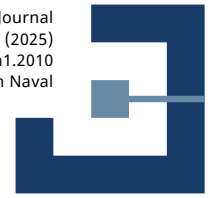
The JALT CALL Journal
ISSN 1832-4215, 21(1), 2010 (2025)
https://doi.org/10.29140/jaltcall.v21n1.2010



Castledown

The impact of COIL on EFL learners' communication confidence and soft skills

Robert Remmerswaal
Sojo International Learning Center, Sojo University, Kumamoto, Japan
robertr@m.sojo-u.ac.jp

Hülya Tuncer
Department of English Language Teaching, Çukurova University, Adana, Türkiye
hulyatncr@gmail.com

Joaquín Naval
Department of Economics, University of Girona, Girona, Spain
joaquin.naval@udg.edu

This paper investigated the effect of a Collaborative Online International Learning (COIL) project on English as a Foreign Language (EFL) learners' self-perceived communication competence (SPCC), communication anxiety (CA), and soft skills. The participants were 157 university students from Japan, Türkiye, and Spain who completed various tasks and created a video on Sustainable Development Goals during the fall of 2022. Students answered two surveys (pre-COIL and post-COIL) through a mixed-method convergent research design. Both surveys included three different 5-point Likert scales (analyzed via *t*-test and convergence analysis), and the post-COIL survey included additional open-ended questions for qualitative analysis. The *t*-test analysis demonstrated a statistically significant increase in EFL students' average SPCC. The convergence analysis showed that students with the lowest and highest SPCC, CA, and soft skills converged towards the average and experienced the greatest change in these traits post-COIL. Considering the qualitative data and social cognitive theory, participation in the COIL assisted students in better understanding their SPCC, CA, and soft skills when working in an international group.

Keywords: self-perceived communication competence (SPCC), communication anxiety, soft skills, COIL, social cognitive theory

Introduction

Collaborative activities have been widely advocated for in foreign-language learning as they provide increased language practice, improve student talk quality, individualize learning, create a positive affective atmosphere, and motivate learners (Long & Porter, 1985). Virtual exchanges (VEs) promote language acquisition through various strategies of negotiation of meaning (Canals, 2023), regardless of linguistic background or type of pairing (Wang, 2004). Synchronous and asynchronous exchanges show similar results to face-to-face exchanges (Blake, 2011; Sotillo, 2000; Yanguas, 2010; Ziegler, 2016). VEs increase students' confidence in their English communication skills (Mestre-Segarra & Ruiz-Garrido, 2022; O'Dowd, 2021), oral skills and motivation to learn English (Canals, 2020), and foreign language enjoyment (Resnik & Schallmoser, 2019). Additionally, VEs allow students to improve digital literacy, cross-cultural understanding, collaboration skills, creativity (Wicking, 2022), and intercultural competence (Hackett et al., 2023).

Collaborative Online International Learning (COIL) falls under the umbrella of VE. It does not have a universally accepted definition, it is generally accepted to be a collaborative exercise of teachers and students, uses online technology and interaction, has potential international dimensions, and is integrated into the learning process (Rubin, 2022). COIL is a relatively new method, with research focused on best practices, case studies, and recommendations on implementations or evaluations (Hackett et al., 2023). Most COILs are five to eight weeks and focus on collaborative project work (Rubin, 2022). The collaborative nature of a COIL is similar to that of a global virtual team, making the interactions situationally and interactionally authentic (Ellis, 2017).

However, current research on COIL and VEs has focused mainly on second-language acquisition (SLA) or intercultural competence. This paper focuses on communication confidence because it strongly predicts an individual's likeliness to communicate (MacIntyre et al., 1998; Samvati & Golaghaei, 2017; Yashima, 2002). Communication confidence is measured using self-perceived communication competence (SPCC) and communication anxiety (CA) (Yashima, 2002). We also focus on soft skills because VEs are a way to develop generic, interrelated, and transferable skills directly related to successful professional practices in the global workforce (Lewis & O'Dowd, 2016). Soft skills are essential in modern societies because they help people adapt to the requirements of dynamic labor markets (European Commission, 2011). Increases in SPCC and soft skills and decreases in CA would indicate that students are better prepared for the global workforce. To our knowledge, there has been no study on the effect of COIL on SPCC, CA, and soft skills.

The present study examines an eight-week COIL project involving three universities in Japan, Türkiye, and Spain through a social cognitive theory (SCT) lens. It investigates whether student responses to SPCC and CA items change during a COIL VE. Moreover, this study explores potential changes in participants' soft skills development. These three measurement tools for SPCC, CA, and soft skills are based on students' perceptions of themselves. To understand the effects of this COIL project on English as a Foreign Language (EFL)

learners, quantitative and qualitative data were collected to answer the following research question:

How does COIL in English affect EFL learners' SPCC, CA, and soft skills?

Literature review

Social cognitive theory

This study focuses on changes in students' perceptions, as most people tend to make decisions following their perceived ability rather than their actual competencies (Bandura, 1997; McCroskey & McCroskey, 1988; Wyatt, 2022). SCT developed by Albert Bandura, is a framework for understanding, predicting, and changing human behavior (Bandura, 2001). It includes key tenets such as observational learning, reciprocal determinism, self-efficacy (SE), and reinforcements. As SE is perception-based, this study utilizes SE to interpret changes in student perceptions post-COIL. SE is a key determinant of behavior and is an individual's belief in their ability to succeed in a specific task (Bandura, 1997). People with high SE are more likely to set challenging goals, persist when a task becomes difficult, and tend to have lower anxiety (Wyatt, 2022).

Four factors affect the development of SE: mastery experiences, vicarious experiences, social persuasion, and physiological and emotional states, with mastery experiences being the most significant influence (Bandura, 1997; Wyatt, 2022). For the completion of a task to be considered a mastery experience, the individual must have believed it was challenging and successfully completed. In contrast, failure in a task can potentially lower SE, especially for those whose SE already measured low or who had limited previous experiences. SE is typically measured using scales that measure a person's beliefs about their ability to complete a specific task (Bandura, 1997; 2006). Researchers can interpret changes in SE by considering the four influencing factors post-intervention. Viewing changes in students' perceptions of their SPCC, CA, and soft skills through an SCT lens can provide a rationale for any changes in these traits post-COIL.

Communication confidence

Communication confidence is an individual's SPCC and lack of CA (MacIntyre et al., 1998; Yashima, 2002). Communication confidence is a strong predictor of willingness to communicate (WTC) (Fushino, 2010; Yashima, 2002) and is correlated to foreign language enjoyment (Zhang et al., 2024). Independently, SPCC measures an individual's beliefs in their ability to communicate in various situations (Leeming et al., 2024). Individuals with higher SPCC engage in more frequent communication (Balouchi & Samad, 2021; Leeming et al., 2024), as individuals who perceive themselves as competent communicators tend to be more willing to communicate (Samvati & Golaghaei, 2017). Additionally, higher SPCC scores correlate with increased English communication frequency inside

and outside the classroom and online (Balouchi & Samad, 2021; MacIntyre & Charos, 1996).

CA, also known as communication apprehension, is a measure of an individual's frequency of nervousness in various situations (McCroskey, 2009). Arnold and Brown (1999) identified anxiety as the main factor that interfered with the learning process and the most common learning disorder. Students with a reduced CA often see a positive impact on their academic achievement regardless of their academic ability (Kim & Choi, 2023). Hashimoto (2002) found a statistically significant negative correlation between CA and SPCC, which supports their interconnected nature in forming communication confidence in Yashima's (2002) model.

Many papers have used SPCC, but there is criticism that in studies outside of the US, the construct fails tests of measurement invariance (Croucher et al., 2020). However, invariance tests favor constructs with low between-group variance, suggesting this criticism may be overstated (Welzel et al., 2023). Another criticism is that SPCC is not a measurement of communication competence, and there is little correlation between SPCC and speaking proficiency (Croucher et al., 2020; Lockley, 2013). This study only concerns student perceptions, which influence behavior (Bandura, 1997).

Soft skills

Soft skills are socio-emotional skills that positively influence personal life, mental health, social success, and the labor market (Organisation for Economic Co-operation and Development, 2015). Current graduates must develop soft skills, equipping them with the transferable skills required for intellectual, professional, and social activities (European Commission [EC], 2011; Llorens et al., 2024; Ministry of Education, Culture, Sports, Science and Technology, 2012). They are non-job specific, closely connected with attitudes, and difficult to quantify and develop (EC, 2011). Other names include 21st-century skills, social skills, cross-disciplinary skills, and interpersonal skills. The EC lists 22 skills transferable across various work and social situations (EC, 2011). However, many of the soft skills from the EC overlapped. Lucisano and du Mérac (2019) narrowed these skills to 10: autonomy, collaboration, commitment, curiosity, empathy, leadership, openness, problem-solving, resilience, and self-confidence. The EC (2011) suggests that educators provide students in higher education with intercultural collaboration opportunities to develop these skills through practical experience. Cooperative learning, as done in this COIL, has been found to increase soft skills due to the practical experience it provides (Gillies, 2016; Llorens et al., 2024), which aligns with the concept of SE.

Effects of a virtual exchange on communication confidence and soft skills

Two recent studies suggest that a VE positively affects WTC (Rahimi & Fathi, 2022; Zhou, 2023). Since WTC and communication confidence are correlated constructs, a change in WTC will also likely influence communication

confidence. Both studies utilized an eTandem model of VE, with the defining trait of students using two languages and acting as both learners of one language and teachers of their native languages (Lewis & O'Dowd, 2016). This model differs from the COIL model, where students communicate with English as a lingua franca. In these studies, students with English proficiency levels of CEFR B1-B2 (Rahimi & Fathi, 2022) and IELTS 6.0 or above (Zhou, 2023) participated in the exchange for several weeks. At the conclusion of the VE, students in the eTandem groups, who were partnered with English speakers outside their country, showed a greater increase in their WTC and speaking skills compared to a control group, which was partnered with their classmates. All students completed the same speaking activities, highlighting the effectiveness of the VE in enhancing communication skills.

VEs often include goals to develop creativity and collaboration skills, though changes in these skills are not frequently measured (Wicking, 2022). Háhn and Radke (2020) hosted a VE to recreate an authentic situation for their students to develop soft skills. Their tourism students worked together to create different promotional materials for their cities. Student feedback demonstrated that groups could overcome their disagreements, they felt the skills they learned were transferable to their future working lives, and they had improved their ability to communicate with people from other cultures. Mestre-Segarra and Ruiz-Garrido (2022) hosted a COIL between their business master's students, which, among other things, considered the application of cross-disciplinary skills. Student expectations for using these skills pre-COIL exceeded their perceptions of using them as measured post-COIL, though students reported the skills being beneficial for the project. The present study looks to expand on this by quantifiably measuring the changes in 10 soft skills and allowing students to address any perceived changes qualitatively. This analysis will allow for a richer interpretation of the quantitative results. Approaching the results through an SCT lens facilitates a nuanced examination of skill acquisition, stability, or disconnect from perceptions. It may also provide insight into which skills require further nurturing.

Methodology

Participants

For eight weeks, from October to December 2022, the COIL took place between 55 aerospace and mechanical engineering students from Sojo University (Japan), 46 pre-service English teachers from Çukurova University (Türkiye), and 56 economics students from the University of Girona (Spain). All three instructors taught their regular classes of English Communication (Japan), Critical Reading and Writing (Türkiye), and International Macroeconomics (Spain) and created a joint project that could feature in all three courses. This study did not consider gender, age, or English language level, which limits the generalizability of this study.

Of the 157 participants, 39 did not submit one or both surveys, though they

participated in all COIL activities and meetings. While additional responses may have changed the overall results, completing the surveys was voluntary, in accordance with the ethical standards of all three universities. Thus, 118 (38 Japanese, 46 Turkish, and 34 Spanish) voluntarily answered pre- and post-COIL surveys (see the questions in Appendix A). Tables 2 to 4 provide the number of participants for each open-ended question. Each author received separate Ethical Committee approval from their universities, and the authors anonymized student data by assigning random alphanumerical values to students before sharing it.

Context

The exchange included individual and group activities but focused primarily on three assignments completed in groups of six to seven students (Table 1). The authors designed pre-COIL tasks limited to information exchange (see O'Dowd & Waire, 2009) as initial ice-breaking activities. Students were assigned a group for the duration of the exchange in Microsoft Teams and introduced themselves via a recorded video. Students also shared their pre-COIL task of creating an infographic on a Sustainable Development Goal (SDG) that interested them. SDGs represent strategies all United Nations Member States adopted to end poverty, improve education and health, reduce inequality, and spur economic growth (United Nations, n.d.).

The authors designed the COIL comparison and analysis tasks and collaboration and product creation tasks, as described in O'Dowd and Waire's (2009) VE task typology. Each assignment began with a COIL-wide video call hosted by the instructors. Additionally, Blooket.com quizzes in the Microsoft Teams meetings acted as ice-breaking activities and a way to introduce assignments during the COIL-wide video calls. This call was followed by the students breaking into their small groups. In the first two assignments, students provided ways to work towards their SDG on campus and in the community. These tasks were designed to prompt discussions via comparison and analysis and keep a record of the group's thoughts (examples in Appendix A). Assignment three was a collaboration and product creation task. Students shared the most relevant information from the first two assignments in a five-minute video. Students organized additional group meetings independently to complete these tasks, typically outside of class time. These independent meetings were a mix of synchronous and asynchronous calls and messages through Teams, Instagram, or WhatsApp. The authors offered support for interactions via Microsoft Teams, but students decided which technology was best for their groups.

Table 1. COIL activities and timeline

Length and focus	Task typology	Tasks	Interaction
Weeks 1–3 Pre-COIL activities	Information exchange	Create an introduction video and an infographic on an SDG, post both to your group's Teams channel, and leave comments.	Asynchronous sharing of media via MS Teams
Weeks 4–6 SDGs at school	Comparison and analysis	Participate in a large group call on Teams. Complete a Blooket.com quiz. Meet with your group and choose an SDG from one of the infographics created by a group member. Complete assignment one synchronously and asynchronously.	Group communication (with 6–7 members) via MS Teams or WhatsApp/ Instagram
Weeks 7–9 SDGs in the community	Comparison and analysis	Participate in a large group call on Teams. Complete a Blooket.com quiz. Meet with your group and choose a company working towards an SDG in your community. Complete assignment two synchronously and asynchronously.	Group communication (with 6–7 members) via MS Teams or WhatsApp/ Instagram
Weeks 10–12 Cumulative video project	Collaboration and product creation	Participate in a large group call on Teams. Complete a Blooket.com quiz. Discuss how your group's SDG can be applied to the lives of other students. Create a 5-minute video with your group synchronously and asynchronously.	Group communication (with 6–7 members) via MS Teams or WhatsApp/ Instagram

Post-COIL, each class connected the COIL project with their individual course objectives. The Japanese students reflected on the experience in class conversations and written assignments. SDGs remained a conversation topic for the duration of the course. For more details on the COIL design and Japanese reflection tasks, see Remmerswaal (2024). The prospective teachers taking Critical Reading and Writing in Türkiye reflected on the COIL group experience in-class, discussing the similarities of group work with classroom management. The students in Spain reflected on the experience during class and considered the financial impact and cost of meeting their proposed methods for meeting the SDGs in their community.

Communication tools

The COIL project was managed primarily through Microsoft Teams. The first researcher created a COIL Team in his institutional instance of Microsoft Teams and invited all participants to join. He assigned all students to one of 27 channels, allowing them to communicate directly with their group members via video call or written text. All three instructors were members of each channel, allowing them to view conversations and meetings and intervene if necessary. Students could also message any of the instructors directly through Teams. The assignments tab within Teams hosted the assignments. The General channel facilitated announcements, supplemental instructions, and COIL-wide video calls. During the first and subsequent assignments, a few groups broke away

from Teams to use other video calling and messaging apps such as WhatsApp and Instagram.

Measurement tools

Self-Perceptions of Communication Competence (SPCC) scale. This study uses the 12 items used by Hashimoto (2002) to measure SPCC in 12 situations, divided into four group sizes (public speaking with 30 people, talking in meetings with 10 people, talking in small groups with five people, and talking in dyads) and three types of participants (strangers, acquaintances, and friends) in which participants rate their perceived communication competence from 0–100. Similar to Ghani and Azhar (2017), this study uses a 5-point Likert scale with written descriptions to remove the ambiguity of the 100-point scale used by Hashimoto (2002). The Likert statements ranged from 1 – *I would never or almost never feel competent* to 5 – *I would always or almost always feel competent* (see the questions in Appendix A).

Communication Anxiety (CA) scale. The CA scale, also from Hashimoto (2002), is a modification of the SPCC scale, where students rank the same 12 situations on how likely they were to feel anxious rather than competent. Once again, we used a Likert scale instead of the numerical scale from 0–100 used by Yashima (2002). The Likert items ranged from 1 – *I would never or almost never feel nervous;* to 5 – I would always or almost always feel nervous (Appendix A).

Soft Skills Self-evaluation Questionnaire (3SQ). The soft skills self-evaluation questionnaire (3SQ) focuses on 10 soft skills: autonomy, collaboration, commitment, curiosity, empathy, leadership, openness, problem-solving, resilience, and self-confidence (Lucisano & du Mérac, 2019). Each skill has four to five items, for a total of 41 items (see Appendix A). Lucisano and du Mérac (2019) examined it for construct validity, reliability, internal consistency, and confirmatory factor analysis with high school students in Italy. Scippo and du Mérac (2021) further validated the items of self-confidence, autonomy, problem-solving, cooperation, and empathy by finding statistically significant correlations between the 3SQ and existing tools measuring the same constructs. This study left the Likert items untouched as a 5-point Likert scale from 1 – *Never or almost never true;* to 5 – *Always or almost always true.*

Design & procedure

The current study adopted a mixed methods convergent design where quantitative and qualitative data were analyzed separately and merged for analysis (Creswell & Clark, 2017). This quantitative data provides insight into overall student trends, while qualitative data provides insights into those trends. The quantitative data comes from pre- and post-COIL surveys containing previously described SPCC, CA, and soft skills scales. The qualitative data was gathered from open-ended question responses in the post-COIL survey. Our convergence

analysis is inspired by the β- and σ-convergence (Barro & Sala-i-Martin, 1992). *T*-tests consider average changes across students, while convergence analysis considers the changes of each individual. We analyze whether the dispersion of SPCC, CA, and soft skills decreased (σ-convergence) or increased and whether those students with lower competence and higher anxiety caught up with those with higher competence and lower anxiety (β-convergence).

Figure 1 illustrates the research design and the interaction between statistical techniques used for quantitative and qualitative analysis.

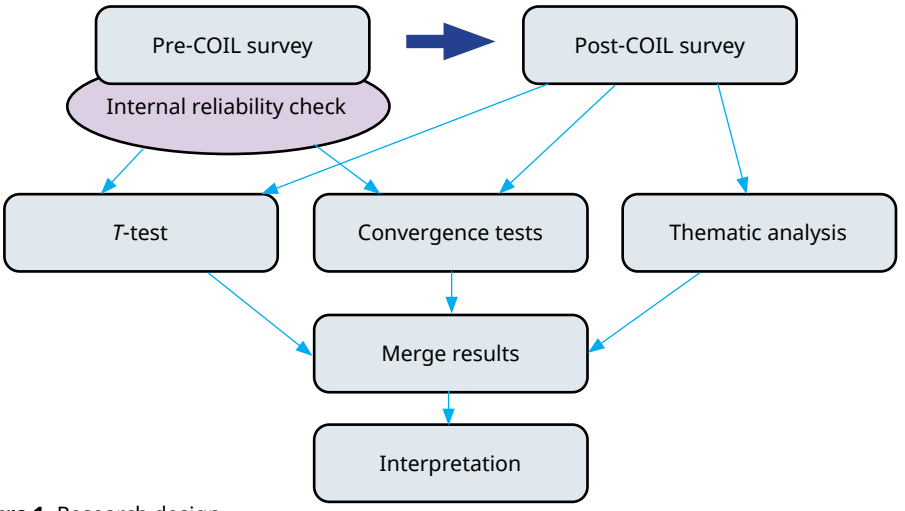


Figure 1. Research design

The pre-COIL survey data was tested for internal reliability and compared to the post-COIL data using a paired samples *t*-test. Statistically significant internal reliability indicates that students interpreted the survey items as related. A statistically significant change in a *t*-test provides evidence that a change occurred. We then calculated the β-convergence and σ-convergence of each scale. Convergence analysis provides more details as to which type of student is the most likely to change and in which direction. For the qualitative data, Japanese and Spanish students' responses were translated as necessary into English via DeepL.com and reviewed by the authors for accuracy, while the Turkish students responded in English. The authors then analyzed the responses inductively via Thematic Analysis (Braun & Clarke, 2006). All three authors coded and created themes for the responses and reached a group consensus when there were discrepancies. The authors reflected on the quotations within each category and chose representative quotes according to the thematic analysis. The authors then interpreted the *t*-test and convergence results with the themes that emerged through the perspective of SCT, considering students' previous experiences, the COIL as a mastery experience, vicarious experiences, social persuasion, and physiological and emotional states.

Because of the depth and breadth of data collected in this study, feedback from the participants, their artifacts related to SDGs, and their written communication may be the subject of future studies.

Results

The Cronbach's α for all three scales had high internal reliability. The SPCC point estimate had a Cronbach's α score of 0.945 overall (see Appendix B). The student CA had Cronbach's α point estimate of 0.936 (see Appendix C). The soft skills' Cronbach's α point estimate was 0.906 (see Appendix D). With this level of internal reliability, the *t*-tests could proceed. Statistical significance in this study is $p < 0.05$, and marginal significance is $p < 0.1$. In a study of this size, marginal significance may warrant further review (Otte et al., 2022).

Self-Perceptions of Communication Competence (SPCC)

In all seven conversation sizes and familiarity settings, SPCC increased (see Figure 2). *Average*, *Acquaintances*, and *Public* had statistical significance ($p < 0.05$, see Appendix E). *Small Group* and *Friend* had a marginally significant increase. By country, Türkiye had a marginally significant increase in *Average*, whilst Spain had a significant increase in *Public* and marginally significant increase in *Acquaintance* (see Appendix E). The averages shown in Figure 2 indicate students' post-COIL exceeded 3 – *I would occasionally feel competent*.

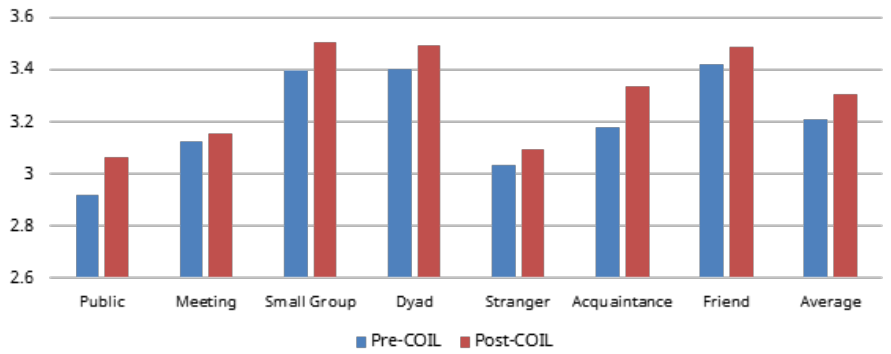
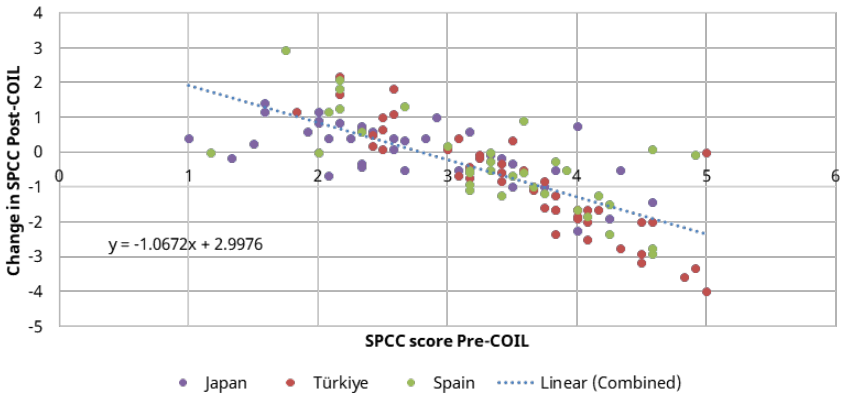


Figure 2. SPCC pre-COIL and post-COIL

Figure 3 demonstrates the difference in SPCC post-COIL compared to initial SPCC levels. The *x*-axis represents the pre-COIL SPCC of each student, and the *y*-axis is the change that occurred, calculated as post-COIL SPCC minus pre-COIL SPCC. Any increases in SPCC are shown as a positive *y* value, and any decreases are a negative *y* value. Those who began with the lowest SPCC are the most likely to have increased their SPCC. The *x*-intercept is 2.8 (near I would occasionally feel competent), which indicates that students who began with SPCC above 2.8 were the most likely to see a decrease in their SPCC post-COIL. Within Figure 3, students increased and decreased on both sides of the *x*-intercept. However, a clear trend emerged in the data, which is represented by the line of best fit labeled Linear. The negative slope of the line of best fit demonstrates the existence of β-convergence (Barro & Sala-i-Martin, 1992). Moreover, we observe σ-convergence because the standard deviation fell from 0.945 to 0.858,

which indicates a fall of dispersion across students' SPCC. Hence, students converged toward the center value.



Note: Linear is a line that best fits all data.

Figure 3. Average SPCC β-convergence

As Table 2 indicates, most students (n = 88, 73.9%) responded positively when asked if they felt more competent in any of the SPCC situations post-COIL. Some students referred directly to the SPCC items, referring to strangers or friends, while others replied with more general reasons for perceiving a change in these SPCC scenarios. The three most frequent (f) reasons students stated for that were related to communication skills (f = 27), self-confidence (f = 26), and speaking skills (f = 25). One representative quote for increased communication skills is, "I felt my communication skills improved when I was able to express my opinions to the members more and more. When I first met them, I could not talk much and just listened, so I could feel their growth even more" (JP181, Japanese participant randomly assigned 181). Self-confidence was often related to leaving their fears behind and trying to communicate: "A little more, since COIL forces you to always use English, and that makes you feel safer and more confident in using it. Besides, it also helps you to see that everyone has their fears and their mistakes speaking English, so my own errors are relatively less important" (SPx2, Spanish participant randomly assigned x2). Within this COIL, speaking skills were also increased: "Because I realized that people will not judge me on my speaking skills, and realized that I'm competent in speaking skills" (TP49, Turkish participant randomly assigned 49).

While the reasons that SPCC did not increase tended to be related to more external reasons for Turkish and Spanish students, those reasons tended to be internally related for Japanese students:

> JP52: I rarely speak up and mostly just listen to what is being said.
>
> TP5: I don't think I have gained that much competence in terms of speaking in a public or meeting since the project was online and it was mostly based on the assignments, and we didn't talk to each other with cameras on so much but just texting via Teams.

SPc1: Since we've communicated more by WhatsApp, we've not done a lot of video calls to speak.

Table 2. Reasons for the presence/absence of an increase in SPCC

SPCC increased	**Japan** ***n*** **= 28**	**Türkiye** ***n*** **= 36**	**Spain** ***n*** **= 24**	**Total** ***n*** **= 88**
Communication skills	14	8	5	27
Self-confidence		11	15	26
Speaking skills	13	4	8	25
Especially for talking with strangers	3	14	1	18
Completing tasks with people who think differently	2	1		3
Listening skills	3			3
Broadening perspective		1	1	2
Writing skills	1		1	2
Problem-solving skills		1		1
Especially for talking with friends		1		1
Learning the importance of cooperation and communication	1			1
No reason	1	3	3	7
Total	38	44	34	116
SPCC did not increase	***n*** **= 10**	***n*** **= 12**	***n*** **= 9**	***n*** **= 31**
Few synchronous meetings		3	2	5
Difficulty speaking English	5			5
Communicating with strangers		3		3
Mostly texted via Teams/WhatsApp		2	1	3
Already competent		1	2	3
Poor communication skills	3			3
Assignment focused		1		1
Cameras were not on		1		1
Not face-to-face		1		1
Group communication/cooperation problems		1		1
Feeling competent is personal		1		1
Difficulty understanding the other's English	1			1
Limited time	1			1
No reason		2	5	7
Total	10	16	10	36

Note: n = the number of students who commented in a section. Some students provided more than one reason, so the total number of comments exceeds the total number of participants.

Communication Anxiety (CA)

Figure 4 shows both increases and decreases in anxiety. Decreases were statistically significant in the whole sample for *Strangers* ($p < 0.05$, see Appendix F). There was a marginally significant decrease in Japanese participants for *Strangers* ($p < 0.1$). Türkiye saw a significant reduction in anxiety for

Acquaintance and a marginally significant reduction in *Meeting*. Spain saw no significant changes (see Appendix F). The average was slightly below 3 – *I would occasionally feel nervous*.

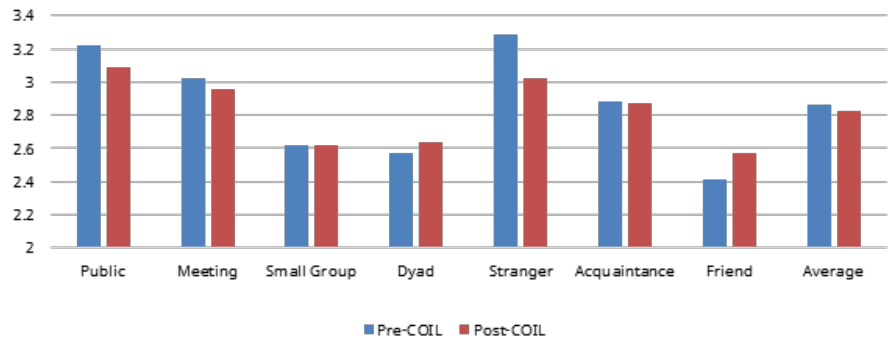
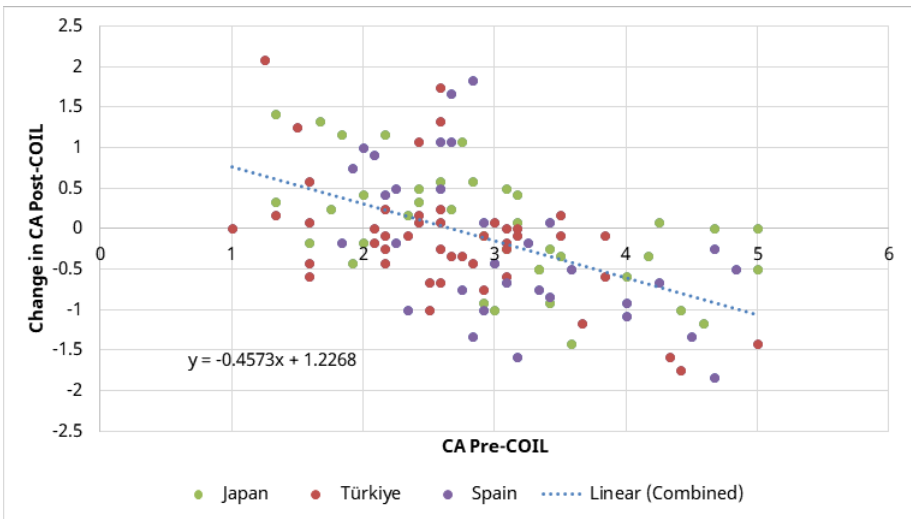


Figure 4. CA pre-COIL and post-COIL

In Figure 5, the negative trendline indicates that those who started with low CA were more likely to increase their anxiety. Those students who began with CA above 2.68 (near *I would occasionally feel nervous*) were likely to see a decrease in their anxiety. The negative trendline indicates a β-convergence (Barro & Sala-i-Martin, 1992). Again, the standard deviation fell from 0.933 to 0.861, indicating a σ-convergence and shows a lower dispersion across students' anxiety speaking English. Students who began as anxious were likely to become less anxious, and those who felt little anxiety were likely to realize they were slightly anxious.



Note: Linear is a line that best fits all data.

Figure 5. Average CA β-convergence

Table 3. Reasons for the presence/absence of a decrease in CA

CA decreased	**Japan n = 18**	**Türkiye n = 31**	**Spain n = 22**	**Total n = 71**
Communication via COIL tasks	7	19	5	31
Especially within a small group of strangers		7	11	18
Realized my potential/gained self-confidence	4	2	6	12
Group members were candid and understanding	2	3		5
Especially with acquaintances	3	2		5
Others did not have perfect English either		2	1	3
Social anxiety decreased		2		2
Communicated mostly in writing		1		1
Others were also nervous		1		1
Others understood us despite mistakes		1		1
Realized the enjoyment of speaking in English	1			1
No reason	4		1	5
Total	21	40	24	85
CA did not decrease	***n* = 20**	***n* = 17**	***n* = 12**	***n* = 49**
Still feeling nervous	11	5	5	21
Not being a nervous person	3	3	2	8
Not having many chances to communicate	1	2		3
First time communicating with foreign people	2			2
Did nothing that would lead to change	1			1
No reason	3	7	5	15
Total	21	17	12	50

Note: n = the number of students who commented in a section. Some students provided more than one reason.

Table 3 demonstrates that most students (n = 71, 59.2%) responded that their CA decreased. Some students referred to the types of interactions from the CA survey directly, such as within a small group of strangers. Others provided general reasons that CA had not changed in any of the situations, such as a general sense of nervousness. The main reasons for a change were due to communication in the various tasks of the COIL (f = 31), "Because we had to communicate with our group members in order to make ideas and complete the assignment and eventually I am now less nervous than before about the situations above" (TP28). Students followed this, referring to their increased comfort in speaking with a small group of strangers (f = 18), similar to the setup of the COIL groups. The following quotation can exemplify this: "Probably COIL has helped me to improve my nerves in expressing successfully in English with strangers" (SPb2). The third most claimed reason was an increase in confidence or self-realization (f = 12), supported by this quotation, "Through the activity, I felt that it is not difficult to present myself. I realized that I used to be nervous because I was thinking too much and that I wouldn't be so nervous if I took action right away" (JP181). The other participants who saw no decrease (n = 49, 40.8%) related that situation to still being nervous or not being nervous in general.

This is reflected in quotes such as, “I couldn’t get rid of the nervousness even after many meetings” (JP51) and “I don’t feel any less nervous as I explained we didn’t talk much” (TP27).

Soft skills

Students answered the 3SQ scale ranging from 1 – *Never or almost never true;* to 5 – *Always or almost always true.* Students saw increases across all soft skills (see Figure 6). *Self-confidence* had the only statistically significant increase ($p < 0.05$), while *Leadership* and *Resilience* had marginally significant increases ($p < 0.1$; see Appendix G). Interestingly, these three skills began and remained in the bottom four skills for students. By country, Japan saw a statistically significant increase in *Self-confidence* while Türkiye saw a statistically significant rise in *Openness* and *Commitment*. Spain experienced a marginally significant increase in *Curiosity* and Türkiye in *Autonomy* and *Empathy*.

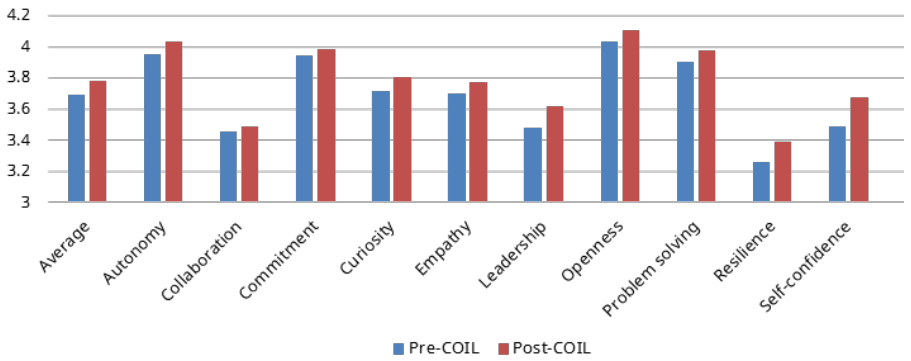
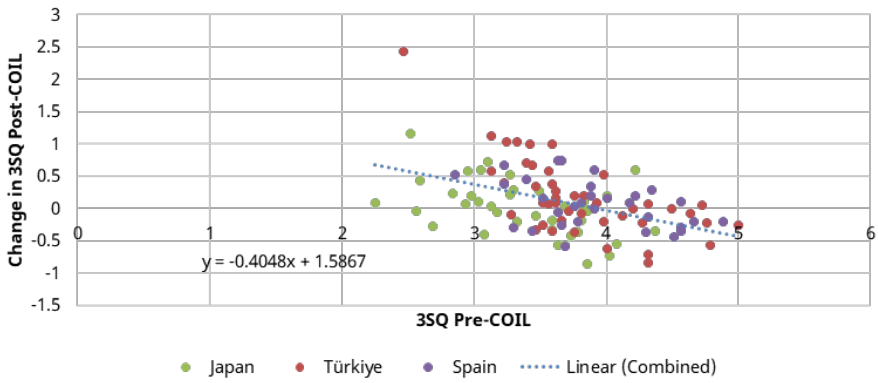


Figure 6. 3SQ pre-COIL and post-COIL

In Figure 7, the β-convergence shows that students who began with a low perception of their soft skills were more likely to see an increase in their skills. In this case, students below 3.92 (near *Usually true*) were likely to have their perception increase in their soft skills. Those above 3.92 were most likely to see a decrease in their perceived soft skills. The dispersion across students’ perception of their soft skills fell from 0.554 to 0.543, again indicating σ-convergence.



Note: Linear is a line that best fits all data.

Figure 7. Average 3SQ β-convergence

Table 4 summarizes the open-ended question asking students if they felt some of the statements became more (or less) true about themselves. Collaboration (*f* = 28), commitment (*f* = 20), and leadership (*f* = 12) were the top three skills students felt increased. The following quotations make this evident:

> JP41: I feel that I am now able to cooperate with others in activities more than before.
>
> TP20: I consider that I try to do my best and similarly we conducted good group activities. We completed our tasks before deadline and we, each person in our group, took responsibility and we worked collaboratively to achieve the desired result.
>
> SPx2: It has helped me to know how to organize work within a group of people and, above all, to work as a team for the same objective.

On the other hand, a few participants felt there was a negative change, especially for leadership (*f* = 8), as in, "I realised that I cannot assign responsibilities in the group, I need to be told what to do from others" (TP18). Some students reported a lack of change, which Japanese students attributed to personal weaknesses, but most did not provide a reason (*f* = 16).

Table 4. Reasons for increase/decrease/no change in perceptions of soft skills

Increase in perceptions		**J** *n* = 24	**T** *n* = 32	**S** *n* = 22	**Total** *n* = 78	**Soft skill subtotal**
Autonomy	Managed time more efficiently		2		2	
	Made decisions for themselves		1	1	2	4
Collaboration	Learned to complete tasks as a team	12	5	4	21	
	Became aware of their skill in group work		3	2	5	
	Learned how to divide group work		1	1	2	28
Commitment	Increased sense of responsibility	2	6	2	10	
	Did their best work	1	3		4	
	Determined to complete the task	2	3	1	6	20
Curiosity	Explored topic in depth/actively sought new information	1	2		3	3
Empathy	Imagined themselves in the situation of a group member	2	2	2	6	
	Involved with the outcome of others	1	1		2	8
Leadership	Realized or improved leadership qualities		6	5	11	
	Assigned group responsibilities			1	1	12
Openness	Considered others' opinions	1	4		5	
	Valued others' contributions		1		1	6
Problem solving	Found different solutions/ problem-solved	4	4		8	
	Able to manage disagreements	1	2		3	11
Resilience	Remained calm/patient in crisis		4	1	5	
	Remained positive in difficulty	1	2		3	8
Self-confidence	Increased self-awareness		3		3	
	Better able to express themselves	1	2		3	
	Increased self-confidence	1			1	7
All soft skills	Grew in all 3SQ statements	3	2		5	5
General growth	Expanded horizon or point of view		1	2	3	
	Due to the various tasks	1	1		2	
	Due to teacher feedback		1		1	6
	No reason		3	7	10	10
Total increase		34	65	29	128	128
Decrease in perceptions		***n* = 1**	***n* = 6**	***n* = 0**	***n* = 7**	**Soft skill subtotal**
Collaboration	Decreased willingness to participate in group activities		2		2	2
Leadership	Difficulty in coordinating the group	1	4		5	
	Preferred others assign tasks	2	1		3	8
Resilience	Difficulty remaining calm in crisis		2		2	2
Total decrease		3	9	0	12	12
Perceptions both increased and decreased		***n* = 1**	***n* = 2**	***n* = 0**	**Total *n* = 3**	

Perceptions did not change	*n* = 12	*n* = 7	*n* = 8	Total *n* = 27
Too few meetings		1		1
Personal weaknesses	4			4
Lack of personal initiative	2			2
My low English level	1			1
Personal lack of enthusiasm	1			1
Prior experiences with group work			1	1
Rigid personality			1	1
No reason	4	6	6	16
Total no change	12	7	8	27

Note: J = Japanese students. T = Turkish students. S = Spanish students. n = the number of students who commented in a section. Some students provided more than one reason. These students were counted in the category "Perceptions both increased and decreased," but those comments were tallied in their respective themes.

Discussion

Several aspects of SPCC, including the average score, had statistically significant increases, but CA only saw a statistically significant change in *Stranger* and not in the average. This expands on the results of Rahimi and Fathi (2022) and Zhou (2023). In those two studies, WTC increased post-VE, and this study found that SPCC, a WTC construct, changed post-VE. In Yashima's (2002) and Hashimoto's (2002) models of WTC, SPCC had a stronger correlation with WTC than CA. This suggests that despite no statistically significant change in average CA, WTC likely increased for students in this study.

The COIL appears to have its greatest effect on students who begin with a low or high SPCC, CA, or soft skills score, as demonstrated by the β-convergences. Both β-convergences and σ-convergences demonstrated that student perceptions were moving closer to the middle, with fewer outliers. The effect of the COIL in reducing outliers in SPCC, CA, and soft skills is explored below. Changes can be classified under three groups: students who increased, remained relatively the same, or decreased post-COIL. We interpreted these changes through a SE lens, as Bandura's (1997) SCT describes.

Students who saw a substantial increase in communication confidence (SPCC and reverse scored CA) and soft skills (a change in that trait of 0.5 or higher) tended to begin with a lower initial score, though a few students started in the middle and on the high end of the scale. These students may have attributed success in the COIL to personal success due to their efforts (Bandura, 1997). It is possible that these students felt they used communication confidence or soft skills at a higher level than they had previously experienced, thus increasing their perceptions (Bandura, 1997). SPCC and CA had several situations, such as speaking to 30 strangers, which only a few students experienced during the COIL. Yet, *Public* (speaking with 30 people) statistically significantly increased for SPCC. This increase could be due to vicarious experiences, where students could now imagine themselves doing these actions due to observing others in the large group call or other group interactions (Bandura, 1997; Wyatt, 2022).

The Spanish and Turkish students often linked increased communication confidence to increased self-confidence in their English, aligning with literature highlighting VEs's potential to build English communication confidence (Mestre-Segarra & Ruiz-Garrido, 2022; O'Dowd, 2021). The Japanese students linked this SPCC increase to their increased communication or speaking skills. Students whose communication confidence increased are more likely to participate in future communication opportunities more frequently (Balouchi & Samad, 2021; Clément et al., 1980; MacIntyre & Charos, 1996; McCroskey & McCroskey, 1988; Samvati & Golaghaei, 2017).

The students with little to no change (a change in that trait below 0.5 but above −0.5) tended to be close to the average scores pre-COIL. These students may not have perceived themselves using the traits the surveys measured or did not attribute group success to their effort (Bandura, 1997). Some students, such as in the case of CA, reported no change due to *not being an anxious person,* or *always being anxious*. Remmerswaal (2024) observed some Japanese students receiving excessive assistance from teammates, which may have also occurred between Turkish and Spanish students. Students receiving this assistance would likely attribute success to external factors, not their contributions, and not see a change in their perceptions of ability (Bandura, 1997; Wyatt, 2022). Similarly, suppose one group member took a leadership role for the duration of the VE. In that case, other students may not have had an opportunity to expand their leadership capacity and would be unlikely to report a change in *Leadership*. Most commonly, Japanese students attributed a lack of change in soft skills to various internal factors or weaknesses. In contrast, the other two groups attributed it to external factors beyond their control. These may result from cultural differences, which may have played a role in many aspects of the COIL project.

The students who saw a decrease in communication confidence or soft skills (a change below −0.5) had a reduction in their SE or beliefs in their abilities. The strength of these students' beliefs pre-COIL largely determines whether the COIL was a negative experience or a learning experience for these students. A decrease indicates that students perceived their ability during the COIL as falling short of their expectations. If students based those expectations on many rich experiences, a decrease would require a substantial failure in the student's mind (Bandura, 1997). However, if a student was basing their expectations on limited experience or little reflection on their previous experiences, a change in perceptions would be relatively easy, as these perceptions were still forming (Bandura, 1997).

It may be the case that most students had never considered their communication confidence and soft skills pre-COIL, and the surveys caused self-reflection and self-evaluation, which led to critical thinking and acquiring a new perspective (Falchikov & Boud, 1989; Saleh, 2019). Several Japanese students indicated that the entire experience was a unique opportunity (Remmerswaal, 2024). Furthermore, anxious language learners tend to underestimate, and relaxed students overestimate their fluency in the second language when compared to the evaluations of independent, bilingual raters (MacIntyre et al.,

1997). Student anxiety levels pre- and post-COIL likely played a factor in their reported SPCC.

There were several forms of disconnect between the qualitative and quantitative data, especially in soft skills. It appears that students did not perceive themselves to be using some of the soft skills despite the instructors observing their use, similar to the findings of Llorens et al. (2024). Birchley (2013) argues that students should be allowed to develop their understanding of soft skills through constructivist theories and experiential learning rather than a teacher-centric approach. The COIL in our study provided a basis for the students to develop their understanding through first-hand experiences. However, having no introduction to soft skills from the instructors may have led to a lack of awareness of the skills they used. For example, *Leadership* had a marginally significant increase ($p < 0.1$) and the third-highest number of positive comments. Yet, it was the skill most mentioned as decreasing due to the COIL project. Reasons for a decrease included a lack of success in a leadership role and a new-found preference to follow a leader.

Another disconnect was in *Self-confidence,* which was the only statistically significant increase ($p < 0.05$), but it was the seventh most-referenced skill when students discussed an improvement in soft skills. Mestre-Segarra and Ruiz-Garrido (2022) also found a disconnect between soft skills and their graduate students. The instructors observed the students using specific skills, such as negotiation, yet students reported not using those skills in their reflections. Considering that example, the disconnect observed in our students may have been a lack of awareness of the soft skills they used in this COIL. Whether students realized an increase in self-confidence or not, the growth seen in the average students' *Self-confidence* score indicates students may be more confident participants in future group projects.

Stranger was an important theme that emerged from this project. CA saw a statistically significant decrease in *Stranger* as a whole group and marginally significant decrease with Japanese students. Turkish and Spanish students reported speaking in small groups of strangers as a reason for a decreased CA. Strangers played a meaningful role in students' attribution of SPCC increases, lack of increase, and CA decreases. Likely, this frequent reference to strangers refers to their group members. Students in this study provided reasons for the decreased anxiety with other non-native speakers such as witnessing the mistakes of others, which helped put the 'fear of evaluation' aside. Lee (2004) and Satar and Özdener (2008) discuss similar benefits when speakers use English with non-native speakers. As time passed, the students became accustomed to each other, and some began to reimagine their group members as acquaintances rather than strangers, leading to decreased anxiety levels. Out of 11 reasons for CA decreases (see Table 3), six of them relate to group members: *Especially within a small group of strangers, Group members were candid and understanding, Especially with acquaintances, Others did not have perfect English either, Others were also nervous,* and *Others understood us despite mistakes.* Students may not have seen a decrease in their overall CA, but with this

understanding of *Stranger*, they may feel less anxious participating in a similar project.

Several groups gravitated towards text-based communication channels like Teams, Instagram, or WhatsApp rather than video-based interactions. While this could have been a choice of convenience due to time differences, part-time jobs, or other responsibilities, it may also have resulted from a low communication confidence. The SPCC and CA scales measured situations implying verbal communication, and the literature indicates that speaking a foreign language induces more anxiety than writing, often stemming from fear of mistakes, evaluation, and low self-confidence (Cheng et al., 1999; Jugo, 2020). These groups tended to have at least one person, often Japanese, who began with low SPCC and high CA, indicating low communication confidence (Yashima, 2002). The extent to which these students influenced group communication methods is unknown. However, these students tended to report a lack of change in communication confidence due to their poor English proficiency and feelings of nervousness (see Tables 2 and 3).

The move to text-based communication was not discouraged, though it was unforeseen. Students had explicit instructions, such as the questions to answer in a report and due dates. However, students had agency to interpret the questions, make their own schedules, use any communication tool, and use any media medium they felt would best help them reach the project goals. Providing agency for students is a way to create a positive learning environment, and Yüce (2023) found agency to have a positive relationship with student SE and SPCC. Increases in these two constructs, in turn, have been seen to increase autonomous learning, academic success, and WTC (Yüce, 2023).

The move to written-based communication led some of their counterparts, mainly Turkish students, to express frustration as they desired synchronous meetings with cameras on, believing this would improve their communication skills. This elucidation strongly aligns with another study in which a few Japanese students stopped participating in synchronous VE sessions because of the feeling that their English language skills were insufficient (Roarty et al., 2023). Despite avoiding video interactions, these groups completed their work successfully and on time. Ultimately, their teamwork and effort led to their success.

Overall, this COIL project resulted in a diverse number of outcomes. The β-convergences and σ-convergences suggest that the experience helped normalize students' perceptions of their abilities, with those at both extremes moving to more moderate self-assessments. This change aligns with SCT's mastery experiences and vicarious learning, as students had opportunities to practice communication skills and observe peers' successes and challenges (Bandura, 1997). Students' varied responses to the COIL experience, whether showing increased, decreased, or unchanged perceptions, can be understood through their initial SE beliefs, actual experiences during the project, and their attribution of success or failure. While some students implied they had little experience with VEs and collaborating with international peers, collecting this information from students prior to a VE may give instructors insight into the

strength of students' SE beliefs. Instructors can then consider students' initial SPCC, CA, and soft skill levels in consideration of their previous experience to provide appropriate scaffolding and support positive development throughout the COIL.

Conclusion

This study investigated the impact of an eight-week Collaborative Online International Learning (COIL) project on university students from Japan, Türkiye, and Spain, focusing on self-perceived communication competence (SPCC), communication anxiety (CA), and soft skills development. The *t*-test results indicated statistically significant increases in several SPCC areas, including average SPCC scores. Although the average CA decrease and soft skills increase were not statistically significant, there was a notable reduction in anxiety when interacting with strangers. Additionally, self-confidence saw a statistically significant rise, suggesting enhanced teamwork capabilities.

Convergence analysis revealed that students became more aware of their abilities through self-evaluation prompted by the project activities. Qualitative data further supported these findings, with many students attributing improvements to opportunities for authentic English communication in small multicultural groups and increased self-confidence. The project's structure, which incorporated synchronous and asynchronous communication, accommodated various comfort levels and learning preferences. Some students, particularly those with initially low SPCC and high CA, preferred text-based communication and reported minimal changes in their perceptions. This underscores the importance of offering diverse communication channels to meet student needs and comfort levels in COIL projects.

Future research should explore different strategies or COIL designs to better support students with persistently low SPCC and high CA and their preferred communication channels. Additionally, investigating the long-term effects of COIL on SPCC, CA, and soft skills through longitudinal studies could provide deeper insights. Replicating this study in various cultural and educational contexts would help generalize the findings. At the same time, interviews with students could offer a more in-depth understanding of their perceptions of these traits and the other participants. For instructors considering a COIL project, we recommend supporting a primary communication channel while allowing multiple channels to cater to student needs and comfort levels, fostering a sense of ownership.

Ultimately, COIL projects have the potential to transform language learning by fostering essential communication skills in an increasingly interconnected world. As graduates increasingly require socio-emotional capacities alongside language proficiency, initiatives blending intercultural collaboration and learning hold significant potential.

Acknowledgments

The authors thank Sojo University for facilitating group interactions through their institutional Microsoft Teams account.

Funding. This work was supported by the R&D&R programme of the Spanish Government [grant number PID2020-113452RB-I00]; and the Generalitat de Catalunya [grant number 2021 SGR 00570].

References

Arnold, J., & Brown, D. (1999). A map of the terrain. In J. Arnold (Ed.), *Affect in language learning* (pp. 1–24). Cambridge: Cambridge University Press.

Balouchi, S., & Samad, A.A. (2021). The effect of perceived competence on second language communication frequency: The mediating roles of motivation, willingness to communicate, and international posture. *Education and Information Technologies, 26,* 5917–5937. https://doi.org/10.1007/s10639-021-10579-z

Bandura, A. (1997). The nature and structure of self-efficacy. *In Self-efficacy: The exercise of control* (pp. 36–78). New York: W.H. Freeman and Company.

Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology, 52*(1), 1–26. https://doi.org/10.1146/annurev.psych.52.1.1

Bandura, A. (2006). Guide for constructing self-efficacy scales. In F. Pajares, & T. Urdan (Eds.), *Adolescence and education: Vol. 5. Self-efficacy and adolescence* (pp. 307–337). Charlotte: Information Age.

Barro, R. J., & Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy, 100*(2), 223–251. https://doi.org/10.1086/261816

Birchley, S. L. (2013). Japanese students' perceptions of soft skills: Teaching professional skills & CLIL. *Bulletin of Toyo Gakuen University, 21,* 187–207. https://doi.org/10.24547/00000205

Blake, R. J. (2011). Current trends in online language learning. *Annual Review of Applied Linguistics, 31,* 19–35. https://doi.org/10.1017/S026719051100002X

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research Psychology, 3,* 1–41. https://doi.org/10.1191/1478088706qp063oa

Canals, L. (2020). The effects of virtual exchanges on oral skills and motivation. *Language Learning & Technology, 24*(3), 103–119. http://hdl.handle.net/10125/44742

Canals, L. (2023). Modified output and metalanguage during conversational interaction: A qualitative look at interactional feedback. *Studies in Second Language Learning and Teaching, 13*(3), 601–626. https://doi.org/10.14746/ssllt.31128

Cheng, Y. S., Horwitz, E. K., & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language Learning, 49*(3), 417-446. https://doi.org/10.1111/0023-8333.00095

Clément, R., Gardner, R. C., & Smythe, P. C. (1980). Social and individual factors in second language acquisition. *Canadian Journal of Behavioural Science, 12*(4), 293–302. https://doi.org/10.1037/h0081081

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research.* New York: Sage publications.

Croucher, S. M., Kelly, S., Rahmani, D., Burkey, M., Subanaliev, T., Galy-Badenas, F., Lando, A. L., Chibita, M., Nyiransabimana, V., Turdubaeva, E., Eskiçorapçı, N., & Jackson, K. (2020). A multinational validity analysis of the self-perceived communication competence scale. *Journal of International and Intercultural Communication, 13*(1), 1–12. https://doi.org/10.1080/17513057.2019.1569250

Ellis, R. (2017). Position paper: Moving task-based language teaching forward. *Language Teaching, 50*(4), 507–526. https://doi.org/10.1017/S0261444817000179

European Commission. (2011). *Transferability of skills across economic sectors: Role and importance for employment at European level.* Luxembourg: Publications Office of the European Union. https://pjp-eu.coe.int/bih-higher-education/images/eul14180_tos_110924_web_with%20erratum-3.pdf

Falchikov, N., & Boud, D. (1989). Student self-assessment in higher education: A meta-analysis. *Review of Educational Research, 59*(4), 395–430. https://doi.org/10.3102/00346543059004395

Fushino, K. (2010). Causal relationships between communication confidence, beliefs about group work, and willingness to communicate in foreign language group work. *TESOL Quarterly, 44*(4), 700-724. https://doi.org/10.5054/tq.2010.235993

Ghani, M., & Azhar, S. W. (2017). Effect of motivation, willingness to communicate (WTC), self perceived communicative competence (SPCC) and L2 anxiety on the frequency of use of English as L2. *Journal of Educational Research, 20*(1), 157–171. https://www.prdb.pk/article/effect-of-motivation-willingness-to-communicate-wtc-self-7342

Gillies, R. M. (2016). Cooperative learning: Review of research and practice. *Australian Journal of Teacher Education, 41*(3). https://doi.org/10.14221/ajte.2016v41n3.3

Hackett, S., Janssen, J., Beach, P., Perreault, M., Beelen, J., & van Tartwijk, J. (2023). The effectiveness of Collaborative Online International Learning (COIL) on intercultural competence development in higher education. *International Journal of Educational Technology in Higher Education, 20*(5). https://doi.org/10.1186/s41239-022-00373-3

Háhn, J., & Radke, K. (2020). Combining expertise from linguistics and tourism: a tale of two cities. In F. Helm, & A. Beaven (Eds.), *Designing and implementing virtual exchange: A collection of case studies* (pp. 11–22). Research-publishing.net. https://doi.org/10.14705/rpnet.2020.45.1111

Hashimoto, Y. (2002). Motivation and willingness to communicate as predictors of reported L2 use: The Japanese ESL context. *Second Language Studies, 20*(2), 29–70. https://www.hawaii.edu/sls/wp-content/uploads/2014/09/Hashimoto.pdf

Jugo, R. R. (2020). Language anxiety in focus: The case of Filipino undergraduate teacher education learners. *Education Research International, 2020,* Article 7049837. https://doi.org/10.1155/2020/7049837

Kim, T. Y., & Choi, E. Y. (2023). English learning anxiety and English learning demotivation in South Korean third-year high school students. *Porta Linguarum An International Journal of Foreign Language Teaching and Learning, 39,* 75–91. https://doi.org/10.30827/portalin.vi39.22305

Lee, L. (2004). Learners' perspectives on networked collaborative interaction with native speakers of Spanish in the US. *Language Learning & Technology, 8*(1), 83–100. http://dx.doi.org/10125/25231

Leeming, P., Vitta, J.P., Hiver, P., Hicks, D., McLean, S., & Nicklin, C. (2024), Willingness to Communicate, Speaking Self-Efficacy, and Perceived Communicative Competence as predictors of second language spoken task production. *Language Learning.* https://doi.org/10.1111/lang.12640

Lewis, T., & O'Dowd, R. (2016). Online intercultural exchange and foreign language learning: A systematic review. In R. O'Dowd, & T. Lewis (Eds.), *Online Intercultural Exchange* (pp. 173–191). New York: Routledge. https://doi.org/10.4324/9781315678931

Llorens, A., Trullols, E., Pérez-Poch, A., & Petrović, N. (2024). Soft skills development in ICT students: an evaluation of teaching methods by university educators. *Cogent Education, 12*(1). https://doi.org/10.1080/2331186X.2024.2437906

Lockley, T. (2013). Exploring self-perceived communication competence in foreign language learning. *Studies in Second Language Learning and Teaching, 3*(2), 187–212. https://doi.org/10.14746/ssllt.2013.3.2.3

Long, M. H., & Porter, P. A. (1985). Group work, interlanguage talk, and second language acquisition. *TESOL Quarterly, 19*(2), 207–228. https://doi.org/10.2307/3586827

Lucisano, P., & du Mérac, E. R. (2019). Soft skills self-evaluation questionnaire (3sq), characteristics and psychometric properties. In P. Lucisano & A.M. Notti (Eds.), *Training actions and evaluation processes. Atti del convegno Internazionale SIRD* (pp. 609–622). Lecce: Pensa Multi Media Editore. https://www.pensamultimedia.it/libro/9788867606344

MacIntyre, P. D., & Charos, C. (1996). Personality, attitudes, and affect as predictors of second language communication. *Journal of Language and Social Psychology, 15*(1), 3–26. https://doi.org/10.1177/0261927X960151001

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K.A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *Modern Language Journal, 82*(4), 545–562.

MacIntyre, P. D., Noels, K. A., & Clément, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning, 47*(2), 265–287. https://doi.org/10.1111/0023-8333.81997008

McCroskey, J. C. (2009). Communication apprehension: What have we learned in the last four decades. *Human Communication, 12*(2), 157–171. https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=6c8abb55b4fbe0e7be285c46c2de534519c437dd

McCroskey, J. C., & McCroskey, L. L. (1988). Self-report as an approach to measuring communication competence. *Communication Research Reports, 5,* 108–113. https://doi.org/10.1080/08824098809359810

Mestre-Segarra, M. Á., & Ruiz-Garrido, M. F. (2022). Examining students' reflections on a collaborative online international learning project in an ICLHE context. *System, 105,* 102–714. https://doi.org/10.1016/j.system.2021.102714

Ministry of Education, Culture, Sports, Science and Technology. (2012). *Higher Education in Japan* (pp. 1–25). https://www.mext.go.jp/en/policy/education/highered/title03/detail03/__icsFiles/afieldfile/2012/06/19/1302653_1.pdf

O'Dowd, R., & Waire, P. (2009). Critical issues in telecollaborative task design. *Computer Assisted Language Learning, 22*(2), 173–188. https://doi.org/10.1080/09588220902778369

O'Dowd, R. (2021) What do students learn in virtual Exchange? A qualitative content analysis of learning outcomes across multiple exchanges. *International Journal of Educational Research, 109,* https://doi.org/10.1016/j.ijer.2021.101804

Organisation for Economic Co-operation and Development (2015). *Skills for social progress: The power of social and emotional skills.* Paris: OECD Publishing. https://doi.org/10.1787/9789264226159-en

Otte, W. M., Vinkers, C. H., Habets, P. C., Van Ijzendoorn, D. G., & Tijdink, J. K. (2022). Analysis of 567,758 randomized controlled trials published over 30 years reveals trends in phrases used to discuss results that do not reach statistical significance. *PLoS Biology, 20*(2), e3001562. https://doi.org/10.1371/journal.pbio.3001562

Rahimi, M., & Fathi, J. (2022). Employing e-tandem language learning method to enhance speaking skills and willingness to communicate: the case of EFL learners. *Computer Assisted Language Learning, 37*(4), 924–960. https://doi.org/10.1080/09588221.2022.2064512

Remmerswaal, R. (2024). SDGs and COIL in Japan: Reflections on design. *KOTESOL Proceedings 2023,* 85-97. https://www.koreatesol.org/sites/default/files/pdf_publications/KOTESOL.Proceedings.2023.pdf

Resnik, P., & Schallmoser, C. (2019). Enjoyment as a key to success? Links between e-tandem language learning and tertiary students' foreign language enjoyment. *Studies in Second Language Learning and Teaching, 9*(3), 541–564. https://doi.org/10.14746/ssllt.2019.9.3.6

Roarty, A., Tuncer, H., & Tang, L. (2023). Adding synchronous sessions to asynchronous virtual Exchange: Insights from the IVEProject. *Intercultural Communication Education, 6*(1), 19–32. https://doi.org/10.29140/ice.v6n1.760

Rubin, J. (2022). Preface to an evolving international educational landscape. In J. Rubin & S. Guth (Eds.), *The guide to COIL virtual exchange: Implementing, growing, and sustaining collaborative online international learning* (pp. 28–45). New York: Routledge. https://doi.org/10.4324/9781003447832

Saleh, S. E. (2019). Critical thinking as a 21st century skill: Conceptions, implementation and challenges in the EFL classroom. *European Journal of Foreign Language Teaching, 4*(1), 1–16. https://doi.org/10.5281/zenodo.2542838

Samvati, G. C., & Golaghaei, N. (2017). Iranian EFL learners' willingness to communicate, self-perceived communication competence, and communication apprehension in L1 and L2: A comparative study. *Journal of English Language Pedagogy and Practice, 10*(20), 50–71. https://jal.tabriz.iau.ir/article_532469.html

Satar, H. M., & Özdener, N. (2008). The effects of synchronous CMC on speaking proficiency and anxiety: Text versus voice chat. *The Modern Language Journal, 92*(4), 595–613. http://www.jstor.org/stable/25173104

Scippo, S., & du Mérac, E. R. (2021). Criterion validation of the scales of autonomy, collaboration, empathy, problem-solving and self-confidence of the 3SQ. Soft skills self-evaluation questionnaire adapted for lower secondary school. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal), 24,* 193–210. https://doi.org/10.7358/ecps-2021-024-scme

Sotillo, S. M. (2000). Discourse functions and syntactic complexity in synchronous and asynchronous communication. *Language Learning & Technology, 4*(1), 82–119. https://www.learntechlib.org/p/88512/

United Nations. (n.d.). *The 17 goals.* Department of Economic and Social Affairs – Sustainable Development. https://sdgs.un.org/goals

Wang, Y. (2004). Supporting synchronous distance language learning with desktop videoconferencing. *Language Learning & Technology, 8*(3), 90–121. http://dx.doi.org/10125/43997

Welzel, C., Brunkert, L., Kruse, S., & Inglehart, R. F. (2023). Non-invariance? An overstated problem with misconceived causes. *Sociological Methods & Research, 52*(3), 1368–1400. https://doi.org/10.1177/0049124121995521

Wicking, P. (2022). Learning-oriented assessment in an international virtual exchange. In S. M. Hilliker (Ed.). *Second language teaching and learning through virtual Exchange,* (pp. 9–33). Berlin: De Gruyter Mouton. https://doi.org/10.1515/9783110727364-002

Wyatt, M. (2022). Self-efficacy. In S. Li, P. Hiver, & M. Papi (Eds.), *The Routledge handbook of second language acquisition and individual differences* (pp. 207–219). New York: Routledge. https://doi.org/10.4324/9781003270546

Yanguas, Í. (2010). Oral computer-mediated interaction between L2 learners: It's about time. *Language Learning & Technology,14*(1), 72–93. http://dx.doi.org/10125/44227

Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal, 86,* 54–66. https://doi.org/10.1111/1540-4781.00136

Yüce, E. (2023). The role of learning climate on self-efficacy beliefs and self-perceived communication competences of EFL learners in higher education. *Porta Linguarum An International Journal of Foreign Language Teaching and Learning, 39,* 231–247. https://doi.org/10.30827/portalin.vi39.24106

Zhang, Q., Song, Y., & Zhao, C. (2024). Foreign language enjoyment and willingness to communicate: the mediating roles of communication confidence and motivation. *System,* 103346. https://doi.org/10.1016/j.system.2024.103346

Zhou, A. (2023). Investigating the impact of online language exchanges on second language speaking and willingness to communicate of Chinese EFL learners: A mixed methods study. *Frontiers in Psychology, 14,* Article 1177922. https://doi.org/10.3389/fpsyg.2023.1177922

Ziegler, N. (2016). Synchronous computer-mediated communication and interaction: A meta-analysis. *Studies in Second Language Acquisition, 38*(3), 553–586. https://doi.org/10.1017/S027226311500025X

Appendix A

Surveys and report questions

https://drive.google.com/drive/folders/1qlFEjKZDjU-uUrw1UKzOAG0-cAZhcEfl?usp=drive_link

Appendix B

SPCC internal reliability test

Cronbach's α estimate	Total	Japan	Türkiye	Spain
Point estimate	0.945	0.953	0.894	0.967
95% CI lower bound	0.925	0.919	0.826	0.944
95% CI upper bound	0.960	0.973	0.939	0.982

Appendix C

CA internal reliability test

Cronbach's α estimate	Total	Japan	Türkiye	Spain
Point estimate	0.936	0.950	0.911	0.958
95% CI lower bound	0.914	0.917	0.859	0.928
95% CI upper bound	0.953	0.972	0.947	0.977

Appendix D

3SQ internal reliability test

Cronbach's α estimate	Total	Japan	Türkiye	Spain
Point estimate	0.906	0.906	0.885	0.889
95% CI lower bound	0.877	0.850	0.824	0.815
95% CI upper bound	0.929	0.944	0.928	0.938

Appendix E

SPCC t-*test*

		Total *df* = 117				Japan *df* = 37			
Pre-COIL	**Post-COIL**	***T***	***Stat***	***z***	***p***	***T***	***Stat***	***z***	***p***
Stranger	Stranger	W	2251.5	−1.252	0.209	S	−.213		0.833
Acquaintance	Acquaintance	W	1885.0	−2.475	**0.013**	W	265.0	−1.068	0.287
Dyad	Dyad	W	2154.5	−1.274	0.203	S	−0.874		0.388
Public	Public	W	1789.5	−2.529	**0.011**	S	−0.534		0.597
Meeting	Meeting	W	1858.0	−0.930	0.352	S	−0.282		0.780
Small Group	Small Group	W	1623.5	−1.858	**0.063**	S	0.117		0.907
Friend	Friend	W	2137.0	−1.634	**0.100**	W	245.0	−0.898	0.371
Average	Average	W	2437.5	−2.375	**0.018**	S	−0.434		0.666

		Türkiye *df* = 45				Spain *df* = 33			
Pre-COIL	**Post-COIL**	***T***	***Stat***	***z***	***p***	***T***	***Stat***	***z***	***p***
Stranger	Stranger	W	315.0	−0.805	0.420	W	204.0	−0.862	0.392
Acquaintance	Acquaintance	S	−1.271		0.210	W	145.5	−1.789	**0.074**
Dyad	Dyad	W	228.5	−1.180	0.241	S	−0.351		0.728
Public	Public	W	260.0	−1.603	0.110	W	153.5	−2.066	**0.039**
Meeting	Meeting	W	258.5	−0.393	0.700	S	−0.400		0.692
Small Group	Small Group	W	181.0	−1.313	0.192	W	135.5	−1.537	0.127
Friend	Friend	W	352.5	−0.523	0.602	S	−0.902		0.374
Average	Average	W	349.5	−1.698	**0.090**	W	196.5	−1.501	0.136

Note. T = Test, where S = Student signed-rank test and W = Wilcoxon signed-rank test. Calculated using a Shapiro-Wilk Test of Normality with JASP software.

Appendix F

CA t-*test*

		Total *df* = 117				Japan *df* = 37			
Pre-COIL	**Post-COIL**	***Test***	***Stat***	***z***	***p***	***Test***	***Stat***	***z***	***p***
Stranger	Stranger	W	3103.0	1.987	**0.046**	S	2.016		**0.051**
Acquaintance	Acquaintance	W	3116.5	1.068	0.284	S	−0.543		0.591
Dyad	Dyad	W	2723.5	1.056	0.291	S	0.537		0.595
Public	Public	W	2398.0	0.256	0.799	S	−0.550		0.586
Meeting	Meeting	W	2700.0	1.359	0.174	S	−0.421		0.676
Small Group	Small Group	W	2614.5	0.487	0.627	S	0.331		0.742
Friend	Friend	W	2524.5	−0.340	0.734	S	−1.110		0.274
Average	Average	W	3608.5	1.290	0.197	S	−0.038		0.970

		Türkiye *df* = 45				Spain *df* = 33			
Pre-COIL	**Post-COIL**	***Test***	***Stat***	***z***	***p***	***Test***	***Stat***	***z***	***p***
Stranger	Stranger	W	515.0	1.101	0.271	S	−0.094		0.926
Acquaintance	Acquaintance	S	2.043		**0.047**	S	0.242		0.810
Dyad	Dyad	W	407.5	0.845	0.402	W	219.0	0.364	0.724
Public	Public	S	0.673		0.505	S	−0.306		0.761
Meeting	Meeting	S	1.790		**0.080**	W	302.0	1.058	0.293
Small Group	Small Group	W	317.5	0.661	0.514	S	−0.254		0.801
Friend	Friend	W	352.5	0.015	0.994	W	280.5	0.309	0.764
Average	Average	W	583.0	1.644	0.101	W	328.5	0.530	0.602

Note. S = Student signed-rank test. W = Wilcoxon signed-rank test. Calculated using a Shapiro-Wilk Test of Normality with JASP software.

Appendix G

3SQ t-*test*

		Total *df* = 113				Japan *df* = 37			
Pre-COIL	**Post-COIL**	***T***	***Stat***	***z***	***p***	***T***	***Stat***	***z***	***p***
Autonomy	Autonomy	W	1490.0	−1.316	0.184	S	−0.696		0.491
Collaboration	Collaboration	W	2343.0	−0.292	0.769	S	0.702		0.487
Commitment	Commitment	W	1957.5	−0.707	0.475	S	0.581		0.564
Curiosity	Curiosity	S	−1.403		0.163	S	−0.868		0.391
Empathy	Empathy	W	1611.5	−0.774	0.436	W	237.5	0.786	0.433
Leadership	Leadership	W	2002.5	−1.649	0.099	S	−0.436		0.666
Openness	Openness	W	1715.5	−0.840	0.396	S	1.493		0.144
Problem Solving	Problem Solving	W	1788.5	−1.042	0.294	S	0.331		0.743
Resilience	Resilience	W	1713.5	−1.657	0.096	S	−0.668		0.508
Self-confidence	Self-confidence	W	1521.0	−2.546	0.011	W	101.0	−2.323	0.019
Average	Average	W	2661.0	−1.460	0.145	S	−0.212		0.834

		Türkiye *df* = 44				Spain *df* = 30			
Pre-COIL	**Post-COIL**	***T***	***Stat***	***z***	***p***	***T***	***Stat***	***z***	***p***
Autonomy	Autonomy	S	−1.751		0.087	W	113.5	0.317	0.762
Collaboration	Collaboration	S	−0.748		0.458	S	−0.542		0.592
Commitment	Commitment	W	163.0	−2.099	0.034	S	0.605		0.550
Curiosity	Curiosity	S	−0.195		0.846	S	−1.853		0.074
Empathy	Empathy	W	226.5	−1.866	0.058	S	0.183		0.856
Leadership	Leadership	S	−1.286		0.205	S	−1.598		0.121
Openness	Openness	W	143.5	−2.448	0.014	W	125.0	−0.395	0.697
Problem Solving	Problem Solving	S	−1.247		0.219	S	−1.438		0.161
Resilience	Resilience	S	−1.418		0.163	W	81.0	−1.199	0.232
Self-confidence	Self-confidence	W	302.5	−1.445	0.148	S	−0.636		0.529
Average	Average	W	368.0	−1.482	0.140	S	−1.10		0.280

Note. T = Test, where S = Student signed-rank test and W = Wilcoxon signed-rank test. Calculated using a Shapiro-Wilk Test of Normality with JASP software.