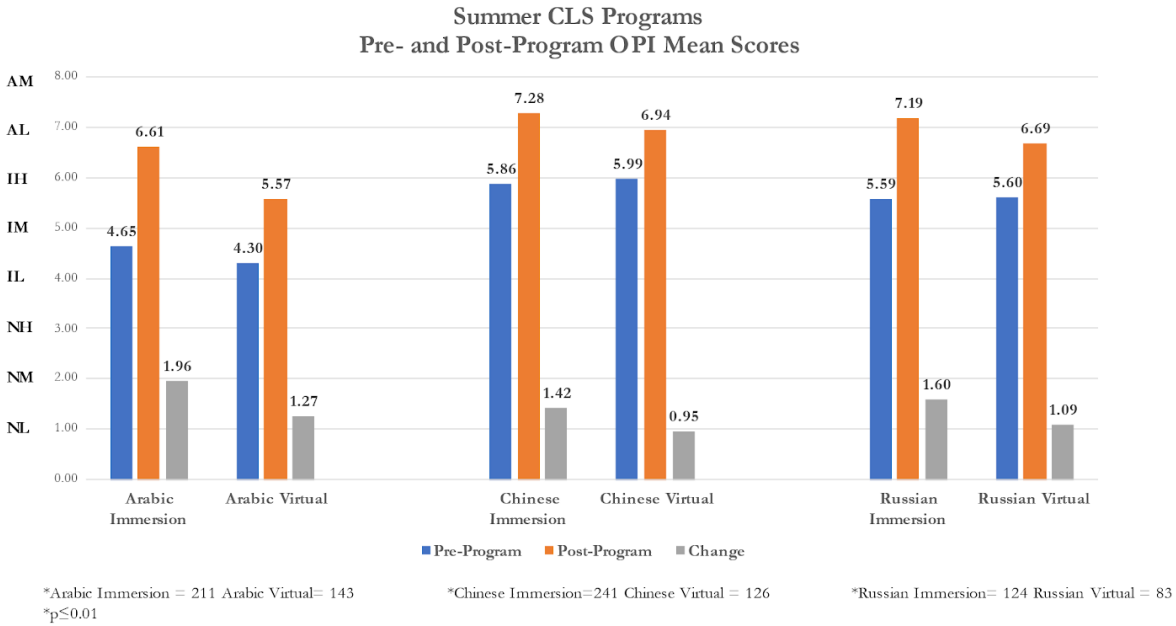
Figure 6
Pre-post OPI score comparisons (in-person vs virtual) NSLI-Y summer program



*Arabic Immersion = 97 Arabic Virtual = 99 *Chinese Immersion = 187 Chinese Virtual = 91 *Russian Immersion = 116 Russian Virtual = 109
p ≤ 0.05

Turning next to a comparison of three CLS Summer programs by individual language (Figure 7), the differences in attainment levels by language as well as across immersion (N=576) versus virtual (N=352) are clearly observable and, as was the case with the NSLI-Y and Flagship programs, statistically significant.⁴

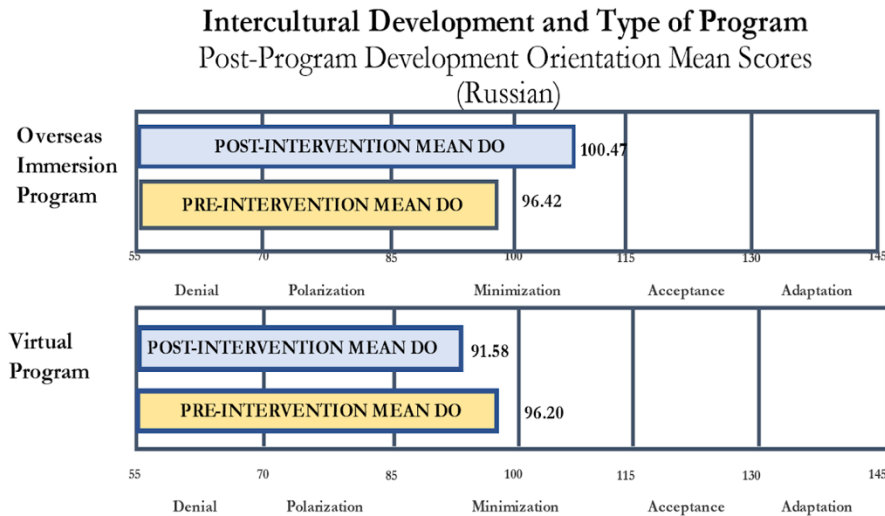
Figure 7
 Pre-post OPI score comparisons (in-person vs virtual) CLS summer programs



Preliminary comparisons of reading and writing outcomes for those programs testing those skills showed smaller differences between virtual and in-person for reading and writing proficiency test results than for speaking, in some cases results that were essentially indistinguishable from the in-person immersion results.

Among the relatively small number of available assessment instruments for evaluating the development of intercultural competencies, the Intercultural Development Inventory (IDI) is widely used in the professions and among universities, study abroad researchers and the U.S. government (Deardorf, 2004; Hammer, 2012). IDI test results are not recommended for use in placement or certification decisions but have been in use in the U.S. and overseas for more than 10 years in measuring individual and group-level change related to the individual’s development orientation and level of ethnocentrism (ranging from denial and polarization to minimization, acceptance and adjustment phases) (Bennett, 1998; Hammer, 2012).

Figure 8
Comparison of intercultural development index pre/post assessment results for virtual and in-person learners



The IDI results shown in Figure 8 should be regarded as preliminary and still in need of replication due to the small sample size. Although all participants in both the in-person and the virtual cohort had studied abroad previously, the pre-program mean IDI score levels reported here are generally high in comparison with those reported by other well-established overseas programs (Baker-Smemoe et al., 2014; Watson et al., 2013). Whether a -4.62-point decline in IDI score for the virtual group (as opposed to a +4.02-point increase for the immersion group) is an anomaly or not, it is clear from the time-on-task data that the online group had far less opportunity for direct contact with target-language peers and community engagement in comparison to their in-person counterparts (See Table 3). As a result, it is not surprising that the IDI scores for this group are lower. If further research confirms these findings, then advocates of “virtual immersion” programs in the future should seriously consider program designs capable of addressing the cultural and intercultural educational goals, as well as the purely linguistic goals of students and stakeholders committed to this form of language and cultural learning.

OBSERVED DIFFERENCES IN SPEECH PRODUCTION OF VIRTUAL VERSUS IN-PERSON IMMERSION STUDENTS

In addition to course examinations and proficiency testing, the year-long L2 Flagship programs require a final public presentation at the conclusion of the year. In the case of the Flagship programs, these presentations take place before a local audience that includes teachers, peers, and internship supervisors and are recorded professionally, transcribed, and archived for later access by the student, and, with permission, by members of the community and also prospective employers. Using the archived video recordings of virtual and in-person student presentations, program advisors and faculty members were requested retrospectively

to review and comment on any differences that distinguished the language and communicative behaviors of the virtual students from those of previous years on the in-person programs. As was evident from Figure 5, the measured proficiency results of the two groups were not great and in some cases the proficiency scores of the students were identical. Nonetheless, the following characteristics were noted:

- Activation of registers slower, less certain than in previous years
- Conversational interactions on Zoom are slower, less spontaneous
- Speech and cultural behaviors lack the ease and familiarity of previous years
- Listening comprehension was slower, based on student response time
- Challenges in perceiving and utilizing tones in Mandarin were observed
- Socio-pragmatic strategy selections and performances less certain
- Awareness of audience expectations, cultural referencing, public demeanor/body language less well-mediated
- Use of collocations, multi-word structures, conventionalized phrases less nativelike than in previous years.

Many of the observed *weaknesses* of the virtual groups in Arabic, Chinese, and Russian are still consistent with ILR-3 and ACTFL Advanced High and Superior assessment guidelines for speaking, but for those planning careers that will require professional level linguistic and cultural competencies, the speech and presentational issues noted above are indeed relevant and worthy of note in the future planning of virtual learning.

POSITIVE CONTRIBUTIONS FROM THE VIRTUAL YEAR EXPERIENCE

On a more positive note, the virtual format was praised by students and faculty for its introduction of instructional management platforms into the overseas programs, as well as for the benefits brought by *flipped* classes, more reading assignments, and oral presentations and more extensive feedback on written work. Students generally rated their virtual interactions with tutors, peers, and homestay hosts highly, and expressed appreciation for the flexibility and special efforts of teachers and support staff in responding to their needs over the course of the virtual year.

Virtual learning has already had a positive effect on planning for the resumption of study abroad L2 programming for 2022-2023 and beyond. Among those “lessons learned”:

- Mobilization of learning technologies in support of overseas curricula
- Introduction of learning management systems in many overseas locations
- More realistic expectations of the linguistic and cultural outcomes online learning can produce
- New opportunities for those who cannot travel abroad for language learning (for health, family, or employment reasons)
- New opportunities for language maintenance for those with previous study abroad
- New opportunities for virtual pre-SA language training to help open the door wider for early-stage learners to study a language overseas, such as the creation of new online summer and academic-year cost-free courses for both high school and

college students for the study of languages like Arabic, Chinese, Farsi, Korean, Russian or Turkish through the National Security Language Initiative for Youth (NSLI-Y) and the Critical Language Scholarship (CLS) Spark Program (NSLI-Y, CLS “Spark”).

For example, both the NSLI-Y and the CLS programs have introduced fully funded pre-study abroad language training for prospective applicants from institutions or school districts where foreign languages are not offered. Similarly, opportunities announced in 2020-2021 for alumni of the CLS program to pursue online training to refresh their language and cultural knowledge was met with such interest by program alumni that the program was expanded by the State Department in 2021-2022.

DISCUSSION

The COVID pandemic triggered a rapid process of virtualization and digitalization of educational activities across disciplines and at all levels of education throughout much of the world. The foreign language and study abroad fields were no exception. The issue posed by COVID in the study abroad field concerned the transferability of the formidable linguistic, cultural, and intercultural growth potential associated with study abroad learning to the virtual setting, and, eventually, the potential reciprocal benefits of the strategies and practices of the virtual classroom for the post-COVID overseas setting.

Successful L2 learners learn to draw on their critical thinking skills, socio-emotional maturity, and general cultural as well as culture-specific experience and procedural knowledge gained from a substantial period of study abroad to engage effectively and appropriately with a range of representatives of the target culture. To explore the extent to which their virtual counterparts were able to do the same, the present study has made use of quantitative and qualitative data sources to compare the evidence of linguistic and cultural gains demonstrated by in-person and virtual students at comparable levels of study of the same languages. The control group was represented by the immersion cohorts of Arabic, Chinese and Russian students, who studied overseas in 2017-2018 and 2018-2019 (N = 1388), while the intervention group completed their training in virtual settings in 2020-2021 (N=770).

Comparison of the speaking proficiencies by language and training levels of in-person and virtual cohorts revealed a relatively small but *consistent gap in mean gain levels across languages*, with participants in the virtual groups attaining on average one proficiency sub-level *lower* in speaking proficiency than the corresponding face-to-face students, when controlling statistically for the pre-program levels of each. By contrast, comparisons of reading and writing outcomes showed smaller differences between virtual and in-person than for speaking with the results in some cases indistinguishable. The assessment of cultural competencies of the virtual students were based in-part on the reports of senior language assessment experts in each of the three languages who were asked to compare the virtual students’ video recorded speech samples with counterparts in the control group. As noted, expert comments are not about proficiency levels (which were certified independently) but comments on the clarity, fluency, and cultural appropriateness demonstrated as they communicate on professional topics with host-country counterparts. Like the IDI results reported above, expert reviewers cited multiple examples of the language and cultural behaviors of the graduates of the virtual programs who were seen to fall short in salient categories in comparison with the in-person groups.

To further clarify the differences in the outcome results reported here between the virtual and in-person learners, this paper also examines the data on culture contact and time-on-

task that are central to the dynamics of overseas immersion learning. The “contact hypothesis,” first advanced by Allport (1954), undergirds the concept of much study abroad thinking today by establishing the link between moderated and extended contact with representatives of a different culture and the reduction of cultural stereotypes and other barriers to mutual understanding. Within the study abroad community, that process has long been viewed as potentially *transformational* for those who undergo it, well before transformational learning was introduced by Mezirow (1994) into modern educational theory. Extended linguistic and cultural contact, as Allport predicted and every language teacher knows, are central to intercultural development. For that reason, the quantity of available time for linguistic and cultural interactions was analyzed for both groups in Table 3. Those comparisons, in turn, have demonstrated the great advantage enjoyed by in-person learners, for example, for multiple forms of community engagement. While formal class hours are nearly the same for virtual and in-person study abroad students, the year-long in-person student who has reached the advanced level spends on average 696 hours in different forms of community engagement (e.g., with host families, friends, internships, volunteer service), whereas the virtual students devoted 168 hours to those same activities over the same period of time.

Given these stark differences in range and levels of cultural contact as well as in available time for engaging them, there can be little surprise that the in-person study abroad students outperform their virtual counterparts in speaking, listening comprehension, and intercultural skills, including identity competence, while the virtual students were seen to hold their own, but not exceed control group members, in other key literacies, such as reading and writing across levels and access to target language social media in the languages under study. The limitations (some of them inherent) of the virtual learning models examined here should be duly noted by program planners and decision makers, but virtual models were nonetheless shown to accomplish a great deal, given the hours and investments required for their implementation. The 2020-2021 virtual immersion experience has demonstrated the potential of online learning and the areas where it can produce results on a par with face-to-face immersion. And in so doing, the operationalization of virtual learning models across languages and at multiple levels due to COVID have now cleared the way for the introduction of a new generation of skill-specific and other much needed language and cultural learning, such as language-maintenance or *refresher* programs, *bridge* programs, content-based offerings in target languages, and low-cost or cost-free introductory courses for new L2 learners in regions of the U.S. without access to the study of foreign language, as is illustrated by the Spark program, the virtual spin-off of the Critical Language Scholarship program and others.

NOTES

¹For the purposes of comparing time-on-task data for virtual and in-person programming, the present study has selected two advanced-level learner cohorts of the same language who studied with the same faculty using the same general syllabus: one in-person (2017-18) and one virtual (2020-21).

²An independent t-test showed a significant difference between the Flagship OPI post program mean scores for the virtual (M= 9.20 SD=.678) and in-person programs (M=9.33 SD=.668; df(255)= p≤0.001, with a small effect d=0.17 95% CI [.076-.329]. An independent t-test showed a significant difference between the NSLI-Y OPI post program mean scores for the virtual (M= 6.21 SD=1.301) and in-person programs (M=6.96 SD=.997; df(255)= p≤0.001, with a small effect d=0.23 95% CI [.397-1.112].

³An independent t-test showed a significant difference between the NSLI-Y Summer Arabic OPI post program mean scores for the virtual (M= 2.02 SD=1.45) and in-person programs (M=1.91 SD=.1.259; df(94)= p≤0.01, with a small effect d=0.20 95% CI [.495-0.27]. An independent t-test showed a significant difference between the NSLI-Y Summer Chinese OPI post program mean scores for the virtual (M= 4.92 SD=1.507) and in-person programs (M=5.33 SD=1.243; df(289)= p≤0.01, with a small effect d=0.30 95% CI [.073-.744]. An independent

t-test showed a significant difference between the NSLI-Y Summer Russian OPI post program mean scores for the virtual (M= 3.79 SD=1.229) and in-person programs (M=3.92 SD=1.151; df(206)= $p \leq 0.001$, with a small effect $d=0.19$ 95% CI [-0.178 -.452].

†An independent t-test showed a significant difference between the CLS Arabic OPI post program mean scores for the virtual (M= 5.57 SD=1.532) and in-person programs (M=6.61 SD=.1.137; df(353)= $p \leq 0.01$, with a small effect $d=0.37$ 95% CI [.745-1.345]. An independent t-test showed a significant difference between the CLS Chinese OPI post program mean scores for the virtual (M= 6.94 SD=1.249) and in-person programs (M=7.28 SD=1.109; t(336)= $p \leq 0.001$, with a small effect $d=0.33$ 95% CI [.096-.596]. An independent t-test showed a significant difference between the CLS Russian OPI post program mean scores for the virtual (M= 6.69 SD=1.313) and in-person programs (M=7.19 SD=1.066; df(206)= $p \leq 0.001$, with a small effect $d=0.47$ 95% CI [.0180-.833].

REFERENCES

- Allport, G. W. (1954). *The Nature of Prejudice*. Addison-Wesley Pub. Co.
- Baker-Smemoe, D., 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* (pp. 21-71). Intercultural Press.
- Bennett, M. J. (1998). Intercultural communication: A current perspective. In M. J. Bennett (Ed.) *Basic concepts of intercultural communication: Selected reading*. Intercultural Press.
- Brewer, M. B., & Gaertner, S. L. (2001). Toward a reduction of prejudice: intergroup contact and social categorization. In R. Brown & S. L. Gaertner (Eds.), *Blackwell handbook of social psychology: Intergroup processes* (pp 451-474). Blackwell.
- Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Multilingual Matters.
- Byrnes, H. (2013). Notes from the Editor. *Modern Language Journal*, 97(4), 825-827.
- Cohen, E. G. (1975). The effects of desegregation on race relations. *Law and Contemporary Problems*, 39(2), 271–299. <https://doi.org/10.2307/1191102>
- Commission on Language Learning. (2017) America's languages: Investing in language education for the 21st century. American Academy of Arts and Sciences. <https://www.amacad.org/publication/americas-languages>
- Davidson, D. E. (2015). The development of L2 proficiency and literacy within the context of the federally supported overseas language training programs for Americans. In J. Bown & T. Brown (Eds.) *To advanced proficiency and beyond: Theory and methods for developing superior second-language ability* (pp. 117-150). University Press.
- Davidson, D. E., Garas, N., & Lekic, M. D. (2016). Assessing language proficiency and intercultural development in the overseas immersion context. In Murphy, D. & Evans-Romaine, K. (Eds.) *Exploring the US Language Flagship Program: Professional competence in a second language by graduation* (pp. 156-176). Multilingual Matters.
- Davidson, D. E., Garas, N. & Lekic, M. D. (2021). Transformative language learning in the overseas immersion environment: Exploring affordances of intercultural development. In Leaver, B. Davidson, D.E, & Campbell, C. (Eds.). *Transformative language learning and teaching* (pp. 109-119). Cambridge University Press. <https://doi.org/10.1017/9781108870788>
- Davidson, D. E., & Lehmann, S.G. (2005). A longitudinal survey of the language learning careers of ACTR Advanced Students of Russian: 1976-2000. *Russian Language Journal*, 55(1), 193-221. <https://scholarsarchive.byu.edu/rlj/vol55/iss1/11>
- Davidson, D. E., & Lekic, M. D. (2010) The overseas immersion setting as contextual variable in adult SLA: Learner behaviours associated with language gain to level-3 proficiency in Russian. *Russian Language Journal*, 60(1), 53-77. <https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1173&context=rlj>
- Deardorf, D. (2004). *Internationalization: In search of intercultural competence*. Duke University Press.
- Dovidio, J., Kawakami, K., & Beach, K. (2001). Implicit and explicit attitudes: Examination of the relationship between measures of intergroup bias. In R. Brown & S. L. Gaertner (Eds.), *Blackwell Handbook of social psychology: Intergroup relations* (pp. 175-197). Blackwell.
- Dirkx, J., Mezirow, J., & Cranton, P. (2006). Musings and reflections on the meaning, context, and process of transformative learning: A dialogue between John M. Dirkx and Jack Mezirow. *Journal of Transformative Education*, 4(2), 123–139. <https://doi.org/10.1177/1541344606287503>

- Ehrman, M. E., & Dornyei Z. (1998). *Interpersonal dynamics in second language education: The visible and invisible classroom*. Sage Publications.
- Freed, B. F., Segalowitz, N., & Dewey, D. P. (2004). Context of learning and second language fluency in French. Comparing regular classroom, study abroad, and intensive domestic immersion programs. *Studies in Second Language Acquisition*, 26(2), 275-301.
- Hammer, M. R. (2012) The intercultural development inventory: A new frontier in assessment and development of intercultural competence. In M. Vande Berg, R. M. Paige, & K. H. Lou (Eds) *Student Learning Abroad* (pp 115-136). Stylus Publishing.
- Kinginger, C. (2008). Language learning in study abroad: Case studies of Americans in France. *Modern Language Journal*, 92(1), 1 -124. <https://doi.org/10.1111/j.1540-4781.2008.00821.x>
- Leaver, B. L., Davidson, D. E., & Campbell, C. (2021). Introduction. In B.L. Leaver, D. E. Davidson, & C. Campbell (Eds.), *Transformative Language Learning and Teaching* (pp 1-10). Cambridge University Press.
- Mezirow, J. (1978). Perspective transformation. *Adult Education Quarterly*, 28(2), 100–110.
- Mezirow, J. (1994). Understanding transformation theory. *Adult Education Quarterly*, 44(4), 222-240.
- Mezirow, J. (2012). Learning to think like an adult: Core concepts of transformation theory. In E. Taylor & P. Cranton (Eds.), *The handbook of transformative learning. Theory, research, and practice* (pp. 73-95). Jossey-Bass.
- OECD, (2020). *Education at a Glance*. Organization for Economic Cooperation and Development. https://www.oecd-ilibrary.org/education/education-at-a-glance-2020_974729f4-en
- Pellegrino Aveni, V. A. (2005). *Study abroad and second language use: Constructing the self*. Cambridge University Press.
- Pettigrew, T. F., & Tropp, L. R. (2006). A meta-analytic test of intergroup contact theory. *Journal of Personality and Social Psychology*. 90(5), 751–783. <https://doi.org/10.1037/0022-3514.90.5.751>
- Quan, T. (2019) Competing identities, shifting investments, and L2 speaking during study abroad. *L2 Journal*, 11(1), 1-19. <http://repositories.cdlib.org/ucclt/12/vol11/iss1/art1>
- Salisbury, M. H. (2011). The effect of study abroad on intercultural competence among undergraduate college students [Unpublished doctoral dissertation]. University of Iowa.
- Watson, J. R., Siska, P., & Wolfel, R. L. (2013) Assessing gains in language proficiency, cross cultural competence, and regional awareness during study abroad: A preliminary study. *Foreign Language Annals*, 43(3), 504-530. <https://doi.org/10.1111/flan.12016>