2024-2-FR01-KA210-SCH-000286488

Final Evaluation Report

Social, Emotional and Cognitive Outcomes in a Cross-National Implementation

Erasmus+
Enriching lives, opening minds.

Cofinancé par l'Union européenne

Co-Financé par l'Union européenne. Les points de vue et avis exprimés n'engagent toutefois que leur(s) auteur(s) et ne reflètent pas nécessairement ceux de l'Union européenne ou de l'Agence exécutive européenne pour l'éducation et la culture (EACEA). Ni l'Union européenne ni l'EACEA ne sauraient en être tenues pour responsables.

Content

I. ETHICAL CONSIDERATIONS

All consent letters were translated into the language of instruction. When selecting the sample and determining participating classes, the teacher was asked which first languages were spoken in the class. Where possible, and with the support of an appropriate translation service, the letters of consent were translated into the students' first languages. This was required in all countries to ensure non-discriminatory practice and compliance with ethical guidelines, and we did not expect potential participants to translate the consent letters themselves, as this would have hindered recruitment. When students received the consent letters in their first language, they were provided with two versions, one in the language of instruction and one in their first language, and, ideally, both copies were signed.

Before the survey in September, all signed consent letters were collected from each school/class. These were kept and organized by class. It was ensured that letters from different classes were not mixed up. After organizing the consent letters by class, an alphabetical list of participating students was created. This list contained only the names of students who had actually submitted a signed consent form. The list of participants was then coded according to the attached instructions. A list linking names and codes was maintained. It was kept locked and secure. In the event of withdrawal of consent, all personal data was destroyed. If questionnaires had already been coded, the relevant code could be identified using this list.

II. METHODOLOGY

The student questionnaires were prepared before the pre-test survey. The scale regarding friendships (contact question) was left blank in the current version. For the survey, all names of students who had provided consent forms were entered on the lines and prepared. The names of classmates without consent were not included under any circumstances. The name entry was done digitally; handwriting was not permitted. Extra lines were removed. Each scale was kept entirely on one page. It was acceptable if whole tables moved to another page, as long as all related items remained together. The questionnaires were prepared according to the number of participants. Trainers generally required the student questionnaires, and the teacher questionnaires corresponded to the number of participating students. Spare questionnaires (approximately five) were brought along in case they were needed.

As this was a cross-national study, it was essential to provide accurate translations of all instruments to ensure the usability and comparability of the data. The translation process followed a structured methodology, including the recruitment of a translation team, forward translation, back translation, and committee consolidation. This procedure was based on established guidelines for cross-cultural research (Cheung et al., 2020), which describe the systematic adaptation of study instruments to different languages and cultural contexts.

Erasmus+
Enriching lives, opening minds.
Cofinancé par l'Union européenne
Co-Financé par l'Union européenne. Les points de vue et avis exprimés n'engagent toutefois que leur(s) auteur(s) et ne reflètent pas nécessairement ceux de l'Union européenne ou de l'Agence exécutive européenne pour l'éducation et la culture (EACEA). Ni l'Union européenne ni l'EACEA ne sauraient en être tenues pour responsables.

III. QUANTITATIVE INSTRUMENTS

1. Students' questionnaire (quantitative)

The student questionnaire was administered in dedicated lessons before the program began and again after it concluded. To ensure robust pre–post comparisons, both administrations of the quantitative student data employed identical scales and items.

Regarding the pre-test phase, the purpose was to assess the initial situation of the students participating in the PBP, so that any subsequent developments could be monitored across social, emotional, and academic levels. The pre-test was carried out before the first project unit and was conducted by the trainer, with the teacher providing support when necessary. It was important that the children completed the questionnaire independently, while receiving clear instructions. The person administering the test guided the children by reading each item aloud and waited until all children had finished responding before moving on to the next item. In this context, care was taken to ensure that students read and ticked the correct line. Line markers, such as bookmarks or rulers, were used to assist when needed. The survey took approximately one lesson to complete. If the questions could not be answered within this time, the session was extended as necessary. Teachers were informed in advance to prevent interruptions.

The student questionnaires included the following content:

Demographic Variables of students

Self-reported data on gender, country of birth, and family language were collected. This information was used to describe the sample and for differentiated analyses.

Sociometric Nominations

Friendships were measured via sociometric nominations: students used a class roster to identify their relationships with peers (Pijl et al., 2008; Bukowski et al., 2009) and rated the extent of shared time on a five-point Likert scale.

Emotional Self-Concept Questionnaire (ESC-Q)

The students' overall emotional self-concept was measured with the Emotional Self-Concept Questionnaire (ESC-Q). Designed with clear and accessible language, the ESC-Q is suitable for learners with academic support needs. The instrument comprises 11 items rated on a 4-point Likert scale and captures four domains: emotion perception, self-concept of emotion regulation, self-concept of emotion expression, and self-concept of emotion coping (Schmidt & Zurbriggen, 2021). The ESC-Q was administered at every quantitative data collection wave.

Perception of Inclusion Questionnaire (PIQ)

The PIQ, developed by Venetz and colleagues (2014), captures three core facets of students' school experience: emotional well-being, social inclusion, and academic self-concept. Each subscale comprises items targeting specific elements of these domains, for example, attitudes toward attending school, relationships with classmates, and perceptions of one's own learning capabilities. Responses are recorded on a four-point Likert scale from strong disagreement to strong agreement, yielding a theoretical midpoint of 2.5. To strengthen construct measurement, each subscale also includes one

negatively worded item. Evidence from most contributing countries suggests the PIQ's satisfactory psychometric properties in primary and/or secondary school settings (e.g.: Guillemot and Hessels, 2022; Kostorz et al., 2022; Zwierzchowska et al., 2022)

Mathematic Self-Efficacy

The shortened scale comprises eleven items that measure mathematical self-efficacy and stress-related reactions (Usher & Pajares, 2009). It covers three dimensions: mastery/competence (items 1–3, 5), social persuasion/feedback (items 6–9), and physiological/emotional state (items 10–11, negatively worded, reverse-coded). In addition, one item measures vicarious experience (item 4). The assessment is based on a 4-point Likert scale (1 = "does not apply at all" to 4 = "applies completely"). Higher total scores indicate a stronger sense of self-efficacy and lower math-specific anxiety. The internal consistency of the scale is satisfactory (Cronbach's $\alpha \geq 0.7$).

2. Event Sampling Method (ESM; quantitative)

The ESM was employed across seven instructional sessions, allowing for momentary assessments of students' emotional and social states at different points throughout each session. During each session, participants were asked to fill out a set of contextual questions regarding the social setting and their engagement in the Bridge game, including: (-) whether they had the choice of playing with specific partners or were assigned to a group, (-) their satisfaction with the group members, (-) whether they had to solve a Bridge-related problem during the session, and if so, whether they were able to solve it, and (-) whether they had played Bridge earlier in the session, and if so, the outcome of the game.

Additionally, the Positive and Negative Activation Scale (PANAS/PANAVAS), was used to measure participants' emotional activation during each session (Schallberger, 2005; Schreiber & Jenny, 2020; Zurbriggen & Venetz, 2018). This scale provides insight into the positive and negative affective states experienced by the students, capturing fluctuations in their emotional engagement throughout the Bridge activity.

The ESM approach offers a dynamic and context-sensitive way to explore the relationships between students' emotional states, social dynamics, and their engagement in the program over time, helping to illuminate how various factors influence emotional well-being and social inclusion in real-time. This description links the Event Sampling Method with the specific variables assessed, providing an overview of the procedure and its relevance to understanding students' social and emotional experiences during the intervention.

3. Teachers' questionnaire (quantitative)

This questionnaire was used at the beginning of the survey and was completed by the students' class teacher, who was assumed to teach the students regularly, know them well, and be able to assess them accurately. While the students completed their questionnaires, the teachers also received their questionnaires. Ideally, the teachers completed the questionnaires during the student survey. If this was not possible due to time constraints, they were given the opportunity to complete the questionnaires before the trainer returned for the first training session. The teacher questionnaires were collected at least immediately before the first training session. It was ensured that the students were not involved in filling out the teachers' questionnaires and that the teacher questionnaires were completed prior to the start of the first training unit.

The teacher questionnaires included the following content:

Demographic Variables of students

Demographic data on the students was collected via questionnaires for teachers: The current school level, information on special educational needs, existing diagnoses, and first languages were recorded. This data was used to characterize the sample and for further analysis.

Mathematic Self-Efficacy

Five of the eleven items on the self-efficacy scale for mathematics were presented to teachers for individual assessment of their students as part of the questionnaires (items 1–3, 9–10).

IV. QUALITATIVE DESIGN

Focus group interviews were conducted to collect qualitative insights on participants' experiences with the intervention program, perceived challenges, and its impact on students. Focus groups were chosen as a method because they allow data collection through group interaction, encourage focused discussion, and create a supportive environment for participants (Morgan, 1996; Krueger, 1994; Abramczyk, 1995).

Separate focus groups were conducted with students, teachers, and/or trainers to capture multiple perspectives on the implementation and effects of the program. In each country, approximately four students participated in the student group. Teacher and trainer groups varied in size but followed the same structured format. All sessions were conducted in person by project team members, lasted about 60 minutes, and were audio-recorded with participants' consent. Recordings were transcribed in the respective country, translated into English, and analyzed centrally in Vienna.

Moderators followed a structured guide designed to maintain a friendly and informal atmosphere, ensure that all participants could contribute, monitor participants' well-being, and remain neutral and non-judgmental throughout. The discussion followed a pre-defined schedule: informal welcome, introduction of moderator and participants, explanation of purpose, and consent for recording. The main discussion covered general experiences and impressions of the project, experiences with preparation and materials, implementation in school routines, perceived impact on students, and suggestions for improvement. Sessions concluded with a summary of key points and informal wrap-up.

Ground rules were emphasized at the beginning of each session: participation was voluntary, participants could skip any questions or leave at any time, all responses were valid, and participants were asked to respect others' views. Guiding questions were used to facilitate discussion while allowing participants to share freely, covering topics such as favorite activities, challenges, observed effects on engagement and well-being, and recommendations for future improvements.

This structured approach ensured that data was collected consistently across all participant groups and countries, while maintaining a safe, supportive environment that encouraged open sharing of experiences from the perspectives of students, teachers, and trainers.

V. PARTICIPANTS

The study included a total of 391 students from three participating countries: Poland, France, and the Netherlands. Participation was structured according to the project The study included 391 students from three participating countries: Poland, France, and the Netherlands. Participation followed the project requirements regarding the number of schools and classes per country. In Poland, data were collected from four schools encompassing six classes, including 130 students. In France, six schools with eight classes participated, including 146 students, and in the Netherlands, three schools with six classes contributed 115 students.

Most students were born in the country of the survey. Among the 358 students with valid data, 88.8% were born in the country where the survey was conducted. Regarding family language, 78.5% spoke the language of instruction at home, while 21.5% spoke a different language. The gender distribution of 355 students with valid data was 54.9% female, 43.9% male, and 1.1% diverse. Most students were monolingual (88.5%), with 10% bilingual and 0.8% trilingual, based on 358 valid cases. Information on special educational needs (SEN) was available for 254 students, with 11% reporting a formal diagnosis and 89% without.

This sample represents a diverse population across multiple schools and classes in each country, reflecting variations in language background, gender, and educational needs. The structured inclusion of a defined number of schools and classes per country ensured that project requirements were met and allowed for systematic cross-national comparisons of student experiences and outcomes.

VI. RESULTS

1. Scale reliability

Reliability of the ESC-Q

The internal consistency of the Emotional Self-Concept Questionnaire (ESC-Q) was evaluated using Cronbach's alpha, based on pairwise complete observations across the entire cross-national sample. Country-specific reliability estimates were not calculated due to sample size constraints and the focus on the overall applicability of the instrument across all participating students.

For the Emotions Perception (EW) subscale, reliability was excellent, with a Cronbach's alpha of 0.941 (95% CI: 0.931–0.950). The Emotions Modulation (EM) subscale also demonstrated excellent internal consistency, with α = 0.929 (95% CI: 0.917–0.940). Considering the ESC-Q as a total scale, internal consistency was extremely high, with a Cronbach's alpha of 0.972 (95% CI: 0.967–0.976).

These results indicate that the ESC-Q is a highly reliable instrument for measuring emotional self-concept in students across multiple countries, with both subscales and the overall scale showing excellent internal consistency.

Reliability of the PIQ

The internal consistency of the Perceptions of Inclusion Questionnaire (PIQ) scales was evaluated using Cronbach's alpha, based on pairwise complete observations across the entire cross-national sample. Country-specific reliability estimates were not calculated due to sample size constraints and the primary focus on assessing the overall applicability of the instrument across all participating students. For the Students' Emotional Well-being (PIQ SEMI) scale, internal consistency was excellent, with a Cronbach's alpha of 0.850 (95% CI: 0.824–0.873). The Social Relationships with Other Students (PIQ SSOI) scale demonstrated moderate reliability, with α = 0.598 (95% CI: 0.528–0.659). Finally, the Academic Self-Concept (PIQ SASC) scale showed somewhat lower reliability, with a Cronbach's alpha of 0.565 (95% CI: 0.490–0.630). Overall, the results suggest that the PIQ scales exhibit acceptable to excellent internal consistency across the three participating countries. The SEMI scale demonstrates strong reliability for measuring students' emotional well-being, while the SSOI and SASC scales show moderate reliability, indicating that these scales are suitable for research purposes within this cross-national student sample, though further studies may consider item-level refinement or cultural adaptation to improve reliability in the social and academic dimensions.

Reliability of the Mathematic Self-Efficacy Scale

The internal consistency of the Self-Efficacy scale was evaluated using Cronbach's alpha, based on pairwise complete observations across the entire cross-national sample. Country-specific reliability estimates were not calculated due to sample size limitations and the primary focus on assessing the scale's overall applicability across all participating students. The Self-Efficacy scale demonstrated good internal consistency, with a Cronbach's alpha of 0.793 (95% CI: 0.761–0.822). These results indicate that the scale is a reliable instrument for measuring students' perceived self-efficacy within this cross-national student sample.

Summary

The internal consistency of the PIQ, Self-Efficacy, and ESC-Q scales was assessed across the entire cross-national sample. All instruments demonstrated acceptable to excellent reliability, providing a solid basis for further statistical analyses. These results ensured that the measures were suitable for examining students' perceptions of inclusion, self-efficacy, and emotional self-concept in the study.

2. Quantitative results

Pre and post comparison

Emotional Self-Concept Questionnaire (ESC-Q)

The analysis of students' emotional self-concept scores from pre- to post-intervention showed no significant change.

For the Emotions Perception (EW) subscale, statistical testing indicated a z-value of 0.084 (df = 297, p = 0.933), with a negligible effect size of +0.005 (SE = 0.060). The mean score increased only minimally by +0.003 points.

For the Emotions Modulation (EM) subscale, the analysis produced a z-value of -1.448 (df = 298, p = 0.149), with a small effect size of -0.084 (SE = 0.058). The mean score decreased slightly by -0.028 points.

For the overall emotional self-concept score, results indicated a z-value of -0.910 (df = 298, p = 0.364), with a negligible effect size of -0.052 (SE = 0.057). The mean score decreased minimally by -0.018 points.

Overall, students' emotional self-concept remained stable following the intervention. These findings suggest that the program did not undermine students' emotional skills. Students maintained their ability to perceive and modulate emotions throughout the course of the intervention, supporting emotional stability while highlighting areas for potential growth and awareness.

In summary, the intervention appears to preserve students' emotional competencies, maintaining stability in both perception and modulation of emotions while providing a safe context to reflect on and manage their emotional experiences.

Perception of Inclusion Questionnaire (PIQ)

The analysis of students' perceptions of inclusion from pre- to post-intervention revealed small but statistically significant improvements in some domains.

For the Emotional Well-Being (SEMI) subscale, scores increased slightly from pre- to post-test (t(274) = 2.186, p = 0.030), with a small effect size (Cohen's d = 0.132, SE = 0.061).

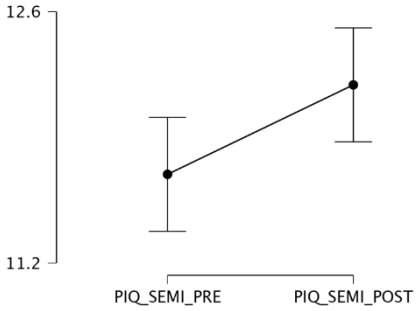

Figure 1: Pre-Post-PIQ-SEMI

Similarly, the Social Relationships (SSOI) subscale showed a modest but significant improvement (t(275) = 2.210, p = 0.028, Cohen's d = 0.133, SE = 0.068).

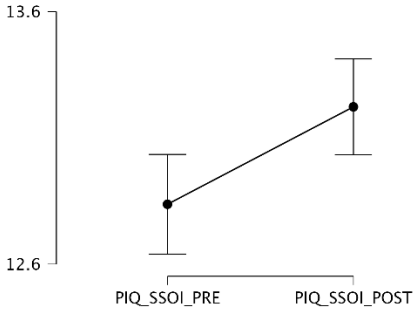

Figure 2: PRE-POST-PIQ-SSOI

In contrast, no significant change was observed for the Academic Self-Concept (SASC) subscale (t(273) = 1.073, p = 0.284, Cohen's d = 0.065, SE = 0.058).

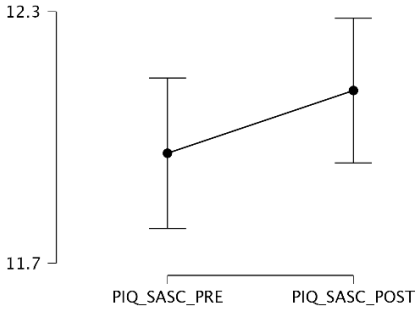

Figure 3: PRE-POST-PIQ-SASC

Overall, these results indicate that the PBAS program slightly enhanced students' emotional and social perceptions of inclusion, while their academic self-concept remained stable. The effect sizes were small, suggesting that changes were modest but consistent across the sample.

Mathematic Self-Efficacy (MSE)

The analysis of students' self-efficacy scores from pre- to post-intervention showed no significant change. Statistical testing indicated a z-value of -0.629 (df = 269, p = 0.530), with a negligible effect size of -0.038 (SE = 0.053). Correspondingly, the mean self-efficacy score increased only minimally by +0.016 points.

Overall, students' math self-efficacy remained stable across most items after the PBAS program. Stress levels showed a slight increase when starting math tasks, but these changes were small. These findings suggest that the PBAS activities did not undermine students' confidence; rather, they appear to have helped highlight potential challenges and stress points, which can be addressed proactively. In this way, the program supports maintaining students' confidence while increasing awareness of areas that may require additional support.

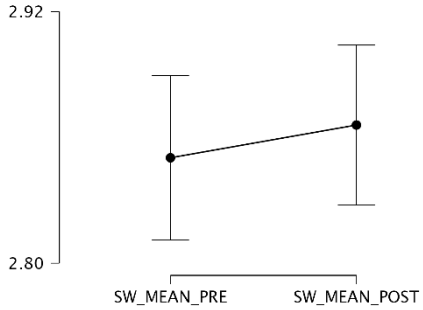

Figure 4: Pre-Post-Comparison MSE

3. Quantitative results

Event Sampling Method

Participants were observed over seven instructional sessions, each involving a Bridge game. Momentary affect was assessed once per session using the Positive and Negative Activation Scale.

Session	N (PA)	M (PA)	SD (PA)	N (NA)	M (NA)	SD (NA)
1	272	3.57	0.69	270	4.39	0.96
2	240	3.62	0.73	239	4.58	0.96
3	258	3.62	0.76	257	4.42	1.02
4	268	3.64	0.71	267	4.28	0.79
5	286	3.58	0.70	284	4.42	0.86
6	290	3.58	0.70	290	4.42	0.84
7	259	3.63	0.78	259	4.34	0.85

Table 1: ESM - Positive and Negative Activation (PA/NA) across Seven Sessions

Across the seven sessions, positive activation (PA) ranged from 3.57 (SD = 0.69) in session 1 to 3.64 (SD = 0.71) in session 4, with the remaining sessions showing values between 3.58 and 3.63 (SD = 0.70–0.78). Negative activation (NA) was consistently higher than PA, ranging from 4.28 (SD = 0.79) in session 4 to 4.58 (SD = 0.96) in session 2, with the other sessions between 4.34 and 4.42 (SD = 0.84–1.02). Both PA and NA remained relatively stable across sessions, showing no clear increasing or decreasing trends. PA scores exhibited moderate variability, whereas NA scores showed somewhat greater dispersion, suggesting that some participants experienced stronger negative activation in certain sessions. Overall, participants reported higher levels of negative than positive activation during the Bridge sessions, which may reflect the cognitive demands of the tasks or social dynamics within the groups.

To examine the relationships between affective states during the seven Bridge sessions and accompanying contextual variables, Pearson correlations were computed. The affective states include Positive Activation (PA) and Negative Activation (NA) at each measurement point (ESM_1 to ESM_7). Contextual variables include free choice of playing partners, satisfaction with group members, success in solving Bridge tasks, and game outcomes. The analysis revealed consistent significant positive correlations among PA scores across sessions ($r \approx 0.3–0.7$, $p < .05$), indicating stability in positive activation throughout the instructional period. Similarly, NA scores were positively correlated over time. Positive Activation scores showed moderate positive associations with free choice of partners and group satisfaction, suggesting that participants who could choose their playing partners or were more satisfied with their group tended to experience higher positive activation. Conversely, NA scores were negatively correlated with these variables, aligning with the interpretation of negative activation as reflecting stress or dissatisfaction. Successful problem-solving in Bridge tasks was positively correlated with PA and negatively with NA, highlighting a link between coping success and more favorable affective experiences. Game outcome displayed weaker but some significant correlations with affective measures, with winning associated with higher positive activation. Overall, these findings emphasize the importance of social and performance-related contextual factors for affective experience in the classroom setting, highlighting the role of positive social interactions and success experiences in shaping emotional well-being during Bridge sessions.

Sociometric nominations

This study aimed to assess the changes in social integration among students across multiple classes following an intervention. The analysis was performed to determine whether there was a significant difference in students' social integration, measured before (Pre) and after (Post) the intervention. The results of the paired-samples t-tests for each class are summarized below.

Important Note: The numbering of the classes in this table does not correspond to the coding used in the research project. Some classes had to be excluded from the analysis due to missing data, which caused variations in the number of students per class. Therefore, the total sample size (N) for each class may differ.

Class	N	Pre Mean (SD)	Post Mean (SD)	Mean Difference (Post–Pre)	Paired t	df	p (two-tailed)
Class 1	20	3.15 (0.82)	2.92 (0.91)	–0.23	1.28	19	0.215
Class 2	22	2.55 (0.64)	2.58 (0.59)	-0.03	-0.26	21	0.801
Class 3	19	2.42 (0.40)	2.66 (0.74)	-0.24	-1,72	18	0.103
Class 4	17	3.07 (0.40)	3.08 (0.76)	-0.01	-0.04	16	0.971
Class 5	17	2.58 (0.49)	2.88 (0.52)	-0.30	-3.35	16	0.004
Class 6	12	3.03 (0.38)	3.31 (0.45)	-0.28	-1.37	11	0.198
Class 7	18	2.36 (0.51)	2.68 (0.60)	-0.31	-1.93	17	0.070
Class 8	17	2.27 (0.52)	2.42 (0.54)	-0.15	-1.23	16	0.238
Class 10	15	2.17 (0.52)	2.07 (0.40)	0.10	0.72	14	0.485
Class 11	16	2.13 (0.67)	2.48 (0.67)	-0.35	-2.34	15	0.034
Class 12	19	2.59 (0.60)	2.81 (0.65)	-0.23	-2.58	18	0.019
Class 15	20	2.22 (0.91)	2.06 (0.65)	0.16	0.81	19	0.429
Class 16	18	2.33 (0.70)	2.33 (0.70)	0.08	0.78	17	0.444

Table 2: Paired t-Test Results for Social Integration Changes by Class

For Class 1, the social integration scores showed a slight decrease from pre- to post-intervention, although the change was not statistically significant (p = 0.215). In Class 2, the mean ratings for social integration remained almost identical before and after the intervention, with a mean difference of only 0.03, which also was not significant (p = 0.801). This suggests no meaningful effect of the intervention in this class. Similarly, Class 3 exhibited a modest increase in social integration after the intervention, but this change was not statistically significant either (p = 0.103). This indicates that while there was some improvement, it was not robust enough to be considered a significant change. For Class 4, no change in social integration was observed, with the pre- and post-intervention means being nearly identical. The lack of statistical significance (p = 0.971) further supports the finding that the intervention had no measurable effect on social integration in this class. In Class 5, a statistically significant improvement in social integration was found (p = 0.004). This indicates that the intervention had a moderate positive effect on students' social integration in this class. Class 6 did not show a significant change (p = 0.198). While there was a small increase in social integration, the p-value suggests that this change was not large enough to be statistically significant. For Class 7, the change in social integration was close to the threshold of statistical significance (p = 0.070), suggesting a potential effect of the intervention that was not sufficiently strong to reach the conventional significance level. Class 8 showed no significant change (p = 0.238), with a small difference between the pre- and post-intervention means. This suggests that the intervention did not significantly impact students' social integration in this class. Class 10 exhibited no significant change in social integration (p = 0.485), indicating that the intervention had no measurable effect on students' social integration in this class. A statistically significant change was found in Class 11 (p = 0.034), where students showed an improvement in social integration following the intervention. This highlights the positive impact of the intervention in this class. In Class 12, a significant improvement was also observed (p = 0.019), indicating that the intervention had a meaningful effect on social integration among students in this class. Class 15 showed a small positive change, but the p-value (p = 0.429) indicated that this difference was not statistically significant, suggesting that the intervention did not significantly impact social integration in this class. Finally, in Class 16, no significant change was

observed (p = 0.444), indicating that the intervention had no measurable effect on social integration for the students in this class.

The analysis reveals that some classes showed significant changes in social integration following the intervention, particularly Classes 5, 11, and 12, where significant improvements were observed. In contrast, other classes exhibited very small or non-significant changes. This suggests that the intervention had differential effects across different classes, with some classes benefiting more than others.

4. Qualitative results

Overview

Dimension	Key finding	Main hurdle	Recommendation
Engagement	High motivation, tournaments drive interest	Long/complex questionnaires reduce attention	Shorter sessions, occasional "just play" time
Learning	Fast gains in rules, trick-taking, basic scoring	Tight time per deal	Fewer deals early; phased scoring guidance
Social	More cooperation, respect, responsibility	Partner choice and dummy acceptance	Flexible pairing, clear role rules
Organization	Volunteers are crucial for smooth delivery	High time/space demands, hard to run alone	Secure volunteer support, easily reconfigurable setup
Materials	Clear step-by-step guides	Cards not robust, missing directional arrows	Durable cards with arrows, more flexible/round mats

Table 1: Overview Qualitative Key Results

Student perspective

France

From the perspective of the French students, Petit Bridge was experienced as fun, engaging, and different from regular lessons. Playing and the tournament were the highlights, with many feeling proud as they improved and understood the rules and roles. Early challenges such as laying out tricks correctly, dealing and following arrows, or accepting the dummy role became manageable with practice. Students described stronger cooperation, better respect for partners, and a sense of responsibility when making decisions that affect the pair's result. They valued the social aspect of sitting together, talking, and playing without screens, and appreciated small rewards that celebrated effort. Many believed the project supported logical thinking and planning and would recommend it to other classes. Their main wishes were to play more often and to have occasional sessions focused purely on the pleasure of playing.

Poland

From the perspective of the Polish students, Petit Bridge was experienced as very exciting, educational, and varied. Playing the game and the final tournament were particularly enjoyable, and many felt proud because they realized they were good at it. At first, there were hurdles such as dealing correctly according to arrows, holding the cards, and understanding why you don't always play the highest card right away, but with practice it quickly became easier. In everyday school life, Petit Bridge provided a welcome change from pure textbook lessons, strengthened the class community, and promoted cooperation, a sense of responsibility, and respect, even when playing with people they didn't like. Many are convinced that the project helps with logical thinking, memory, and thoughtful decision-making, and they clearly recommend it to other classes. Wishes for further development mainly concerned more frequent lessons and, in some cases, fixed partners. The children would change little in terms of content.

Teacher perspective

France

From the perspective of French teachers and volunteers, Petit Bridge at School was found to be motivating, socially empowering, and easy to integrate into lessons. The students quickly got involved and showed clear progress in card handling, understanding the rules, and cooperative play, even among children who were initially reserved or challenging. Tournaments in particular increased motivation, and small prizes made successes visible. Initial indications of beneficial thought processes were noted, such as planning, weighing alternatives, simple additions to 10, and mental arithmetic for scoring, without expecting immediate effects on general math performance. The participants cited time and organizational effort, large groups, and the need for continuous support from volunteers as key obstacles. Suggestions for further development included clearer guidelines for laying out the stitches, more adaptable mats, easily accessible bilingual materials, and short projectable problem situations as a ritual. The key results are high acceptance, visible gains in competence, successful tournament implementation, and concrete approaches for optimization for future runs.

The Netherlands

From the Netherland's perspective, the overall picture is clearly positive. Teachers report that the students were highly motivated, made rapid progress in card hierarchy, trick-taking, understanding roles

(declarer/dummy) and simple scoring, and a final class tournament ran in a structured manner and on schedule. The teacher found the program to be beneficial, well suited to self-directed learning, and conducive to cooperation and concentration. The materials were considered practical and continued to be used in class. At the same time, some obstacles became apparent: the extent of new learning content regarding the program and sometimes the time slots per (Bridge) deal were tight, and self-chosen partner constellations occasionally slowed down the flow of the game. The following ideas for further development were mentioned: allowing more time for the early phases of the game or reducing the number of deals, giving clear and consistent instructions on laying out and rotating cards, flexible partner assignment (either by the teacher or by choice), using visual timers, and providing more guidance on scoring or introducing it step by step. The key results are therefore: high acceptance, visible gains in competence, successful tournament implementation, and clear indications for organizational optimizations for future runs.

Volunteer perspective

The Netherlands

From the perspective of the Netherlands' volunteers, the project was perceived as enjoyable, innovative, and easy to implement. The children were very enthusiastic and were able to play bridge after just a few weeks, even some of the younger students. The volunteers were considered patient, flexible, and crucial to the smooth running of the project. There is curiosity about the research results, but also skepticism as to whether social and mathematical skills can be validly measured in such a short time. The clear materials and step-by-step instructions were particularly praised, as was the broad participation, including children with lower academic achievement. Many took cards home and continued playing with their families. Regarding the social impact, existing patterns became apparent, such as preferences for the dummy role or difficulties in cooperation, without any major changes in behavior being observed. For the future, the volunteers see potential applications in gifted education alongside chess and in regular classes as a playful, math-related activity. This requires the continued involvement of volunteers, as teachers do not always feel confident enough to continue on their own. It is recommended that Bridge be implemented in the classroom with a clear age guideline from grade 5 onwards, improved cards with directional arrows, and a revised manual.

Poland

From the perspective of the Polish trainers, Petit Bridge at School was rated very positively. The children were motivated, able to concentrate well, and enjoyed working together. Tournament moments and playing together without using screens strengthened communication, team spirit, and acceptance of the rules. Initial dynamics in the choice of partners became apparent, partly based on ability and partly on friendships, but over time these were accepted in favor of random pairings. The preparatory workshops and clearly structured lesson scenarios were seen as a great help, even though time was short in individual lessons and organizational activities such as rearranging tables took up teaching time. The low durability of the cards was particularly criticized. There is a desire for more robust card sets that are easily available in Poland. The volunteers cite the successful introduction of many children to card and trick-taking game principles, visible social gains through playing and talking.

VII. IMPLICATIONS

Does the Petit Bridge program influence students' cognitive skills, social and emotional inclusion, and engagement?

To connect the objectives of this study to the quantitative results and provide a comprehensive discussion, we can highlight how the program influenced the three main areas: engagement, social inclusion, and cognitive skills.

Cognitive skills:

Despite the positive activation related to social inclusion and emotional well-being, the cognitive outcomes related to mathematics (as assessed through the Mathematic Self-Efficacy Scale or MSE) remained relatively stable, showing no significant change post-intervention: The quantitative data focused on mathematic self-efficacy showed no significant changes from pre- to post-intervention. Mean scores increased minimally (+0.016 points), and effect sizes were negligible (Cohen's d = -0.038, SE = 0.053). This indicates that students' confidence in performing mathematical tasks remained stable after the program. This finding suggests that, while the program may have promoted students' engagement and social inclusion, its direct impact on cognitive skills—specifically in mathematical tasks—was minimal. The lack of significant changes in self-efficacy might indicate that while students enjoyed the program and showed improved social integration, the cognitive challenges associated with math tasks were not directly addressed by the intervention, or the intervention duration was not long enough to produce observable changes in cognitive confidence.

Social and emotional skills:

In terms of social inclusion, results from the Perception of Inclusion Questionnaire (PIQ) indicated small but statistically significant improvements in students' emotional well-being (SEMI, $t(274) = 2.186$, $p = 0.030$, $d = 0.132$) and social inclusion (SSOI, $t(275) = 2.210$, $p = 0.028$, $d = 0.133$). Academic self-concept (SASC) remained stable ($t(273) = 1.073$, $p = 0.284$, $d = 0.065$). Emotional self-concept, measured by the ESC-Q, showed excellent internal consistency across all countries, confirming the reliability of the instrument for these dimensions. This result aligns with the event sampling method (ESM), where positive activation correlated positively with factors like group satisfaction and free choice of playing partners. This suggests that students' social connections during the intervention were influenced by these contextual factors, leading to stronger feelings of inclusion within their groups. Furthermore, the increase in social inclusion aligns with findings from the focus groups, where students across countries reported stronger cooperation, respect, and responsibility, especially in the context of the game. This social component of the program appears to have played a key role in fostering inclusion, as evidenced by both the statistical data and qualitative feedback.

Engagement:

The Quantitative Results section reveals that students' engagement, as inferred from emotional well-being and social relationships, was positively influenced by the intervention, albeit to varying degrees. The Perception of Inclusion Questionnaire (PIQ) showed statistically significant improvements in the Social Relationships (SSOI) and Emotional Well-Being (SEMI) subscales, with moderate effect sizes. These improvements suggest that the intervention provided students with a more positive social environment, which could have contributed to increased engagement. Although the Mathematic Self-

Efficacy scale did not show significant changes, the slight increase in engagement-related areas (e.g., emotional well-being) indicates that students likely felt more comfortable and involved, even though mathematical confidence remained unchanged.

How do participants experience the program in practice?

Students: Across France and Poland, students described the program as fun, engaging, and different from regular lessons. Highlights included playing the game, participating in tournaments, and feeling proud of their improvement. Students reported stronger cooperation, respect, and responsibility, and valued the social aspects of group play without screens. Challenges like understanding rules or holding the cards were quickly overcome through practice.

Teachers: Teachers in France and the Netherlands observed that students were highly motivated, gained competence in game mechanics and scoring, and benefited from structured tournaments. Challenges included time constraints, large group management, and the need for volunteer support. Teachers highlighted opportunities for program optimization, such as clearer instructions and more adaptable materials.

Volunteers/Trainers: Volunteers emphasized that the program was enjoyable, inclusive, and easy to implement. They noted that students quickly learned the game and enjoyed participating, including those with lower academic achievement. Trainers in Poland and the Netherlands reported positive effects on communication, cooperation, team spirit, and rule acceptance.

Which areas of the program could be improved for future implementation?

Suggestions included shorter or phased sessions, flexible partner assignments, clearer instructions for card layout and scoring, more robust materials (cards and mats), and structured volunteer support. Focus groups highlighted the need to balance instructional time with playful engagement to maintain motivation.

Does the program maintain or enhance students' confidence while identifying challenges?

Quantitative and qualitative data together suggest that students' self-efficacy and academic confidence were maintained. Minor increases in stress during math tasks were observed, but these did not undermine confidence. The program provided opportunities for students to recognize challenges, such as working in pairs or managing tournament tasks, in a supportive environment. This awareness could be addressed proactively in future sessions.

Summary

The evaluation demonstrates that the Petit Bridge program supports students' social and emotional development, maintains cognitive confidence, and provides engaging experiences that foster cooperation and responsibility. While some effects were modest, all instruments used (PIQ, ESC-Q, Mathematic Self-Efficacy) were reliable across countries, and qualitative feedback from students, teachers, and volunteers confirms the program's acceptability and practical benefits. In conclusion, while the Petit Bridge program did not produce significant changes in mathematical self-efficacy, it had a positive impact on social inclusion and emotional well-being. These results suggest that the program was more effective in enhancing social and emotional skills, possibly by providing a structured yet fun environment for students to collaborate and interact with peers. This aligns with findings from the focus

groups, where students reported stronger cooperation, responsibility, and enjoyment, which indicates a successful intervention in improving the social dynamics of students. In terms of engagement, the program seems to have sparked students' interest and involvement, especially through the social aspects of playing together. While cognitive outcomes such as math self-efficacy did not show significant improvement, future iterations of the program could consider more focused strategies for cognitive development alongside the emotional and social components. For example, incorporating specific activities that directly target mathematical skills or stress management in math tasks could enhance students' cognitive engagement and boost their confidence in this area. Overall, these results underscore the value of the Petit Bridge program in enhancing social inclusion and emotional well-being, while also suggesting areas for potential refinement to achieve broader cognitive gains. Future research could focus on enhancing the cognitive aspects of the Petit Bridge program, particularly by incorporating targeted mathematical activities to complement the social and emotional components. Additionally, longitudinal studies could examine the long-term effects of the program on social inclusion and emotional well-being. Exploring cultural adaptations and individual differences in future studies could provide valuable insights into how the program performs across different contexts and student groups, helping to refine and optimize its effectiveness for a broader audience.

VIII. REFERENCES

Abramczyk (1995): Focus groups are a phenomenological method that collects qualitative data in a group situation.

Bukowski, W. M., Motzoi, C., & Meyer, F. (2009). Friendship as process, function, and outcome. In K. H. Rubin, W. M. Bukowski, & B. Laursen (Eds.), Handbook of peer interactions, relationships, and groups (pp. 217-231). The Guilford Press.

Cheung, H., Mazerolle, L., Possingham, H. P., Tam, K. P., & Biggs, D. (2020). A methodological guide for translating study instruments in cross-cultural research: Adapting the 'connectedness to nature'scale into Chinese. *Methods in Ecology and Evolution*, *11*(11), 1379-1387.

Guillemot, F., & Hessels, M. G. (2022). Validation of the student version of the perceptions of inclusion questionnaire on a sample of French students. *European Journal of Special Needs Education*, *37*(5), 850-865.

Krueger (1994) describes focus groups as planned discussions that take place in a friendly and non-threatening environment to collect data on a particular topic. The atmosphere is a catalyst for data collection.

Morgan (1996) defines focus groups as a research method that collects data through group interaction on a topic determined by the researcher. The emphasis is on group interaction and focused discussion.

Kostorz, K., Zwierzchowska A., & Ziemba, M. (2022). Effects of the COVID-19 pandemic on the Perception of Inclusion in school education and physical activity among polish students. *Frontiers in Psychology, 13*, 80791. https://doi.org/10.3389/fpsyg.2022.880791

Pijl, S. J., Frostad, P., & Flem, A. (2008). The social position of pupils with special needs in regular schools. Scandinavian Journal of Educational Research, 52(4), 387-405.

Schallberger, U. (2005). *Kurzskalen zur Erfassung der Positiven Aktivierung, Negativen Aktivierung und Valenz in Experience Sampling Studien (PANAVA-KS) [Short scales for the assessment of Positive Activation, Negative Activation and Valence in Experience Sampling Studies (PANAVA-KS)]*. Zürich: Psychologisches Institut der Universität Zürich.

Schmidt, P., & Zurbriggen, C. (2021). Entwicklung und erste empirische Überprüfung des Fragebogens zur Erfassung des generellen emotionalen Selbstkonzepts in der frühen Adoleszenz (FEGESK-A) [Development and preliminary empirical evaluation of a questionnaire assessing the general emotional self-concept in early adolescence (QAGESC-A)]. Empirische Pädagogik, 35(2), 183–200.

Schreiber, M., and Jenny, G. J. (2020). Development and validation of the 'lebender emoticon PANAVA' Scale (LE-PANAVA) for digitally measuring positive and negative activation, and valence via emoticons. *Personality Individual Diff.* 160:109923. doi: 10.1016/j.paid.2020.109923

Usher, E. L., & Pajares, F. (2009). Sources of self-efficacy in mathematics: A validation study. *Contemporary educational psychology*, *34*(1), 89-101.

Venetz, M., Zurbriggen, C., & Eckhart, M. (2014) Entwicklung und erste Validierung einer Kurzversion des „Fragebogens zur Erfassung von Dimensionen der Integration von Schülern (FDI 4-6)" von Haeberlin,

Moser, Bless und Klaghofer. Empirische Sonderpädagogik, 6, 99-113. Available from: https://doi.org/10.25656/01:9247

Zurbriggen, C., and Venetz, M. (2018). “Diversität und aktuelles emotionales erleben von schülerinnen und schülern im inklusiven Unterricht [Diversity and current emotional experience of students in inclusive classrooms],” in *Emotionen und Emotionsregulation in der Schule und Hochschule*, eds G. Hagenauer and T. Hascher (Münster, NY: Waxmann), 87–102.

Zwierzchowska, A., Kostorz, K., Rosołek, B., & Tomińska-Conte, E. (2022). Validation of the perceptions of inclusion questionnaire including PE teachers’ opinion as part of an innovative use of the tool. *Sustainability*, *14*(6), 3241.

IX. VALIDATION AND SIGNATURES

This final evaluation report was prepared within the framework of the Erasmus+ KA210-SCH project "Petit Bridge at School" (2024-2-FR01-KA210-SCH-000286488).

The undersigned confirm the scientific integrity of the data analysis and the accuracy of the findings presented in this document.

universität wien

Project Manager

Katharina-Theresa Lindner
Post-doc Researcher – University of Vienna

Signature:

Researcher

Susanne Schwab
Professor – University of Vienna

Signature:

Erasmus+
Enriching lives, opening minds.

Cofinancé par l'Union européenne

Co-Financé par l'Union européenne. Les points de vue et avis exprimés n'engagent toutefois que leur(s) auteur(s) et ne reflètent pas nécessairement ceux de l'Union européenne ou de l'Agence exécutive européenne pour l'éducation et la culture (EACEA). Ni l'Union européenne ni l'EACEA ne sauraient en être tenues pour responsables.