# Assessing the Impact of Sport for Development on Youth Employability Skills

A Rigorous Impact Evaluation from Albania

# IMPRINT

URL links:
This publication contains links to external websites. Responsibility for the content of the
listed external sites always lies with their respective publishers.
When the links to these sites were first posted, GIZ checked the third-party content to
establish whether it could give rise to civil or criminal liability. However, the constant
review of the links to external sites cannot reasonably be expected without concrete
indication of a violation of rights. If GIZ itself becomes aware or is notified by a third
party that an external site it has provided a link to gives rise to civil or criminal
liability, it will remove the link to this site immediately. GIZ expressly dissociates itself
from such content.

Responsibility:
GIZ is responsible for the content of this publication. The views expressed in this paper
are purely those of the authors and may not in any circumstances be regarded as
stating an official position of GIZ.
The designation of geographical entities in this paper, and presentation of material, do
not imply the expression of any opinion whatsoever on the part of GIZ,
concerning the legal status of any country, territory, or area, or of its authorities,
or concerning the delimitation of its frontiers or boundaries.

Digital publication only

Bonn, September 2024

# CONTENT

# LIST OF ABBREVIATIONS

| | |
|---|---|
| **BMZ** | German Federal Ministry for Economic Cooperation and Development |
| **GIZ** | Deutsche Gesellschaft für Internationale Zusammenarbeit |
| **NEETs** | Not in Education, Employment, or Training |
| **NGO** | Non-Governmental Organisation |
| **PE** | Physical Education |
| **RCT** | Randomized Controlled Trial RCT |
| **RIE** | Rigorous Impact Evaluation |
| **SDG** | Sustainable Development Goal |
| **S4D** | Sport for Development |
| **S4E** | Sport for Employability |
| **VET** | Vocational Education and Training |

# ABSTRACT

To evaluate whether GIZ's Sport for Development (S4D) approach is an effective tool for fostering employability among youth, GIZ, in collaboration with the Faculty of Social Sciences of the University of Tirana, assessed the impact of S4D on communication competences, self-confidence, cooperation competences, and goal orientation of young people in Albania. The study aimed to address the broader question of whether S4D can contribute to achieving Sustainable Development Goal (SDG) Target 8.6, which seeks to substantially reduce the proportion of youth not in employment, education, or training. The intervention was implemented in ten high schools within the municipality of Tirana, with a focus on improving key employability skills among Albanian youth.

The results demonstrate that the S4D intervention had a generally positive effect on various competencies, including communication skills, self-confidence, cooperation competences, and goal orientation, with the intervention group maintaining or slightly improving these skills over time. In contrast, the comparison group, which did not receive the intervention, experienced notable declines across all areas. While communication and goal orientation remained stable in the intervention group, self-confidence showed significant improvement, highlighting the program's effectiveness. However, cooperation competences declined in both groups, though the decline was more pronounced in the comparison group, indicating that the intervention was only partially effective in this area. Overall, the S4D program played a protective role, helping to preserve key skills and prevent the deterioration observed in the comparison group. These findings underscore the importance of targeted interventions in youth development and suggest a need for refining and expanding S4D programs to fully sustain critical competencies over time.

# BACKGROUND

Youth unemployment is a significant problem in Albania, affecting both youth themselves and society as a whole. Unemployment among youth (ages 15-24) in 2023 amounted to 28.2% in Albania. However, when young persons are not in employment nor participating in formal education or trainings, they run an increased risk of becoming disconnected from the labour market and facing social exclusion with effects on their whole adult life. This is why many scholars and labour market experts also take into consideration the rate of NEETs: Young persons in an economy aged 15-24 not in education, employment, or training. The share of NEETs amounted in Albania to 24.1% in 2021 (ILO n.a.).

Gender disaggregated data shows slight differences between males and females: In Albania, youth unemployment in 2023 among females amounted to 27% and among males to 29.1%. The share of NEETs in Albania amounted to 22.2% among males and to 26% among females in 2021. (ILO n.a.).

Inconsistent quality of education and training, which often does not meet the requirements of the labour market, is frequently mentioned as one of the main causes of the high youth unemployment rate in the region. This high unemployment level in turn is one of the main reasons for the outflow of youth from the region, presenting a growing problem because the young generation in particular has the potential to play a decisive role in social, economic, and cultural cooperation as well as in the reconciliation of the entire Balkan region (GIZ 2022). Albania has taken measures to improve employment opportunities for youth. For instance, the so-called Youth Guarantee was introduced in Albania in 2021. Inspired by the implementation of Youth Guarantee Schemes in other European countries, these policies are a commitment to support NEETs. This commitment entitles youth to receive a good quality offer of employment, traineeship, apprenticeship, or continued education within four months of leaving school or becoming unemployed (ILO 2022b). But despite some progress achieved in terms of overall employment, young people, still experience high rates of unemployment and inactivity (European Commission 2020).

Youth in Albania face many challenges in finding good quality jobs that match their skills and aptitudes. Often, education systems fail to provide students with appropriate skills for the labour market, and career advisory services are underdeveloped and lacking a systematic support for career orientation and soft skills development. If at all, offers on skills development are made by non-governmental organisations (NGOs) or international organizations. Public employment services are ineffective in assisting young people into work and in consequence, a large proportion relies on family or political connections to obtain a job or work in the informal sector. Furthermore, there is a shortage of jobs available, and the COVID-19 crisis has brought new job creation to a halt affecting the job prospects of young people. The prevalence of temporary job contracts among young people is a further source of insecurity (RCC, 2021).

# SPORT FOR EMPLOYABILITY INTERVENTION LOGIC

Employability "is the combination of all factors which enable [a young person] to progress towards or get into employment, to stay in employment and to progress during a career" (CEDEFOP, 2011: 46). This combination of factors includes the possession of basic educational skills, vocational qualifications, technical or job-specific knowledge plus the individual′s personal qualities, attitudes, and attributes, usually called soft – or life skills. The International Labour Organisation (ILO) describes four core competencies for good employability: learning to learn, communication, teamwork and problem-solving (Brewer, 2013). The biggest impact of sport may certainly be expected in relation to the development of life and soft skills.

Sport for Employability (S4E) is not a stand-alone concept. It rather represents a specific focus within the broader approach of Sport for Development (S4D). It includes all measures where sport is used in a targeted manner as a tool to promote the different aspects of employability at any stage of career pathways. It is important to note, that employability not only develops through formal education, but also through informal learning and personal development. The individual environment of youth plays a very important role in this regard, sport can help to reach youth who would be hard to reach through other channels. This is due to the fact that sport can draw attention to almost any issue and is considered an attractive activity for the majority of youth. Especially the final years of school and the transition into university, vocational education or work are associated with great challenges and uncertainties for young persons. Disorientation, setbacks, and frustration can be just as much a part of this phase o f life as joyful anticipation, big dreams, and important developmental steps. The stronger and more stable youth are in their personalities, the better they succeed in their transition to adult life. Sport can help build trusting relationships with the target group and strengthen their health, wellbeing, confidence, and resilience. Especially youth from vulnerable groups may benefit from such empowerment, as it provides them with a more solid foundation for the specific challenges of this developmental stage (GIZ 2022).

Building on sport's unique ability to reach out to youth and build trusting relationships with them, physical activity and sport can also be used as the starting point from which youth can be connected to other supporting agencies. In connection with sport-related events or activities, youth can be brought into contact with universities, career counsellors or potential employers in a non-formal, low-barrier environment. However, the most important function of sport in terms of promoting employability lies in its educational potential and the opportunities it offers to teach life skills in a very effective and targeted way (GIZ 2022).

Depending on the specific context, purpose or occupational sector there are long lists of potentially relevant life skills which can be linked to a young person's level of employability: Among many other capabilities, these descriptions often include skills such as adaptability, communication, confidence, conflict resolution, creativity, critical thinking, decision-making, dedication, emotional intelligence, empathy, flexibility, honesty, integrity, leadership, organization, perseverance, politeness, problem-solving, punctuality, reliability, respect for rules, self-discipline, self-motivation, teamwork, tolerance, willingness to learn etc. In addition, many of these skills are inter-related – for example, to be a strong leader one also needs to have good communication and organizational skills. The question of which life skills should actually be developed through an S4E program in order to increase the employability of young people should take several perspectives into account. First, the selection of relevant skills can be based on theoretical considerations

derived from research findings and expert knowledge. Second, the skills that are required and expected by potential employers in the different occupational sectors must be considered and finally, the individual skills deficits identified by trainers, mentors, and the participants themselves should also be taken into account. This means that the choice of life skills to be developed in an S4E program should always consider the specific circumstances (GIZ 2022).

Sport for Development can help youth to equip themselves with a wide range of soft or life skills that match the actual labour market demands. These skills are an indispensable prerequisite for employment. For youth without any previous work experience, they are a key resource to improve their employment prospects, but unfortunately even the best skills portfolio is no guarantee for employment in a tight and competitive labour market as it is characteristic of the Western Balkans. This should not diminish the motivation to develop and implement S4E programs, but it must be taken into account with regard to the expectation management of program developers and youth alike (GIZ 2022).

In the following, the terminology S4D will be used to describe the Sport for Employability intervention in Albania. While it is strictly speaking a S4E intervention, the broader term S4D is more common and in order to avoid unclarity, will be used in this report as a more general wording.

# S4D IMPLEMENTATION IN ALBANIA

On behalf of the German Federal Ministry for Economic Cooperation and Development (BMZ), the Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH has been implementing Sport for Development in the Western Balkans since 2017. GIZ builds capacities of local coaches, teachers and other practitioners to identify and leverage the potential of sport for children's and youth's development.

The S4D implementation targeting the improvement of employability competences among youth in Albania took place in ten high schools in the municipality of Tirana. Ten physical education (PE) teachers and ten social workers and school psychologist were trained in Sport for Employability. In Albania, social workers and psychologists are integrated into Albanian schools and were paired to the respective PE teacher in their school during the S4D implementation. While PE teachers took over the responsibility of facilitating physical exercises and sport, social workers/ psychologists were mainly responsible for the reflection part of the S4D sessions, creating safe spaces for critical reflection and discussion of the competences and topics targeted in the sessions.

For the implementation, GIZ partnered with the non-governmental organisation Epoka e Re in Albania that conducted the S4D training for PE teachers and social workers/ psychologist and accompanied and monitored the implementation in the schools. The target group consisted of youth in the first and third grade of high school, in the age group of 15 to 16 years. Participants attended 90-minute S4D sessions as after school activities once per week for the duration of one school year. In total, 24 sessions were held. The targeted competences of the S4D intervention were communication, self-confidence, cooperation and goal orientation. Additionally, students had the opportunity to realise two days of job shadowing in local businesses or public institutions in order to gain practical insights.

# RESEARCH DESIGN

To analyse whether GIZ's S4D approach is a meaningful tool to foster employability among youth, GIZ in collaboration with the Faculty for Social Science of the University of Tirana, evaluated if S4D impacts the communication competences, self-confidence, cooperation competences and goal orientation of young persons in Albania; and thus examined the following question:

Can S4D contribute to the realization of SDG Target 8.6: By 2020 substantially reduce the proportion of youth not in employment, education or training?

A quasi-experimental, longitudinal study design was used to examine possible impacts of Sport for Development (S4D) on youth' employability skills. The rigorous impact evaluation (RIE) consisted of one intervention group and one comparison group as well as two points of measurement. The baseline was conducted in October/ November 2023 and the endline in May/ June 2024.

While the comparison group did not participate in any type of S4D activities, the intervention group participated in Sport for Development activities once per week for one school year. Schools for the S4D intervention were selected via an open call and then selected via randomization. Participation criteria for schools were the following: High school; availability of an adequate number of students to participate in the intervention and comparison group; permission of municipalities to cooperate with schools; interested PE teacher and social worker/psychologist.
The selection of students in the intervention was carried out via an open call within each school to register for voluntary participation. The criterium was to be in the age group of 15 to 16 years. A random sample was taken from the registered students which were sampled into intervention and comparison group. It was planned to sample 50 students per school (25 into the intervention group and 25 into the comparison group). However, due to the too small numbers of students registered in some schools, it was not always possible to fulfil the quota of 50 students per school. In such cases, all registered students were selected and thus a randomization was not possible. In a few schools, less than 50 students registered. In order to obtain a somewhat equal size of intervention and comparison group (250 students per group were initially planned), more students in other schools were selected. This didn't allow for an equal representation of the schools within the sample, which will be discussed in the chapter on limitations later. While the study was planned as a Randomized Controlled Trial, due to these circumstances, it could only be conducted as a quasi-experimental study.

The allocation ratio between intervention and comparison group is 1:1. From 630 students registered in total, 235 students were selected for the intervention group and 259 students for the comparison group. In total, 494 students participated in the study. Not all 630 students could participate in the S4D intervention or the study due to limited spots for students in S4D training groups (25 students per S4D training group) in order to ensure a learning effect. The gender ratio is 1:1 with a slightly higher representation of females: 53% females and 47% males in both intervention group and comparison group.

SPORT FOR
**DEVELOPMENT**

For the study, a standardized questionnaire was developed taking into consideration the specific project intervention and local context. After a pilot, the questionnaire was readjusted to fit the age group and cultural context, was translated into Albanian and was digitized. For the data collection, 20 students from the psychology department of the University of Tirana were trained as interviewers to conduct interviews with tablets.

The study was approved by GIZ's data protection unit. The anonymity of the participants is guaranteed by GIZ, and the General Data Protection Regulation by the European Union is applied. As the study's target group are minors, approval for participation was sought by parents/ legal guardians beforehand. Additionally, schools, directors, and staff were informed about the process and the schools' and teachers' approval was obtained to conduct a data collection on their premises.

To analyse S4D's impact on communication, a scale with 15 Likert scale items was created measuring communication competences revolving around four competences: non-verbal communication, verbal communication, subtle communication, and communication in conflict situations. As these competences do not necessarily mutually depend on each other we do not expect Cronbach's Alpha to show internal consistency of the scale. One person can be great in communication verbally but have low non-verbal communication skills. However, we classify all four competences as the construct of communication competences. Cronbach's Alpha in the baseline is 0.422 and in the endline 0.5 – confirming our expectations.

**Reliability Statistics**

| Cronbach's Alpha | N of Items |
|---|---|
| ,422 | 15 |

**Figure 1**

**Reliability Statistics**

| Cronbach's Alpha | N of Items |
|---|---|
| ,500 | 15 |

**Figure 2**

To compare communication competences between the intervention and comparison group and within each group over time, a mixed between-within ANOVA (also called split-plot ANOVA, between-within ANOVA, or mixed factorial ANOVA) was conducted. It was decided not to conduct a MANOVA (Multivariate Analysis of Variance) as the four dependent variables do not relate to each other: communication; self-confidence; cooperation; goal orientation.

We fulfil the prerequisite for conducting mixed between-within ANOVAs with the dependent variables being interval-scaled and the independent variable/ between-subjects factor nominal-scaled with two independent groups. The within-subjects factor is time (two measuring points) and is independent and nominal-scaled. Regarding outliers, the questions were programmed as Likert scales with pre-defined answer options in order to prevent outliers. As ANOVA is a quite robust analytical method against violations of the normality assumption, especially with large sample sizes and balanced designs, the normality assumption can be neglected (Tabachnik & Fidell 2007; Salkind 2010): The sample size for communication is equally distributed among both groups and rather large.

The same applies to variance homogeneity which is tested through Levene's Test of Equality of Error Variances. While variance homogeneity can be neglected with large sample sizes and balanced designs, in our case we still fulfil the assumption of homogeneity. Homogeneity of variances was asserted using Levene's Test based on median which shows that equal variances can be assumed ($p = 0.142$ in the baseline and $p = 0.285$ in the endline). We use Levene's Test based on median as it is more robust then based on mean.

**Levene's Test of Equality of Error Variances[a]**

| | | Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| Communication.1 | Based on Mean | 2,024 | 1 | 348 | ,156 |
| | Based on Median | 2,166 | 1 | 348 | ,142 |
| | Based on Median and with adjusted df | 2,166 | 1 | 344,618 | ,142 |
| | Based on trimmed mean | 2,037 | 1 | 348 | ,154 |
| Communication.2 | Based on Mean | 1,063 | 1 | 348 | ,303 |
| | Based on Median | 1,145 | 1 | 348 | ,285 |
| | Based on Median and with adjusted df | 1,145 | 1 | 344,903 | ,285 |
| | Based on trimmed mean | 1,113 | 1 | 348 | ,292 |

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + IntorControl
   Within Subjects Design: MeasuringPoints

**Figure 3**

The following graph shows a scale from one to four, one representing low communication skills and four depicting comprehensive communication skills. We observe a consistency in the communication skills within the intervention group over time and a decrease within the comparison group. The effect size is small with partial eta squared = 0.012.

**Box's Test of Equality of Covariance Matrices[a]**

| | |
|---|---|
| Box's M | 9,100 |
| F | 3,015 |
| df1 | 3 |
| df2 | 24319019,357 |
| Sig. | ,029 |

As we have a mixed design, we also check for homogeneity in covariance by using Box's Test of Equality of Covariance Matrices. Since the power of Box's Test is dependent on the number of cases, the test becomes more significant the larger the sample is. Some authors therefore recommend not testing the Box's Test at a .05 significance level, but at 0.025 or 0.01 (Mertler, 2004) or 0.001 (Verma, 2015; Warner, 2012). In our case, homogeneity in covariance is violated with Box's Test of Equality of Covariance Matrices showing p = 0.029. However, with a large sample size and a balanced design, this assumption can be neglected.

**Figure 4**

The assumption of sphericity can be neglected, as it only applies for procedures with measurement repetition that have more than two stages. In the present case there are only two measurement points of time.

The mixed between-within ANOVA conducted to assess the impact of Sport for Development on communication competences across two time periods (pre-intervention, post-intervention) shows that there is a statistically significant interaction between group affiliation and time, Greenhouse-Geisser $F(1.00, 348.00) = 4.193$, $p = 0.041$, partial $\eta_i^2 = 0.012$.

**Tests of Within-Subjects Effects**

Measure:   MEASURE_1

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| MeasuringPoints | Sphericity Assumed | ,232 | 1 | ,232 | 4,341 | ,038 | ,012 |
| | Greenhouse-Geisser | ,232 | 1,000 | ,232 | 4,341 | ,038 | ,012 |
| | Huynh-Feldt | ,232 | 1,000 | ,232 | 4,341 | ,038 | ,012 |
| | Lower-bound | ,232 | 1,000 | ,232 | 4,341 | ,038 | ,012 |
| MeasuringPoints * IntorControl | Sphericity Assumed | ,224 | 1 | ,224 | 4,193 | ,041 | ,012 |
| | Greenhouse-Geisser | ,224 | 1,000 | ,224 | 4,193 | ,041 | ,012 |
| | Huynh-Feldt | ,224 | 1,000 | ,224 | 4,193 | ,041 | ,012 |
| | Lower-bound | ,224 | 1,000 | ,224 | 4,193 | ,041 | ,012 |
| Error(MeasuringPoints) | Sphericity Assumed | 18,564 | 348 | ,053 | | | |
| | Greenhouse-Geisser | 18,564 | 348,000 | ,053 | | | |
| | Huynh-Feldt | 18,564 | 348,000 | ,053 | | | |
| | Lower-bound | 18,564 | 348,000 | ,053 | | | |

**Figure 5**

The following graph shows a scale from one to four, one representing low communication skills and four depicting comprehensive communication skills. We observe a consistency in the communication skills within the intervention group over time and a decrease within the comparison group. The effect size is small with partial eta squared = 0.012.
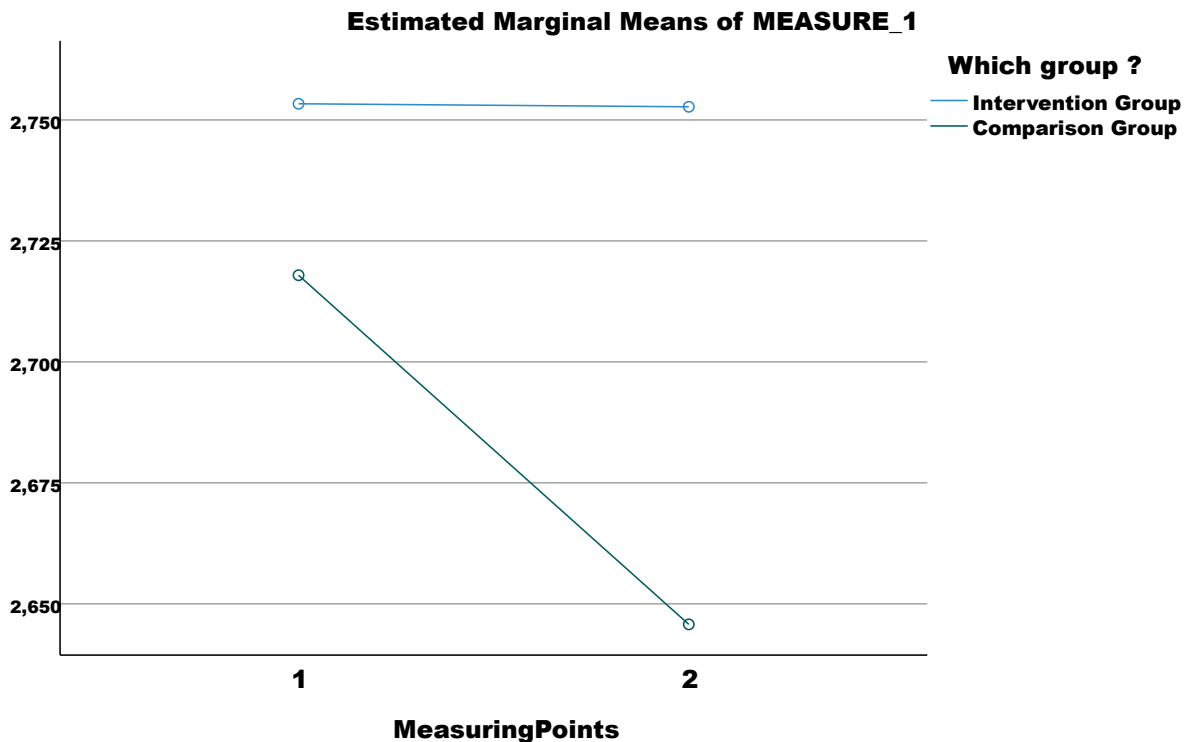


**Figure 6**

The results indicate that communication skills remained stable over time within the intervention group, as shown by consistent scores. This suggests that the intervention has helped maintain communication skills among participants. In contrast, the comparison group showed a decline in communication skills over the same period, suggesting that without the S4D intervention, students' communication abilities deteriorate or do not improve. The stability of communication skills in the intervention group suggests that the S4D intervention was effective in either enhancing or maintaining these skills over time. In contrast, the decline observed in the comparison group indicates that students who did not receive the intervention experienced a reduction in their ability to communicate effectively, which could be due to various factors such as a lack of structured support or practice. This divergence between the two groups underscores the importance of targeted interventions in fostering essential skills like communication, especially in critical developmental stages. Additionally, the results may imply that, without intervention, students are at risk of losing these skills or the confidence to use them over time, potentially affecting their social interactions and academic performance. The consistent gap between the two groups could suggest that the S4D intervention played a protective role, preserving communication skills that would otherwise degrade in its absence.

To analyse S4D's impact on self-confidence, a scale with 25 Likert scale items was created. Cronbach's Alpha shows good internal consistency of the scale for both baseline and endline with values of 0.683 (baseline) and 0.742 (endline).

**Reliability Statistics**

| Cronbach's Alpha | N of Items |
|---|---|
| ,683 | 25 |

**Figure 7**

**Reliability Statistics**

| Cronbach's Alpha | N of Items |
|---|---|
| ,742 | 25 |

**Figure 8**

To compare self-confidence between the intervention and comparison group and within each group over time, a mixed between-within ANOVA was conducted. The normality assumption is neglected as the sample size is large and it is a balanced design. Homogeneity of variances was asserted using Levene's Test based on median which shows that equal variances can be assumed in the baseline (p = 0.905) as well as in the endline (p = 0.358).

**Levene's Test of Equality of Error Variances[a]**

| | | Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| SelfConfidence.1 | Based on Mean | ,009 | 1 | 347 | ,925 |
| | Based on Median | ,014 | 1 | 347 | ,905 |
| | Based on Median and with adjusted df | ,014 | 1 | 342,627 | ,905 |
| | Based on trimmed mean | ,016 | 1 | 347 | ,901 |
| SelfConfidence.2 | Based on Mean | ,839 | 1 | 347 | ,360 |
| | Based on Median | ,848 | 1 | 347 | ,358 |
| | Based on Median and with adjusted df | ,848 | 1 | 345,138 | ,358 |
| | Based on trimmed mean | ,865 | 1 | 347 | ,353 |

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

  a. Design: Intercept + IntorControl
     Within Subjects Design: MeasuringPoints

**Figure 9**

**Box's Test of Equality of Covariance Matrices[a]**

| | |
|---|---|
| Box's M | 1,787 |
| F | ,592 |
| df1 | 3 |
| df2 | 24802725,526 |
| Sig. | ,620 |

**Figure 10**

Checking for homogeneity in covariance in the case of a large sample, as recommended by Mertler¬ (2004), Verma (2015) and Warner (2012), the Box's Test is tested at a 0.001 significance level. The assumption of homogeneity of covariance is asserted with p = 0.620.

The assumption of sphericity can be neglected, as this only applies for procedures with measurement repetition with more than two stages. In the present case there are only two measurement points of time.

The mixed between-within ANOVA conducted to assess the impact of Sport for Development on self-confidence across two time periods (pre-intervention, post-intervention) shows a significant interaction between group affiliation and time, Greenhouse-Geisser $F(1.00, 347.00) = 8.729$, $p = 0.003$, partial $\eta^2 = 0.025$.

**Tests of Within-Subjects Effects**

Measure:   MEASURE_1

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| MeasuringPoints | Sphericity Assumed | ,158 | 1 | ,158 | 2,484 | ,116 | ,007 |
| | Greenhouse-Geisser | ,158 | 1,000 | ,158 | 2,484 | ,116 | ,007 |
| | Huynh-Feldt | ,158 | 1,000 | ,158 | 2,484 | ,116 | ,007 |
| | Lower-bound | ,158 | 1,000 | ,158 | 2,484 | ,116 | ,007 |
| MeasuringPoints * IntorControl | Sphericity Assumed | ,554 | 1 | ,554 | 8,729 | ,003 | ,025 |
| | Greenhouse-Geisser | ,554 | 1,000 | ,554 | 8,729 | ,003 | ,025 |
| | Huynh-Feldt | ,554 | 1,000 | ,554 | 8,729 | ,003 | ,025 |
| | Lower-bound | ,554 | 1,000 | ,554 | 8,729 | ,003 | ,025 |
| Error(MeasuringPoints) | Sphericity Assumed | 22,038 | 347 | ,064 | | | |
| | Greenhouse-Geisser | 22,038 | 347,000 | ,064 | | | |
| | Huynh-Feldt | 22,038 | 347,000 | ,064 | | | |
| | Lower-bound | 22,038 | 347,000 | ,064 | | | |

**Figure 11**

The following graph shows a scale from one to four, one representing low self-confidence and four depicting high self-confidence. We observe a positive development within the intervention group over time and a decrease within the comparison group. This demonstrates that S4D increases self-confidence among youth and has a statistically significant, small effect on self-confidence.
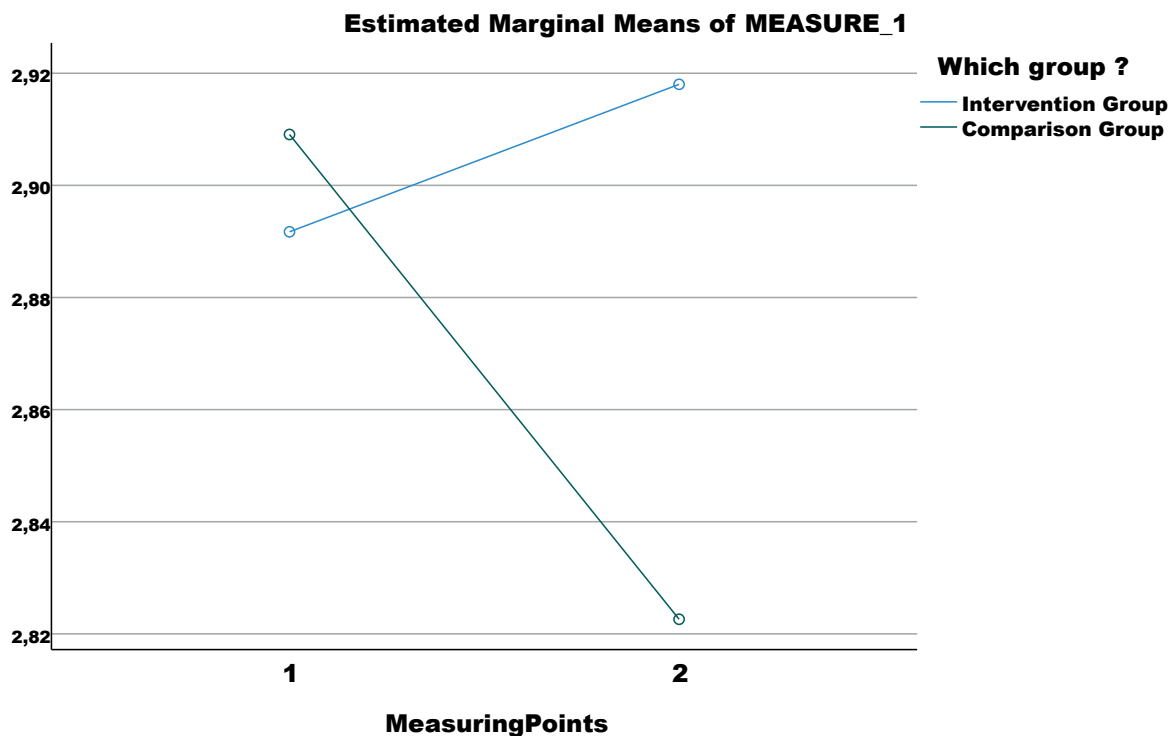


**Figure 12**

The results suggest that the S4D intervention contributed to enhancing self-confidence among the participating youth. In contrast, the comparison group experienced a decline in self-confidence, which implies that without the S4D intervention, students may be vulnerable to losing self-confidence over time. The difference in trends between the two groups highlights the positive influence of the S4D program on personal development. The findings also underscore that while the changes in self-confidence were not large, they were consistent and meaningful. This suggests that even modest improvements in self-confidence through S4D can be valuable for youth development. The decline in the comparison group further emphasizes the role of the intervention in fostering self-confidence. It also raises concerns about external factors that may negatively impact self-confidence in adolescents who are not receiving targeted support. The overall findings demonstrate the effectiveness of S4D in promoting self-confidence, a critical factor for personal growth and success. This makes a strong case for the continued implementation and refinement of such programs aimed at empowering youth.

We measure cooperation competences among youth by creating a scale of 12 variables since cooperation is a latent construct. Cronbach's Alpha shows good internal consistency of the scale for baseline and endline with values of 0.644 (baseline) and 0.727 (endline).

**Reliability Statistics**

| Cronbach's Alpha | N of Items |
|---|---|
| ,644 | 12 |

**Figure 13**

**Reliability Statistics**

| Cronbach's Alpha | N of Items |
|---|---|
| ,727 | 12 |

**Figure 14**

To analyse cooperation competences and cooperative behaviour between the intervention and comparison group and within each group over time, a mixed between-within ANOVA was conducted. The normality assumption is neglected as the sample size is large and it is a balanced design. Homogeneity of variances was asserted using Levene's Test based on median which shows that equal variances can be assumed in the baseline ($p = 0.150$) and in the endline ($p = 0.147$).

**Levene's Test of Equality of Error Variances[a]**

| | | Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| Cooperation.1 | Based on Mean | 2,118 | 1 | 346 | ,146 |
| | Based on Median | 2,076 | 1 | 346 | ,150 |
| | Based on Median and with adjusted df | 2,076 | 1 | 344,864 | ,150 |
| | Based on trimmed mean | 2,132 | 1 | 346 | ,145 |
| Cooperation.2 | Based on Mean | 2,241 | 1 | 346 | ,135 |
| | Based on Median | 2,109 | 1 | 346 | ,147 |
| | Based on Median and with adjusted df | 2,109 | 1 | 342,105 | ,147 |
| | Based on trimmed mean | 2,256 | 1 | 346 | ,134 |

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + IntorControl
   Within Subjects Design: MeasuringPoints

**Figure 15**

**Box's Test of Equality of Covariance Matrices[a]**

| | |
|---|---|
| Box's M | 5,447 |
| F | 1,804 |
| df1 | 3 |
| df2 | 24072607,314 |
| Sig. | ,144 |

**Figure 16**

Checking for homogeneity in covariance in the case of a large sample, as recom¬mended by Mertler (2004), Verma (2015) and Warner (2012), the Box's Test is tested at a 0.001 significance level. With $p = 0.144$ homogeneity in covariance is asserted.

The assumption of sphericity can be neglected, as this only applies for procedures with measurement repetition that have more than two stages. In the present case there are only two measurement points of time.

The mixed between-within ANOVA conducted to assess the impact of Sport for Development on cooperation competences across two time periods (pre-intervention, post-intervention) shows a statistically significant interaction between group affiliation and time, Greenhouse-Geisser $F(1.00, 346.00) = 4.618$, $p = 0.032$, partial $\eta^2 = 0.013$.

**Tests of Within-Subjects Effects**

Measure:   MEASURE_1

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| MeasuringPoints | Sphericity Assumed | ,905 | 1 | ,905 | 9,530 | ,002 | ,027 |
| | Greenhouse-Geisser | ,905 | 1,000 | ,905 | 9,530 | ,002 | ,027 |
| | Huynh-Feldt | ,905 | 1,000 | ,905 | 9,530 | ,002 | ,027 |
| | Lower-bound | ,905 | 1,000 | ,905 | 9,530 | ,002 | ,027 |
| MeasuringPoints * IntorControl | Sphericity Assumed | ,439 | 1 | ,439 | 4,618 | ,032 | ,013 |
| | Greenhouse-Geisser | ,439 | 1,000 | ,439 | 4,618 | ,032 | ,013 |
| | Huynh-Feldt | ,439 | 1,000 | ,439 | 4,618 | ,032 | ,013 |
| | Lower-bound | ,439 | 1,000 | ,439 | 4,618 | ,032 | ,013 |
| Error(MeasuringPoints) | Sphericity Assumed | 32,858 | 346 | ,095 | | | |
| | Greenhouse-Geisser | 32,858 | 346,000 | ,095 | | | |
| | Huynh-Feldt | 32,858 | 346,000 | ,095 | | | |
| | Lower-bound | 32,858 | 346,000 | ,095 | | | |

**Figure 17**

The following graph shows a scale from one to four, one representing low cooperation competences and four depicting high cooperation competences. We observe a small decrease within the intervention group over time and a stronger decrease within the comparison group.
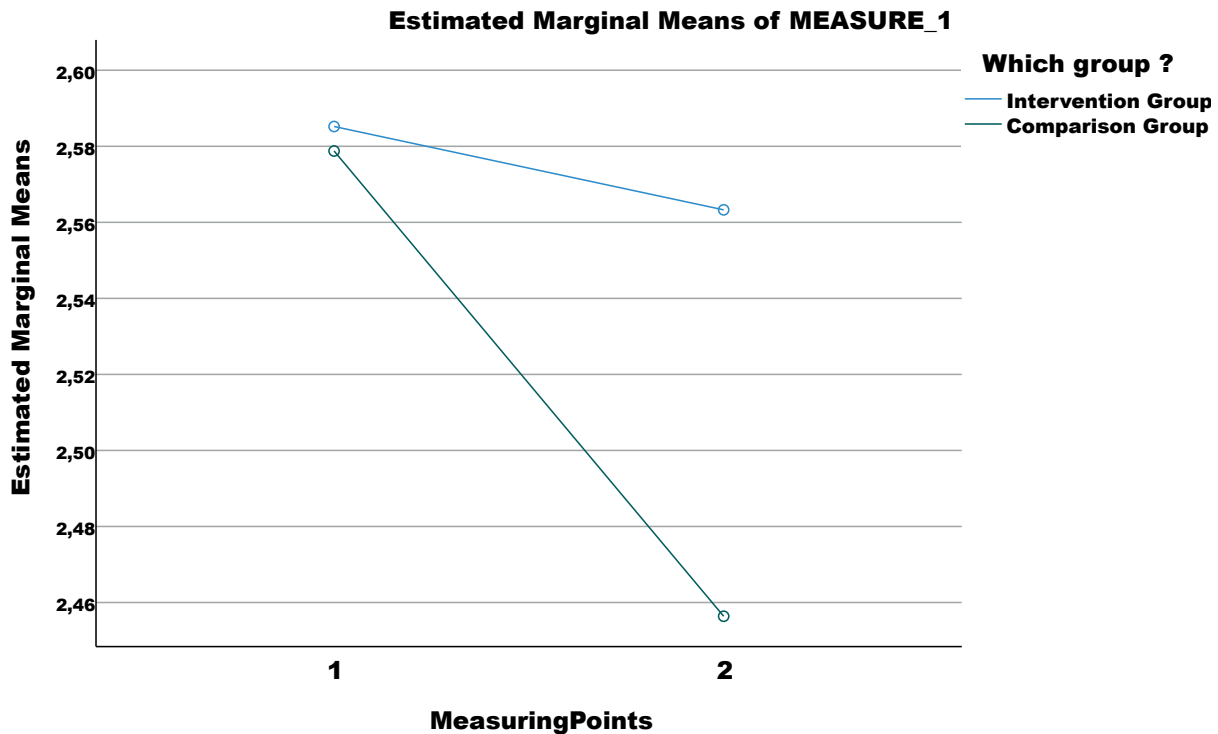


**Figure 18**

The results suggest a decline in cooperation competences over time for both the intervention and comparison group, with the intervention group experiencing a smaller decrease and the comparison group showing a more pronounced decline. The smaller decrease within the intervention group implies that the S4D intervention may have mitigated the loss of cooperation competences, though it wasn't enough to completely prevent the decline.

The stronger decrease in the comparison group indicates that, without the intervention, participants were more susceptible to a significant reduction in their ability to cooperate effectively with others. This suggests that while the intervention had some effect, it may not have been fully sufficient in maintaining or enhancing cooperation competences over time. The results highlight the potential challenges in sustaining such skills in adolescents, even with structured programs in place. Additionally, these findings raise the possibility that other factors, external to the intervention, may have contributed to the overall decline in cooperation competences across both groups. This could include environmental, social, or developmental changes that affect adolescents' abilities to collaborate. Despite the small decline in the intervention group, the results underscore the need for more robust or extended interventions to better support cooperation skills development.

We measure goal orientation by creating a scale of 20 variables since goal orientation is a latent construct. Cronbach's Alpha shows very good internal consistency of the scale for both baseline and endline with values of 0.790 (baseline) and 0.814 (endline).

**Reliability Statistics**

| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
|---|---|---|
| ,790 | ,789 | 20 |

**Figure 19**

**Reliability Statistics**

| Cronbach's Alpha | Cronbach's Alpha Based on Standardized Items | N of Items |
|---|---|---|
| ,816 | ,814 | 20 |

**Figure 20**

To analyse goal orientation between the intervention and comparison group and within each group over time, a mixed between-within ANOVA was conducted. The normality assumption is neglected as the sample size is large and it has a balanced design. Homogeneity of variances was partially asserted using Levene's Test based on median which shows that equal variances can be assumed in the baseline (p = 0.136) but not in the endline (p = 0.010). However, with a balanced design and large sample, this can be neglected.

**Levene's Test of Equality of Error Variances[a]**

| | | Levene Statistic | df1 | df2 | Sig. |
|---|---|---|---|---|---|
| GoalOrientation.1 | Based on Mean | 3,021 | 1 | 346 | ,083 |
| | Based on Median | 2,235 | 1 | 346 | ,136 |
| | Based on Median and with adjusted df | 2,235 | 1 | 337,346 | ,136 |
| | Based on trimmed mean | 2,825 | 1 | 346 | ,094 |
| GoalOrientation.2 | Based on Mean | 7,427 | 1 | 346 | ,007 |
| | Based on Median | 6,738 | 1 | 346 | ,010 |
| | Based on Median and with adjusted df | 6,738 | 1 | 323,899 | ,010 |
| | Based on trimmed mean | 7,358 | 1 | 346 | ,007 |

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

a. Design: Intercept + IntorControl
   Within Subjects Design: MeasuringPoints

**Figure 21**

**Box's Test of Equality of Covariance Matrices[a]**

| | |
|---|---|
| Box's M | 10,933 |
| F | 3,622 |
| df1 | 3 |
| df2 | 24072607,314 |
| Sig. | ,012 |

The assumption of homogeneity of covariance is violated with p = 0.012. Since there is a balanced design and large sample size, this violation can also be neglected.

The assumption of sphericity can be neglected, as this only applies for procedures with measurement repetition that have more than two stages. In the present case there are only two measurement points of time.

The mixed between-within ANOVA conducted to assess the impact of S4D on goal orientation across two time periods (pre-intervention, post-intervention) shows a significant interaction between group affiliation and time, Greenhouse-Geisser F (1.00, 346.00) = 9.294, p = 0.002, partial $\eta^2$ = 0.026.

**Tests of Within-Subjects Effects**

Measure: MEASURE_1

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. | Partial Eta Squared |
|---|---|---|---|---|---|---|---|
| MeasuringPoints | Sphericity Assumed | ,814 | 1 | ,814 | 8,955 | ,003 | ,025 |
| | Greenhouse-Geisser | ,814 | 1,000 | ,814 | 8,955 | ,003 | ,025 |
| | Huynh-Feldt | ,814 | 1,000 | ,814 | 8,955 | ,003 | ,025 |
| | Lower-bound | ,814 | 1,000 | ,814 | 8,955 | ,003 | ,025 |
| MeasuringPoints * IntorControl | Sphericity Assumed | ,844 | 1 | ,844 | 9,294 | ,002 | ,026 |
| | Greenhouse-Geisser | ,844 | 1,000 | ,844 | 9,294 | ,002 | ,026 |
| | Huynh-Feldt | ,844 | 1,000 | ,844 | 9,294 | ,002 | ,026 |
| | Lower-bound | ,844 | 1,000 | ,844 | 9,294 | ,002 | ,026 |
| Error(MeasuringPoints) | Sphericity Assumed | 31,438 | 346 | ,091 | | | |
| | Greenhouse-Geisser | 31,438 | 346,000 | ,091 | | | |
| | Huynh-Feldt | 31,438 | 346,000 | ,091 | | | |
| | Lower-bound | 31,438 | 346,000 | ,091 | | | |

The following graph depicts goal orientation on a scale from one to four, one depicting low goal orientation and four representing high goal orientation. We observe a consistency within the intervention group over time and a decrease within the comparison group.
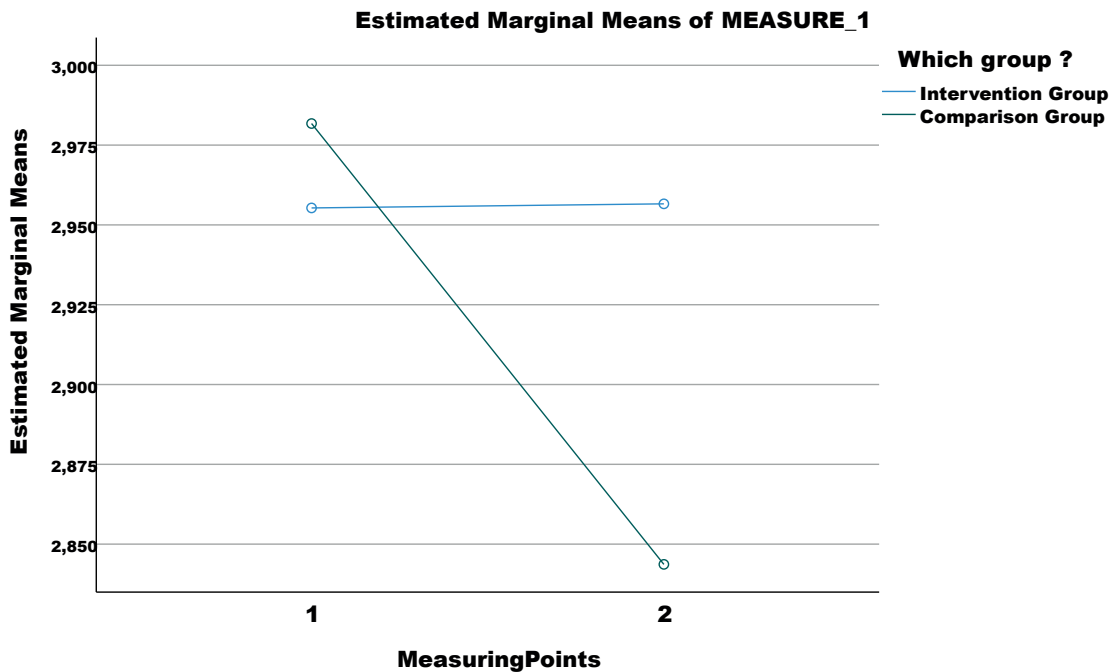
**Estimated Marginal Means of MEASURE_1**



**Figure 24**

The results suggest that goal orientation remained stable within the intervention group, which indicates that the S4D intervention has played a role in maintaining participants' focus and ability to set and pursue goals. In contrast, the comparison group exhibited a decline in goal orientation, suggesting that without the intervention, participants' capacity to stay goal-directed diminished over time. The stability in the intervention group is important, as it highlights the potential effectiveness of S4D in preventing a decline in this critical skill. Goal orientation is essential for academic success and personal development, and the intervention appears to have provided participants with the tools or motivation to sustain it. Meanwhile, the decrease in the comparison group may reflect a lack of structured support or guidance in fostering goal-oriented behaviour.

This difference between the groups underscores the positive impact of the S4D intervention, as it seems to shield participants from the negative trend observed in the comparison group. It also suggests that goal orientation may naturally decrease over time in the absence of targeted interventions, possibly due to environmental or developmental factors. These findings emphasize the importance of programs designed to cultivate and sustain goal-oriented skills, especially during formative years.

With the choice of a quasi-experimental, longitudinal study design instead of a Randomized Controlled Trial (RCT) we cannot exclude the possibility of a selection effect and cannot control for all disruptive factors distorting an unbiased assessment. Randomization was not feasible in this case due to the small sample size, which limited the statistical power required for meaningful random group assignments. Furthermore, dividing the participants into randomized groups could have led to unequal distributions of key variables, thus compromising the integrity of the study outcomes. Given the constraints, the quasi-experimental design allowed us to observe changes over time within the same cohort, though it does come with the trade-off of reduced control over confounding variables. We acknowledge this limitation to our study design and tried to control for it by choosing two points of measurement and ensuring consistency in the S4D implementation. By doing so, Stockmann, R. (2007) argues that there are hardly any differences to a RCT in terms of design quality.

The unequal representation of schools in the sample presents several limitations. Firstly, schools with higher enrollment in the study may disproportionately influence the results, leading to potential bias. This could mean that certain characteristics or interventions that are more prominent in those overrepresented schools will have a larger impact on the overall findings, potentially distorting the results. Secondly, schools with fewer participants may not be adequately represented in the analysis, which could mask any unique contextual factors that might affect the study's generalizability. Finally, the lack of randomization and uneven distribution of students across schools impairs to some extent the study's internal validity, as certain schools may inherently differ in terms of socioeconomic background, teaching quality, or other key variables, which may influence the study outcomes in ways unrelated to the intervention itself. These factors limit the ability to draw clear, causal inferences across the entire population.

Conducting questionnaires with youth always raises the issue of social desirability. We noticed a tendency towards "better answers", high approval rates and higher values in the response behaviour of the participants. This is a well-known phenomenon in social sciences and psychology. Due to social desirability, respondents and especially young persons try to give a predominantly positive description of one's own person and to correspond to what the interviewer or other persons involved supposedly expect from them. This can be done by means of an exaggerated mention of desirable behaviour or by means of an understated mention of undesirable behaviour. Orientation is provided by social norms (Kreuter, F. et al. 2008). Additionally, youth often reflect their response behaviour differently after an intervention, knowing more about the different topics which sometimes even leads to supposedly negative results. In the present study, this may explain to some degree the decreased levels in some competences, and it is assumed that participants reflected their own competences more critically after the S4D intervention. We also find the phenomenon of social desirability. We control for it by having a comparison group and by using quantitative analysis methods that are able to still identify significant results and sufficient effect sizes.

# CONCLUSION

The results across multiple competencies - communication skills, self-confidence, cooperation competences, and goal orientation - demonstrate a consistent trend where the intervention groups either maintained or showed slight improvements in their skills, while the comparison groups generally experienced declines. For communication skills, the intervention group showed stability over time, suggesting the intervention helped maintain these abilities, while the comparison group saw a deterioration. A similar pattern emerged for self-confidence, where the intervention group displayed a positive development and the comparison group declined, highlighting the statistically significant impact of the intervention, albeit a small one.

For cooperation competences, both groups saw a decline, but it was more pronounced in the comparison group, suggesting that while the S4D intervention may have mitigated some loss, it was not fully effective in sustaining these skills. Lastly, in terms of goal orientation, the intervention group remained consistent over time, while the comparison group showed a noticeable decrease, indicating the intervention's role in preventing a decline in this critical area.

Overall, the intervention appears to have played a protective role across all domains, helping to either preserve or slightly improve skills in areas like communication, self-confidence, and goal orientation, even though cooperation competences showed some decline. In contrast, the comparison group, not participating in S4D, experienced more significant decreases, underscoring the importance of S4D programs in fostering and maintaining essential skills during key developmental stages. These results highlight the need for ongoing support and potentially more robust interventions to fully sustain these competencies over time.

The overall findings emphasize the critical role of targeted S4D interventions in supporting youth development across various competencies. While the S4D intervention showed effectiveness in maintaining skills like communication and goal orientation, as well as enhancing self-confidence, the decline in cooperation competences, even within the intervention group, suggests there may be limitations in the scope or intensity of such programs. This points to the need for refining such interventions to better address areas where improvement was minimal or where deterioration still occurred. Moreover, the pronounced decline in the comparison group underscores the potential negative trajectory for adolescents who do not receive structured support. These results make a compelling case for the expansion and strengthening of S4D programs aimed at fostering essential life skills, ensuring that youth are equipped to succeed academically, socially and in the working environment.

Brewer, Laura. (2013). *Enhancing Youth Employability: What? Why? and How? – Guide to Core Work Skills.* Geneva: ILO Skills and Employability Department.

CEDEFOP (2011). *Empowering the young of Europe to meet labour market Challenges.* Findings from study visits 2009/10. Luxembourg: Publications Office of the European Union.

European Commission (2020). *Albania 2020 Report.* albania_report_2020.pdf (europa.eu).

GIZ (2022). *Sport For Employability Manual for Project Developers in the Western Balkan Region.* https://www.sport-for-development.com/imglib/downloads/Manuale/Western%20Balkan/Publikacija%20 SPORT%20FOR%20EMPLOYABILITY_12.2022%20final.pdf.

GIZ (2023). *Sport Creates Change. Impacts of Sport for Development on Gender Equality, Psychosocial Wellbeing and Social Cohesion in the Kurdistan Region of Iraq.* Sport creates change_Final Report KRI (1).pdf.

ILO (n.a.). *ILO Modelled Estimates and Projections database* ( ILOEST ) ILOSTAT. https://ilostat.ilo.org/data/

ILO (2022b). *Decent Work Country Programme 2023–26 Albania.* wcms_865259.pdf (ilo.org).

Kreuter, F., Presser, S, & Tourangeau, R. (2008). *Social desirability bias in CATI, IVR, and web surveys.* Public Opinion Quarterly, 72(5), 847-865.

Mertler, C. A. (2004). *Advanced and Multivariate Statistical Methods: Practical Application and Interpretation* (3rd ed.). Pyrczak Publishing.

RCC (2021). *Study on Youth Employment in the Western Balkans.* Regional Cooperation Council | Study on Youth Employment in the Western Balkans (rcc.int).

Salkind, N. J. (2010). *Encyclopedia of Research Design* (Vol. 2). Los Angeles: Sage.

Stockmann, R. (2007). *Handbuch zur Evaluation.* Eine praktische Handlungsanleitung. Münster: Waxmann Verlag GmbH.

Tabachnik, B.G. & Fidell, L.S. (2007). *Using Multivariate Statistics* (5th ed.). Boston: Pearson Education.

Verma, J. P. (2015). *Repeated Measures Design for Empirical Researchers* (1st ed.). Wiley.

Warner, R. M. (2012). *Applied Statistics: From Bivariate Through Multivariate Techniques* (2nd ed.). Sage Publications Ltd.