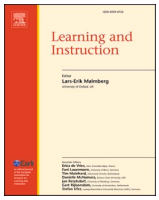
Learning and Instruction 95 (2025) 102042

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc



# Effects of SI-PASS on a high-risk course – A randomized controlled trial

Ninon Puttaert [a,b,*], Dylan Dachet [c,d], Ariane Baye [c,d], Etienne Quertemont [a,e], Laurent Leduc [f,g], Anne-Sophie Nyssen [a,b], Dominique Verpoorten [f,g]

[a] Faculty of Psychology, Speech Therapy and Education Sciences - Department of Psychology, University of Liège, Belgium
[b] Bât. B32 Ergonomics and Work Intervention, Quartier Agora, Place des Orateurs 2, 4000, Liège, Belgium
[c] Faculty of Psychology, Speech Therapy and Educational Sciences - Department of Educational Sciences, University of Liège, Belgium
[d] Bât. B32 Department of Educational Sciences, Quartier Agora, Place des Orateurs 2, 4000, Liège, Belgium
[e] Bât. B32 Quantitative Psychology, Quartier Agora, Place des Orateurs 2, 4000, Liège, Belgium
[f] Institut de Formation et de Recherche en Enseignement Supérieur (IFRES), University of Liège, Belgium
[g] Bât. B63B IFRES, Quartier Urbanistes 1, Traverse des Architectes 5B, 4000, Liège, Belgium

ARTICLE INFO

*Keywords:*
SI-PASS
Randomized controlled trial
Experimental study
Peer-learning
Academic performance
Psychology

ABSTRACT

*Background:* SI-PASS is a structured academic support program employing successful later-year students to facilitate peer-learning sessions attached to high-risk courses, specifically here statistics for psychology at ULiège. The research translates as: How much does this method improve academic performance and impact socio-affective perceptions in first-year students?
*Aims:* This study compares academic performance and socio-affective variables of first-year students in the experimental condition and the control condition.
*Sample:* Participants were 245 freshmen for the experimental trial, and 985 for the quasi-experimental trial.
*Methods:* Participants were assigned to either participate to the SI-PASS scheme or not during the first semester following a randomized controlled trial with stratified random assignment method. Then, the whole cohort enlisted for the course was analyzed to validate supplementary hypotheses.
*Results:* The first step of the study reveals no difference between both groups on academic performance, unless the level of attendance is considered; nor does it identify any significant impact on socio-affective variables. The second step consisted in the comparison of the experimental group and the group of students who chose not to volunteer for SI-PASS and resulted in significant improvement in academic performance in favor of SI-PASS.
*Conclusions:* This delivered significant results in favor of the program but only when attendance is considered, thus offering empirical evidence that a genuine experimental design is likely to mitigate the effects found in a tradition of quasi-experimental designs. These results are valuable for the SI-PASS community, where randomized trials are still scarce, and for higher education institutions seeking evidence-based assistance.

## 1. Introduction

Research shows that the transition from secondary school to university is a critical stage in educational journeys (Johnston, 2010; Mayhew et al., 2016; Kovač, 2015). This shift implies significant academic adjustments (Aspelmeier et al., 2012; Van Rooij et al., 2018; Wasylkiw, 2016) or adaptations (De Clercq et al., 2023) to various essential requirements (Trautwein & Bosse, 2017): students must acclimate to new teaching and learning methods (Noël & Parmentier, 1998), effectively manage a range of resources (Tampakis & Vitoratos, 2009) and services (Lowe & Wright, 2024), and build connections with peers to foster a sense of belonging within the academic community (Brooman & Darwent, 2014; Thomas, 2012). As failure and dropout rates remain very high, up to 60% of a student cohort in French-speaking Belgium (Brunet et al., 2021), supporting academic performance and success remains a major priority for higher education institutions (Upcraft et al., 2005; van der Zanden et al., 2018; Vanthournout et al., 2012), as student attrition carries direct psychological, organizational, and financial consequences (Galand et al., 2005; Jacquemart et al., 2023). Among the available forms of academic support (Delnoij et al.,

* Corresponding author. University of Liège, Bât. B32 Quartier Agora, Place des Orateurs 2, 4000, Liège, Belgium.
*E-mail addresses:* ninon.puttaert@uliege.be (N. Puttaert), ariane.baye@uliege.be (A. Baye), equertemont@uliege.be (E. Quertemont), laurent.leduc@uliege.be (L. Leduc), asnyssen@uliege.be (A.-S. Nyssen), dverpoorten@uliege.be (D. Verpoorten).

https://doi.org/10.1016/j.learninstruc.2024.102042
Received 15 December 2023; Received in revised form 23 August 2024; Accepted 21 October 2024
Available online 3 November 2024

2020; Feldman, 2018), SI-PASS (Supplemental Instruction Peer Assisted Study Sessions) stands out as one of the most widespread programs, with a well-established presence in over 30 countries (Malm, 2021).

### 1.1. Practice of SI-PASS

SI-PASS is a peer assisted learning program in which trained and supervised second-year students (usually called *leaders*) facilitate freshmen's 1-h weekly discussions on the content attached to a "high-risk course" (Martin & Arendale, 1993). The role of the leader is not to teach the group but to organize the interactions so that the students get used to learning together. SI-PASS uses a supplemental and non-remedial approach to learning by providing students with additional opportunities to engage with course material and to develop disciplinary study skills and strategies (Strømmen-Bakhtiar et al., 2021).

### 1.2. Research on SI-PASS effects on performance

In 2022, SI-PASS celebrated its 50th anniversary since its inception at the University of Missouri-Kansas. Throughout the years, much research on peer-learning brought to the program an aura of efficacy regarding academic success. This overarching positive trend is noticeable in two comprehensive studies:

- Dawson et al.'s systematic review (2014) covers the period 2001–2010. Among the 29 studies included, some aim to measure the effectiveness of the SI-PASS on three dimensions of the academic success:
  - Final course grades (sixteen studies): sixteen studies used this dependent variable to compare students who were engaged in the program and those who were not.
  - Course completion (five studies): while no precise count of studies that included this variable is specified, the authors do clarify that some of them resort to a dichotomous variable (pass/fail) for the high-risk course as a primary measurement.
  - Assessment tasks (three studies): three studies focused on the identification of improvements in assessments other than the final marks.
- Bengesai et al.'s meta-analysis (2023) covers the period 2010–2021 with a wider scope. They report that among the 1645 studies on peer tutoring found in the considered period, 21 studies were specifically focused on the SI-PASS program, collectively demonstrating, through a comparative analysis of participant against non-participant groups in quasi-experimental designs, a moderate impact of the program of Hedges' $g = +.40$ [.30-.50] on achievement.

Besides this positive perspective on the effectiveness of the SI-PASS program, both reviews highlight significant methodological flaws. They report inadequate control of the measured variables and emphasize the need for more robust experimental designs. Dawson et al. argue that, despite the listed positive effects, "none is supported by a gold standard study involving random assignment to groups and sufficient detail about methodology, participants, and the SI intervention in practice" (2014, p. 635). Similarly, Bengesai et al. express caution about the existing research, "Despite the optimism surrounding peer learning in recent years and the increase in the number of studies, very few control studies have evaluated its effect on academic performance in the past decade" (2023, pp. 12–13).

Between 2001 and 2010, Dawson et al. (2014) identify only one study with a proper controlled design: in 2009, Parkinson (2009) created an experimental and a control group based on voluntary participation. However, the contribution was based on a relatively small sample ($n = 63$) and the SI-PASS meetings described were not connected to a "high-risk" specific course but to the whole program of study for a degree in Biotechnology. Even if these studies report positive results, Dawson et al. (2014) highlight the difficulty of verifying their results due to lacks in the methodological information provided. Only seven studies allow to compute effect size. These effects ranged from $d = .29$ to $d = .60$. Note that these quasi-experimental studies do not account for the self-selection bias. The only experimental study cited (Parkinson, 2009) provides partial methodological information, both on the numbers of participants and on the criteria for assigning them to the experimentation. Finally, Dawson et al. (2014), point out the conflicts of interest of many authors of the studies carried out.

To be complete, Paloyo et al. (2016) conducted a study at an Australian university that assessed an SI-PASS program, a study that was neither included by Bengesai et al. (2023) nor by Dawson et al. (2014): a randomly selected subgroup of students, enrolled in fourteen first-year courses, over three semesters, was encouraged to attend SI-PASS sessions through an incentive (here, a chance to enter a draw and win a gift certificate of AUD 5000). The study found that students who attended the sessions had increased final grades and were more likely to pass the course than students who did not attend the sessions. This design, however, cannot be compared to the present study which sticks to the common SI-PASS principle of fully voluntary participation.

To ensure thorough coverage, the authors inspected literature produced from 2021 (the endpoint of Bengesai et al.'s review) to August 2023. This effort left unchanged Bengesai et al.'s conclusion that "very few control studies have evaluated the effect of SI-PASS on academic performance during this past decade" (2023, p. 13), with one noticeable exception: Dekker et al. (2023) who, through a sound controlled randomized trial, obtain significantly higher grades in favor of SI-PASS participants. While the methodology is similar to the present study, differences arise in the following aspects: a) focusing on a specific course, calculations were conducted across a mix of 10 courses, b) the research was conducted in an institution that would not be categorized as a "university" in Belgium, but as "vocationally-oriented" higher education, c) participants were offered seven meetings of 100 min, in contrast to twelve meetings of 60 min, and d) the average age of subjects was 20.57 years, as opposed to 18.76 in the present study.

### 1.3. Research on SI-PASS effects on socio-affective dimensions

Dawson et al.'s systematic review (2014) identifies five studies demonstrating the beneficial influence of the SI-PASS program on students' well-being and sense of belonging/connectedness to the university. By providing additional support and guidance, the scheme helps alleviate stress and anxiety, thereby improving the overall well-being for students (Abegglen & Morris, 2015). In Martin and Arendale (1992) or Bronstein (2007), students engaging in peer-led sessions expressed decreased anxiety and increased confidence in their academic skills. These studies also revealed the program's heightened effectiveness among students at risk of dropout or those with a history of low academic achievement before enrollment (Hafer, 2001; Stansbury, 2001). Furthermore, Dobbie and Joyce (2008) proposed that students' increased comfort in asking questions freely during SI-PASS sessions contributes to these positive outcomes.

SI-PASS can also help foster a sense of community and connection among students, as these sessions are typically held in a group setting which provides students with the opportunity to interact with peers and build relationships with other students taking the same course. This sense of community can help enhance social integration, which is important for student well-being (Hayes & Fulton, 2019). This aspect is underscored in Verpoorten et al. (2021) with a notable 70% of students reporting SI-PASS meetings as positive experiences for socialization. Regarding students' engagement with the discipline, a study by Hockings (2009) showed that students who participated in peer-learning programs reported feeling more engaged and motivated in the course. The study also found that the program was particularly effective for students who were struggling with the course material or who had low levels of motivation prior to enrolling in the course. By providing

additional support and guidance, peer-assisted learning can help to improve student engagement and motivation, which can have a positive impact on overall well-being (Hager, 2018). However, in contrast to the aforementioned observational studies, the randomized controlled trial conducted by Dekker et al. (2023) did not unveil significant differences between groups concerning well-being and the sense of belonging scores.

### 1.4. Geographical and disciplinary gaps in SI-PASS research

The need for more research in diverse geographical regions, where SI-PASS is introduced, pertains both in Dawson et al.'s state-of-the-art for the studied period ("The majority of studies using entry scores were conducted in an American context", 2014, p. 623) and, for the subsequent timeframe, in Bengesai et al. ("[…] most of the studies were conducted in the USA […]", 2023, p. 13). Again, Dekker et al. (2023) represents a noticeable exception as it takes place in the Netherlands.

Disciplines and content domains also call for refined research on SI-PASS. Regarding statistics for psychology (the high-risk course selected for this study), the research landscape appears rather depleted. Only one research on SI-PASS in the context of a psychology course has ever been documented (Guarcello et al., 2017). Two quasi-experimental studies specifically focused on statistics were conducted, but for a statistical class in a business school (Szal & Kennelly, 2017) and for a chemistry class (Miller, Oldfield, & Bulmer, 2004). Two of the three experimental studies involving randomly assigned students were carried out, for one, within the disciplines of chemistry and mathematics (Parkinson, 2009), and, for the other, in an education faculty with no included course related to psychology or statistics (Dekker et al., 2023). The third one (Paloyo et al., 2016) mentions three courses in statistics for business and one in statistics for psychology. However, as the data combines fourteen courses, no specific metrics for this course is available.

### 1.5. Hypotheses

The distinctive contribution of this study is threefold: a randomized controlled trial targeting a course of statistics for psychology in a Belgian context. Four hypotheses were tested.

#### 1.5.1. Hypothesis 1: Effects on academic performance

**Hypothesis 1a**. Freshmen from the experimental group will perform significantly better at the final exam than those from the control group. Hypothesis 1a was tested controlling for students' motivation, since participants from both groups volunteered for the SI-PASS scheme. Hypothesis 1b: Students from the experimental group will also outperform students who opted not to participate (freshmen in the same cohort who did not volunteer for the experiment). The combination of these two predictions helps determine if experimental and quasi-experimental methods will produce convergent or contrasting outcomes.

#### 1.5.2. Hypothesis 2: Effects of the quality of implementation

**Hypothesis 2a**. Students from the experimental group who attended a minimum of four meetings are expected to demonstrate significantly better performance compared to those with lower attendance. Early concerns, as noted by McCarthy et al. (1997), highlight the limitations of binary variables for attendance (i.e., attending or not attending), suggesting the adoption of a discrete variable approach. Following this recommendation, Cheng and Walters (2009) reported that attending all meetings in their setting corresponds to a ten-fold increase in the odds of academic success. Previous quasi-experimental studies conducted at ULiège indicated a potential effectiveness threshold of six or seven participations. However, since these studies were kept internal, Hypothesis 2a chooses to adopt the figure presented in a published paper by Gattis (2002) which asserts that individuals who attended four or more SI-PASS meetings performed better than those who opted not to attend. Beyond this tentative cut point of four meetings, hypothesis 2a also operates a correlational analysis between discrete attendance levels and final exam results. To consider the nested nature of the measures, an additional multilevel analysis was conducted to explore possible influence of group ownership in relation to a Hypothesis 2b, which posits that exam performance varies across different SI-PASS groups. This analysis aims to understand how group dynamics or leader-dependent aspects influence outcomes.

#### 1.5.3. Hypothesis 3: Effects on socio-affective dimensions

Students from the experimental group will perceive a significant improvement in their psychosocial well-being and academic adjustment. Along with the five studies mentioned in Dawson et al.'s systematic review (2014), the study inspects these effects through "socio-affective" variables. It allows to simultaneously deal with general well-being (positive/negative affects and somatization) and students' academic adjustment, including belonging/connectedness to the university.

#### 1.5.4. Hypothesis 4: Effects on academic engagement

**Hypothesis 4a**. Students attending SI-PASS sessions will use more intensively than the control group the online exercises recommended by the teacher. This hypothesis is introduced to get an insight about the place of a SI-PASS scheme in the larger learning "ecology" (Barron, 2004; Normak et al., 2012; Siemens, 2003) of a course and to explore patterns of engagement (Cole & Spence, 2012; Howard et al., 2019; Meehan & McCallig, 2015) with two different learning aids.

**Hypothesis 4b**. Similar to Hypothesis 1b, an extra comparison is performed between SI-PASS participants and those who also signed up for the course but opted not to participate in the scheme. Likewise, it allows to see whether the experimental and the quasi-experimental methodologies yield similar or divergent results.

## 2. Methodology

### 2.1. Intervention

ULiège has a four-year experience in implementing regular SI-PASS schemes across four faculties (Verpoorten et al., 2021; Verpoorten & Jérôme, 2022). The groups of voluntary students are supervised by 2nd or 3rd year students (SI-PASS leaders), who undergo an initial 6-h training, receive a handbook outlining expected roles, activities, and mindset, and get reflective exercises and debriefings as in-service training, To secure a fair level of adherence to the SI-PASS specifics, two certified program supervisors conduct site visits to each leader during the semester, using an observation grid (Appendix 1) grounded in recommendations and quality standards from American and European SI-PASS centers. Analysis of these field observations over the years reveals minimal variability, indicating a good understanding and application of the instructions by the leaders. To further support compliance, one individual and two collective debriefings take place within the semester. Those monitoring actions allow assuming a reliable level of implementation fidelity (James Madison University, 2023), aligning with the fundamental tenets of the SI-PASS approach. Feedback questionnaires, collected from both leaders and participants, corroborate overall compliance to the unique structured form of peer learning that SI-PASS embodies.

### 2.2. Targeted high-risk course

At the Faculty of Psychology, the course of "Descriptive and inferential psychostatistics is known by students as a "killer course" with a high rate of failure. The lecturer is aware of this and has, over the years, in addition to the 2-h lecture per week, created a supplemental aid for students: weekly online exercises were offered, featuring questions akin to those anticipated in the final exam. A total of 96 exercises were

distributed across twelve chapters. Immediate automatic feedback was provided to students, helping to identify and correct mistakes. Students had the freedom to attempt these online exercises at their convenience. Despite those efforts, the lecturer noticed that a significant number of students still appeared underprepared (in terms of "time on task") for the final exam. This is the reason why he welcomed a SI-PASS scheme in his course, without canceling the online exercises. Hypothesis 4 explores the complementarity/competition of this aid with SI-PASS sessions.

### 2.3. Schedule and participants

The experiment (Ethical Approval 2122-098) began on the welcome day of the academic year 2022–2023, introducing freshmen from the Faculty of Psychology to the SI-PASS program and the details of the upcoming experiment. All students interested in taking part in the study were informed that they would be allocated either to the experimental group (EG) which would benefit from the scheme during the first semester or to the control group (CG) which would participate to the program during the second semester. Much understanding was expressed by the students for two reasons: on the one hand, as future psychologists, they understood the value of a rigorous experiment and were keen on taking part in it as participants. On the other hand, students assigned to the control group were warranted a priority for participating in the SI-PASS scheme taking place in the next semester in a different course. Doing so alleviated a common criticism addressed to randomized trials, that is depriving a group of a possibly beneficial pedagogical intervention. Everything being clear, the candidates took the pre-test (Appendix 2) before being randomly assigned to the EG or CG. Students from the EG were also randomly assigned to one of the thirteen groups headed by one of the thirteen hired leaders. As with all SI-PASS programs at ULiège, the groups were designed to be stable: students assigned to a specific leader were not allowed to switch groups during the semester. This policy aims to foster a sense of identification among group members, including the leader, and to cultivate mutual trust and consistent working habits deemed to develop a sense of belonging. The fixed groups convened twelve meetings throughout the semester, adhering to a recommended schedule where each course session was followed by a meeting later that week to discuss the course content.

### 2.4. Research design

The research implements a randomized controlled trial (RCT) with a stratified random assignment and a waiting list control group (Slavin, 2007) based on a convenience sample of students interested in participating in the intervention. The randomization of volunteers addresses the primary criticism raised by McCarthy et al. (1997) and reiterated by Dawson et al. (2014, p. 611) regarding motivation and the danger of self-selection bias in experiments: "More than a decade after McCarthy et al.'s (1997) critique of the literature, there is still limited research that controls for motivation, as distinct from prior academic achievement, when evaluating the effectiveness of supplemental instruction". In this case, both experimental and control group included students wishing to participate in the intervention. Additionally, in line with Dawson et al.'s (2014) recommendation to include students from diverse backgrounds, the study considered the following variables linked to students' profiles during stratified randomization to form equivalent pairs of students: gender, socio-economic status, and logical reasoning abilities. Students were not tested after the second semester because the targeted course was different from the course chosen in the first semester, both in terms of its subject matter (*i.e.* it did not deal with statistical analyses) and its nature (*i.e.* it did not systematically involve exercises and practical work) Such comparison would not have made much sense.

The dependent variables are the grade obtained by students at the final exam (Hypotheses 1 and 2), students' perceptions on academic adjustment and well-being (Hypothesis 3), and the advantage that students took off the other aid (online exercises) offered in the course during the semester (Hypothesis 4). The independent variable is the exposure to a SI-PASS scheme offered in the high-risk course of statistics for psychology. A sensitivity analysis in G*Power indicated that with a power of .80, a sample size of 245 (120 in control group and 125 in treatment group) would detect group differences with an effect size of Cohen's $d = .36$ (*t*-test, difference between independent group means, two-tailed test and an alpha of .05). With the current sample size of 184 students (92 in control group and 92 in treatment group) after attrition, the detectable effect size increases to .42.

### 2.5. Data sources

#### 2.5.1. Sociodemographic data

The pretest collected self-reported data on gender, age, and profession of both parents. Socio-economic status (SES) was inferred from the highest occupation of either the mother or father, using the one-digit ISCO-08 code (ILO, 2012) and converted into four categories.

#### 2.5.2. Socio-affective data: SACQ

The pretest encompassed an adjusted French-translated version (Carayon & Gilles, 2005) of the "Student Adaptation to College Questionnaire" (SACQ, Dahmus et al., 1992), intended to measure possible influences of SI-PASS in the areas of social well-being and academic adjustment. Considering possible concerns related to a potential North American-centric bias, all item formulations were double-checked by linguistic and education experts. The final questionnaire (Appendix 3) was structured as follows:

- Academic Adjustment: evaluates the student's ability to face academic standards.
- Application: perception of efforts put into academic work.
- Performance: perception of academic achievements.
- Environment: satisfaction towards academic environment (this dimension is only present in the post-test because its items made no sense at the very start of a first year at university).
- Social Adjustment: evaluates the student's ability to face social standards.
- Attachment: perceived level of engagement towards the academic institution.
- Personal-emotional adjustment: this dimension was replaced by three dimensions from the Multidimensional and Systemic Survey on Climate and Well-being at School (MSSCWS) – upper secondary students (Francotte et al., 2023), namely positive affects/negative affects, climate, and somatization. This substitution was necessary since this SACQ sub-scale probes into the respondent's personal experience at the university, which would have been incongruent given that the pre-test was conducted on their official first day at university.

#### 2.5.3. Logical thinking: TOLT

Since the intervention was implemented within a statistics course, controlling students' logical skills and proficiency in probabilities, two predictors of students' performance in statistics (Jiang, XU, Garcia, & Lewis, 2010; Vázquez & de Anglat, 2009) was the most relevant measure one could take. The "Test of Logical Thinking" (TOLT, Tobin & Capie, 1981) was administered, after personal contact with the authors, as a section of the baseline pretest. The ten-items questionnaire was designed to gauge formal reasoning ability within a limited completion time. Internal consistency analyses ($KR\text{-}10 = .87$) attests for the good reliability of the TOLT test.

#### 2.5.4. Participation data

Throughout the semester, each SI-PASS leader was tasked with maintaining a detailed record of student participation. This documentation served to monitor attendance, tally the total number of sessions

attended by each student, and pinpoint the specific meetings they participated in. This data collection was essential for analyzing and testing Hypothesis 2. Students' involvement in the weekly online exercises was monitored through the institutional Learning Management System. Basic learning analytics allowed for the calculation of individual students' participation rate (the number of exercises each attempted to solve).

#### 2.5.5. Final exam grades

The final exam grades (scores from 0 to 20), provided by the course lecturer, were used to measure the students' achievement in the targeted course. The Cronbach's reliability coefficient for the 9-items exam is $\alpha = .84$, thus indicating good internal consistency.

#### 2.5.6. Post-test on socio-affective data

The post-test is identical to the pre-test, excluding only demographic items and the TOLT test. The comparison of pre- and post-test for the French version of the SACQ (Carayon & Gilles, 2005) yielded satisfactory internal consistency scores (Appendix 4), confirming the good psychometrical properties of this slightly adapted instrument. Similar analyses were applied to the MSSCWS scale dimensions producing consistent Cronbach alphas (Appendix 5).

### 2.6. Data processing

All statistical procedures were executed with a double-blind approach, as both researchers did not know which conditions was experimental and decided to analyze data on their own before comparison. Minor disparities noted between the two independent researchers (in terms of rounding of found value to the hundredth) were attributable to the statistical tools employed (SAS 9.4 and Jamovi 2.3.21 with packages 'JMV', 'Descriptives', 'corrMatrix', 'ttestIS', 'contTables') and to the management of attrition.

## 3. Results

### 3.1. Analytical sample

On the welcome day (September 14th), 245 students expressed interest in being part of the SI-PASS scheme attached to the targeted high-risk course and took the pretest (Appendix 9). Among those, 184 passed the final exam (92 in each group). For Hypothesis 1 analyses, all the students who took the final exam, even those who did not take the post-test on socio-affective variables as well (26 in the control group and 15 in the experimental group) were kept to maximize power. The post-test on socio-affective variables was filled in by 160 students (86 in the experimental group, 74 in the control group). For Hypothesis 3, only the subjects with complete datasets on socio-affective variables were kept. This loss of participants impacted in a non-significant way the groups equivalence established at baseline (Tables 1 and 2). The large proportion of women in the sample reflects a common reality in Belgian psychology faculties. Correlation tables can be found in appendixes 6 and 7, confirmatory factor analyses for the two scales in pre-test and post-test can be found in appendix 8. Results of the confirmatory factor analysis performed on the pre-test display a $X^2$ (1106) = 2163.6, $p < .001$, CFI = .634, TLI = .61, SRMR = .096, RMSE = .078; and on the post-test a $X^2$ (1297) = 2494, $p < .001$, CFI = .678, TLI = .658, SRMR = .097, RMSE = .078.

### 3.2. Hypothesis 1 – Academic performance

To test Hypothesis 1a, a *t*-test was performed between the experimental and the control group. Results show that the model is nonsignificant ($t(182) = -.18$, $p = .86$), with a higher mean in favor of the experimental group ($M = 6.48$, $SD = 4.36$). Effect size is Hedges' $g = .03$. These results do now allow the rejection of the null hypothesis. When contrasting the academic performance of the experimental group with that of all students enrolled in the course (Hypothesis 1b) who took the exam but who did not volunteer for the scheme, the *t*-test, this time, displays a significant result: ($t(696) = 3.15$, $p = .002$). Effect size is Hedges' $g = .26$.

Taking the mark at the exam as dependent variable, a further multiple linear regression was performed, using the following variables as predictors: experimental group, SES category, TOLT score. Results show that the general model is statistically significant ($F(3,180) = 20.87$, $p < .001$) with an $R^2 = .26$. However, when taken apart, only the TOLT score reaches statistical significance ($F(1) = 63.15$, $p < .001$). The other two predictors, namely the experimental group and the socioeconomic status

**Table 1**
Groups baseline equivalence on sociodemographic variables used for the stratified randomization.

| Characteristic | *n* | Control group Sample % or *M (SD)* | Treatment group Sample % or *M (SD)* | Test | *p* |
|---|---|---|---|---|---|
| Gender | 182 | 7.69 % | 12.09 % | $X^2(2, 182) = 1.30$ | .52 |
| Age | 184 | 19.25 (4.18) | 18.36 (1.13) | t(104)* = 1.97 | .05 |
| SES | 184 | 1.87 (1.02) | 1.99 (1.02) | t(182) = .43 | .43 |
| TOLT (Score) | 184 | 8.85 (3.89) | 9.27 (4.20) | t(182) = .48 | .48 |
| TOLT (Success rate) | 184 | 55.43 % | 57.61 % | $X^2(1, 184) = .09$ | .77 |

*Note*. The percentages indicated for the gender variable represent the percentage of boys in our sample; the percentages indicated for the TOLT (Success Rate) variable represent the percentage of success in our sample; * Welch's approximation was preferred to the classic Student's *t*-test to take unequal variances into account.

**Table 2**
Groups baseline equivalence on socio-affective variables.

| Dimension | *n* | Control group *M (SD)* | Treatment group *M (SD)* | *t* | *p* |
|---|---|---|---|---|---|
| Performance | 156 | 16.5 (3.35) | 16.0 (3.13) | *t*(154) = .99 | .33 |
| Environment | NA | NA | NA | NA | NA |
| Social Adjustment | 154 | 19.0 (4.72) | 20.3 (4.63) | *t*(152) = 1.77 | .08 |
| Application | 157 | 29.3 (3.57) | 29.2 (3.65) | *t*(155) = .23 | .82 |
| Attachment | 158 | 19.0 (5.72) | 18.0 (5.72) | *t*(156) = 1.01 | .31 |
| Affects | 157 | 20.2 (4.43) | 20.1 (4.64) | *t*(155) = .15 | .88 |
| Climate | 158 | 18.6 (5.09) | 18.0 (4.22) | *t*(156) = .82 | .41 |
| Somatization | 156 | 23.3 (5.14) | 22.6 (5.54) | *t*(154) = .82 | .41 |

*Note.* NA = not applicable and relates to the sub-dimension that was only measured at post-test. Maximum *n* was used here to best represent participants' data.

**Table 3**
Linear Regression of the independent variables Group (SI-PASS vs. Control), Socio-Economic Status (SES) and logical thinking (TOLT) on the dependent variable "performance at the final exam".

| Predictor | Estimate | *SD* | *t* | *p* |
|---|---|---|---|---|
| Intercept | 1.44 | 1.07 | 1.36 | .18 |
| Intervention (SI-PASS 2 – Control 1) | .11 | .59 | .19 | .85 |
| SES | .07 | .3 | .25 | .81 |
| TOLT | .57 | .08 | 7.55 | <.001 |

score fail to reach statistical significance (Table 3).

### 3.3. Hypothesis 2a – Attendance level

The linear regression calculated on participation in SI-PASS meetings exhibits a positive correlation ($r(90) = .28$, $p = .007$) between students' attendance at SI-PASS meetings and their exam performance (Fig. 1).

Linear contrasts performed between all levels of attendance show that students must have attended at least 5 meetings to see their academic performance significantly improve ($t(90) = 2.33$, $p = .022$) with a mean difference of 2.08 on a 20-point scale. Fig. 2 displays the mean differences between the control group (0; $n = 92$), those who attended fewer than 5 meetings (1; $n = 40$, $M = 5.3$, $SD = 3.82$), and those who attended at least 5 (2; $n = 52$, $M = 7.38$, $SD = .63$).

### 3.4. Hypothesis 2b – Group ownership

The multilevel analysis does not detect any differences when group ownership is considered, as reflected in the dotplot (Fig. 3) where each individual grade is given as a dot and the group mean grade as a diamond. Table 4 presents results of multilevel analyses conducted. It shows that there are no between group (leader) variance and that the only independent variable that influences students' exam results is the score these same students obtained on the TOLT, the logic test used as a pre-test. Analyses were performed on R (version 4.3.1) using the lme4 package (Bolker, 2024) and following the recommendations of Finch et al. (2019).

### 3.5. Hypothesis 3 – Socio-affective dimensions

Table 5 displays the effect sizes of all dimensions in the post-test when comparing the two groups. No $p$ value reaches significance level, thus failing to reject the null hypothesis.

### 3.6. Hypothesis 4 – Patterns of engagement

No statistical difference was found in engagement with the online exercises (number of online exercises completed, irrespective of the accuracy of the answers) between the groups ($t(182) = .48$, $p = .63$). Yet, a correlation matrix calculated for SI-PASS students only reveal a significant relationship between their attendance at meetings and their interaction with online exercises ($r(90) = .33$, $p = .001$) (the more students attend meetings, the more they engage with online exercises, and/or the more students engage in online exercises, the more they attend meetings). Students attending more meetings also engage more frequently with online exercises (Table 6). As expected, there's a robust and significant correlation between students' interaction with online exercises and their exam performance ($r(184) = .75$, $p < .001$).

An exploratory analysis, comparing students who participated in the program to all other students enrolled in the targeted course, was conducted for this hypothesis similarly to the first one. This time, the $t$-test between SI-PASS students ($n = 184$, $M = 24.31$, $SD = 19.97$) and all other students ($n = 801$, $M = 10.76$, $SD = 17.54$) turns out to be significant ($t(983) = 9.2$, $p < .001$, $d = .72$). Results with these two groups are again different from the ones obtained within the randomized controlled trial.

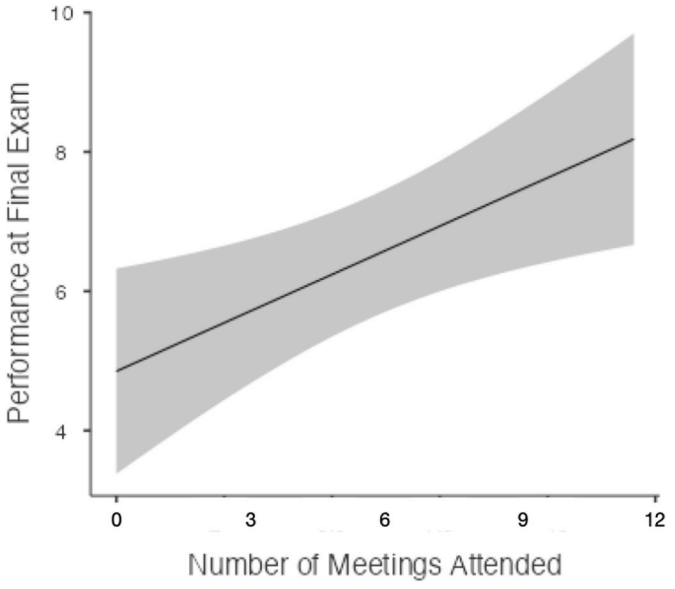



**Fig. 1.** Correlation between Performance at Final Exams and Number of attended Meetings within the Experimental Group.

**Table 4**
Two-level regression analyses of treated effects of performance at exam.

| Effect | Parameter | Performance at exam | | | |
|---|---|---|---|---|---|
| | | Model 1 | Model 2 | Model 3 | Model 4 |
| Fixed effects | | | | | |
| Intercept | $\gamma_{00}$ | 6.42 (.34) | 6.24 (1.07) | 1.38 (1.11) | 1.56 (1.31) |
| Intervention (=1) | $\gamma_{01}$ | | .12 (.68) | −.12 (.59) | −.11 (.59) |
| Tolt | $\gamma_{02}$ | | | .58*** (.07) | .57*** (.08) |
| HISEI | $\gamma_{03}$ | | | | −.07 (.30) |
| Random effects | | | | | |
| Leader variance | $e_{0ij}$ | .00 (.00) | .00 (.00) | .00 (.00) | .00 (.00) |
| Student variance | $\mu_{0j}$ | 20.97 (4.58) | 21.08 (4.59) | 15.73 (3.97) | 15.82 (3.98) |
| Total variance | $e_{0ij} + \mu_{0j}$ | 20.97 | 21.08 | 15.73 | 15.82 |
| Goodness of fit | | | | | |
| AIC | | 1087.1 | 1096.7 | 1036 | 1038.2 |
| BIC | | 1089.0 | 1101.9 | 1052 | 1057.4 |
| Deviance | | 1081.1 | 1081.0 | 1026.2 | 1026.2 |
| Model of reference | | | Model1 | Model1 | Model 1 |
| $\chi 2$ fit improvement (Df) | | | .03 (1) | 54.85 (2) | 54.91 (3) |
| *p-value* | | | .86 | <.001 | <.001 |

Note. Standard errors are presented in parentheses. All p values in this table are two-tailed.

**Table 5**
Comparison of Control and Experimental Groups in Post-Test on socio-affective variables.

| Dimension | $n$ | Control $M$ $(SD)$ | Treatment group $M$ $(SD)$ | $t$ | $p$ |
|---|---|---|---|---|---|
| Performance | 159 | 16.3 (3.52) | 15.6 (3.36) | $t(157) = 1.12$ | .27 |
| Environment | 160 | 15.04 (2.37) | 14.44 (3.06) | $t(158) = 1.37$ | .17 |
| Social Adjustment | 157 | 21.46 (4.78) | 21.69 (4.54) | $t(155) = -.31$ | .76 |
| Application | 159 | 16.22 (5.85) | 16.51 (5.06) | $t(157) = -.34$ | .74 |
| Attachment | 160 | 27.46 (5.13) | 27.48 (4.94) | $t(158) = -.02$ | .98 |
| Affects | 155 | 22.31 (4.60) | 21.75 (4.61) | $t(153) = .75$ | .45 |
| Climate | 151 | 16.79 (4.54) | 16.64 (4.69) | $t(149) = .2$ | .84 |
| Somatization | 153 | 23.66 (4.58) | 22.80 (5.76) | $t(151) = 1.01$ | .32 |

Note. Maximum n was used here to best represent participants' data.

**Table 6**
*Means of Online Exercises done according to Meeting Attendance.*

| Meetings Attendance | *n* | *M* | *SD* |
|---|---|---|---|
| 0 | 15 | 9.8 | 10.9 |
| 1 | 9 | 18.45 | 19.23 |
| 2 | 7 | 23.95 | 23.44 |
| 3 | 3 | 25.13 | 19.58 |
| 4 | 6 | 19.08 | 15.25 |
| 5 | 3 | 28.25 | 10.58 |
| 6 | 6 | 11.86 | 19.16 |
| 7 | 2 | 34.62 | 31.24 |
| 8 | 5 | 35.35 | 15.58 |
| 9 | 8 | 39.63 | 28.1 |
| 10 | 9 | 19.6 | 13.8 |
| 11 | 12 | 35.47 | 25.1 |
| 12 | 7 | 25.61 | 15.05 |

## 4. Discussion

Research on SI-PASS effects lacks studies employing randomized trial experiments. The present study developed such an experimental design deemed to get more valid results than most of the prevailing quasi-experimental studies. Results related to each hypothesis are now discussed.

### 4.1. Hypothesis 1 – Academic performance

In contrast to the predominantly quasi-experimental methodological paradigm of the field (Armentor, 2019; Fayowski & MacMillan, 2008; Moore & LeDee, 2006), a rigorously designed randomized experimental plan failed to demonstrate higher performance among students participating in SI-PASS. The first hypothesis cannot be confirmed. However, when reverting to a quasi-experimental approach—comparing SI-PASS participants with students who took the exam without prior engagement in the peer-learning scheme (i.e., all enrolled students excluding those in the control group), a significant performance disparity appears. This result aligns with Cheung et al.'s (2016) observation that quasi-experimental studies tend to report larger effect sizes than randomized experiments, raising broader concerns about the influence of research design on the evaluation of educational programs. When looking for a possible explanation, the role of intrinsic motivation in academic success could be invoked as one potential confounder. While all participants in the randomized trial were volunteers, the "all-comers" group likely included a higher proportion of students with uncertain commitment, potentially more prone to underperformance, thereby accentuating the contrast with the experimental group and leading to significant differences. These findings invite to caution when interpreting results from quasi-experimental studies within SI-PASS research. They also highlight the need for careful consideration, especially in the context of leader training and scheme promotion: tempting taglines making simplistic links between SI-PASS and academic success might be deceptive, unless nuance is added to the discussions.

### 4.2. Hypothesis 2 – Attendance threshold and group ownership

Results regarding the second hypothesis come as a nuance to the lack of effect recorded for the first one, by revealing a cut point (five participations) from where effects can be significantly traced in the exam grade (Fig. 2). Strictly speaking, the second hypothesis, which had located this inflection point at four participations, following Gattis' proposal (2002), is not confirmed. However, its rationale subsists. The transformative power of SI-PASS is neither automatic nor instantaneous. As already noticed by McCarthy et al. (1997), the binary dyad "participated" versus "did not participate" misses the point that a certain level of engagement is needed to produce effects. This observation is further substantiated by Fig. 1, which portrays a trend demonstrating that although lower levels of engagement fail to attain statistical

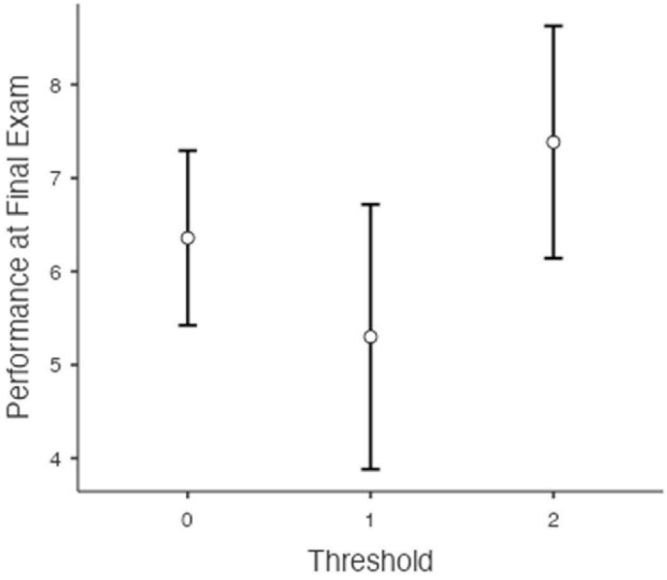



**Fig. 2.** Mean (95 IC95%) Differences between Control Group (0), Experimental Group <5 Meetings (1), and Experimental Group ≥5 (2) in attendance to SI-PASS Meetings.

significance, students who exhibit heightened commitment to attending SI-PASS meetings tend to perform better in the exam. The implications are immediate: any endeavors aimed at improving the support to students should not involve altering the SI-PASS scheme itself but directed towards encouraging students to elevate their level of participation in the weekly sessions.

The current body of literature does not conclusively fix the exact number of SI-PASS sessions students must attend to significantly benefit from the program. Dancer et al. (2007) tend to relate the optimal number of meetings needed to contribute to academic success to factors such as gender, local vs. international student status, etc. In any case, the optimal attendance level likely varies across different contexts. This context-dependent variation is illustrated by evidence from the only other randomized controlled trial so far, which contrasts with the findings of the present study. Dekker et al. (2023) noted a substantial impact on performance after > four 100-min sessions for about 35.7% of participating students, in contrast to the five 60-min sessions required in this study. Determining an ideal number of sessions for substantial improvement thus remains an open question for future research. This inquiry is complex, as it involves exploring nuanced aspects of the learning process. For instance, it remains conceivable that even minimal participation in SI-PASS sessions, if oriented towards strategic learning – a deviation from the true spirit of SI-PASS – could result in improved exam grades. In contrast, the effects that take longer to manifest might be associated with deeper understanding (Phan, 2009). An intriguing angle is explored by Malm (2021) in this respect. They investigate the surface/deep nature of SI-PASS through descriptive verbs used by participants, offering insights into the qualitative aspects of the program. Rigorous multidimensional documentation of various SI-PASS implementations could shed light on the impact of different intervening factors. With their meta-analysis of 28 peer-led learning research studies conducted from 1993 to 2017, Zha et al. (2019) offer an interesting example of such refined investigation as their work deliberately focuses on specific considerations: effects of the leader's training, task type, and session duration. In the same vein, a recent qualitative study by Dekker et al. (2024) focused on the understanding of perceived efficiency of SI-PASS on fourteen students. Three elements emerged from the interviews, namely: the use of effective study techniques, collaborative learning, and pedagogical climate. Besides these contextual parameters, refining the profiles of the students for whom this specific support can be

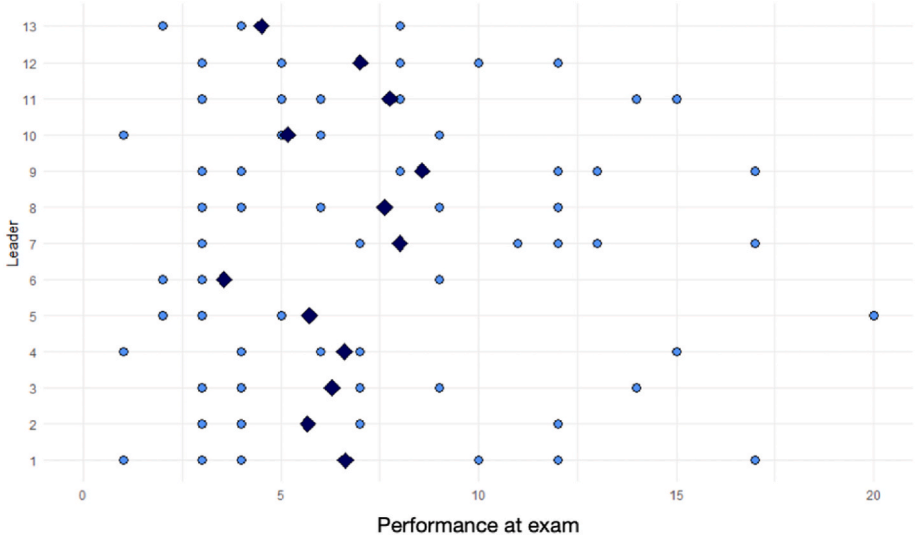



**Fig. 3.** Dotplot of students' performance at the final exam in relation to their leader.

beneficial also opens the door to personalizing support programs, a critical aspect of enhancing educational outcomes.

### 4.3. Hypothesis 3 – Socio-affective dimensions

Previous studies (Abegglen & Morris, 2015) have emphasized that one of the effects of the program relates to the social ties students build from weekly interactions with their peers. However, the dimensions tested in this research using both the SACQ and MSSCWS questionnaires revealed no significant difference between the control and experimental groups, consistent with the findings of Dekker et al. (2023). This outcome does not validate the third hypothesis. Should this be surprising or disappointing? Not necessarily. When summoning the classical conditions for a successful first year as historically established by Tinto (*i.e.* academic integration and social integration), SI-PASS is right from the start oriented towards the former, with socialization gains as welcomed by-products or fringe benefits which naturally emanate from the assembly of unfamiliar students during weekly sessions and the inherent SI-PASS approach fostering discussion and interaction. Expecting too much of socialization outcomes from a program primarily centered on mastering disciplinary academic content and mindset could overlook the fact that other peer-support initiatives, such as buddy or mentoring programs specifically tailored to enhance social integration (Larose & al., 2011), are more fittingly employed for such a purpose.

### 4.4. Hypothesis 4 – Profiles

Generally, SI-PASS schemes are part of a broader array of support mechanisms available for a course, potentially functioning in either complementarity or competition. What is the place of the SI-PASS program in such a learning ecology? Results show that there is no significant difference between both groups in terms of engagement with the online exercises. However, SI-PASS students who show the greatest levels of participation in meetings are also those who engage the most with online exercises. From these findings, one could think that those who display motivation for one course support tend to display it across other supports (Meehan & McCallig, 2015), pointing towards a congruence in the behavioral pattern displayed. In contrast, when a quasi-experimental approach is used to compare students who participated in SI-PASS to those who took the exam without showing initial interest in the peer-learning scheme (*i.e.* not the control group), a significant difference emerges between the groups. Further research is needed to explore this phenomenon and attest for the existence of intrinsic motivation in the consistent use of academic aids made available to students (Van Nuland et al., 2010). This internal disposition is proposed as a plausible explanation, though it should be noted that "motivation" is a broad term encompassing various constructs such as self-efficacy, attribution, goal orientation, and task value (Roozen et al., 2024; Trevino & DeFreitas, 2014; Vu et al., 2024), which this study was not designed to disentangle. Moreover, while motivation has been extensively studied in educational contexts (Galand & Bourgeois, 2006; Lieury & Fenouillet, 2013; Viau, 1994), its influence has been questioned. For instance, the meta-analysis conducted by Credé et al. (2010) suggests that class attendance is a stronger and more reliable predictor of academic success, with only a weak to moderate relationship with student motivation (p. 281, 287). While motivation is discussed here not directly in relation to grades, but as the underlying factor driving two forms of observable (non-) engagement in the course (SI-PASS attendance and online exercises), it would be worthwhile to incorporate class attendance (Leclercq & Glowacki, 2005) into future research on SI-PASS. This is especially relevant considering that at the University of Liège, it is part of the leaders' official duties to encourage students to attend classes before joining their SI-PASS group, ensuring they fully benefit from such collaborative gatherings. Unfortunately, data on class attendance was not collected in the present study. In their extensive literature review of student support actions, Delnoij et al. (2020), while acknowledging peer mentoring as a promising intervention to resolve the problem of non-completion, highlight, on the motivation side, the "contradictory results" (p. 7) regarding the link between motivational factors and retention in (online) higher education, which underscores the need for more precise research on this aspect in the context of SI-PASS schemes. This call for further investigation is echoed in the recent review by Van Vu et al. (2024), which focuses on the "motivation/achievement" reciprocity and concludes that essential factors in this dynamic may not yet have been fully captured, motivation being one potential confounder.

## 5. Limitations and further work

The primary limitation concerns the sample size. Although it isn't exceedingly small, it inherently stems from a finite student population associated with a relatively modest-sized faculty. This would not be a serious problem if other studies had found large effects. However, as indicated by the sensitivity analysis, the sample size of the present study only allows to detect medium effect sizes ($d = .42$) with an adequate power of .80 and is therefore barely acceptable for the average effect found in quasi-experiments ($g = .4$). The achieved power to detect

similar effect as the one found in Dekker et al.'s RCT study was significantly lower (the achieved power to detect an effect size of d = .26 was exactly .42 with the current sample size), such that the results of the present study cannot rule out real effect sizes in the range of those observed in Dekker et al.'s study. It is also worth noting that the Intra-class Correlation Coefficient (ICC) was zero, which fortunately maintained the study's statistical power. Had the ICC been greater than zero, the power would have been significantly reduced. Taken together, this experiment's constraints might also indicate that the real effect size could be closer to the $g = .26$ (or even lower, depending on attendance) found than the moderate $g = .4$ usually found in quasi-experimental designs. Consequently, the study doesn't fully overcome the warnings raised by McCarthy et al. (1997), Ashwin (2003), or Kochenour et al. (1997) regarding the "anecdotal" and "small or non-representative" sample sizes frequently seen in SI-PASS research. Recent developments in the field of educational measurement suggest that in the realm of education, many interventions might yield effects, but at a very modest magnitude (Hill et al., 2008; Kraft, 2020). Thus, test sensitivity becomes even more crucial than ever before. A second limitation relates to the composition of the sample. The higher representation of female participants can be attributed to the demographic distribution within the Faculty of Psychology. This gender skew might also tangentially echo findings from Peterfreund et al. (2008) and Hodges et al. (2001a) Hodges et al., 2001a, suggesting that more females tend to enroll in SI-PASS programs. Consequently, the generalizability of results to faculties with dissimilar gender distributions can be hampered. A third limitation touches upon the second factor of the SACQ scale, Academic Adjustment (Environment), which could only be assessed in the post-test phase. It might have been beneficial to establish a pre-test for this aspect, perhaps through inquiries about students' anticipation regarding the upcoming semester. Another limitation concerns previous education: students' GPA could not be obtained, and proxies were used instead. The last limitation is related to engagement with the course markers. The current study uses two of them: attendance to SI-PASS sessions and diligence for doing the online exercises. The attendance to the lectures would have been a natural and critical indicator of engagement as well. Unfortunately, this presence is never tracked at ULiège.

Future investigations should focus on more detailed observations regarding the content, structure, interactions within SI-PASS meetings. These extra insights would trigger the development of new measures and comparability options of schemes' quality against performances and profiles. Furthermore, exploring the dynamics of competition and complementarities among various learning support mechanisms in high-risk courses warrants thorough examination. In their article, Dekker et al. (2023) also introduce an interesting section "Costs-effectiveness of the intervention", presenting various assessments of the program. This goes beyond the typical focus on efficacy to include a consideration for efficiency. In lending support to this line of inquiry, which extends beyond SI-PASS to encompass all support programs and educational public policies, a rough calculation is offered for comparative purposes. In the present study, the cost per student, accounting solely for leaders' payment, is 99€ per term, with additional expenses for training, supervision, monitoring, reporting, and administrative services yet to be factored in. One last element that needs further investigation relates to intrinsic motivation. For both the first and the fourth hypotheses, when a supplementary analysis is performed with the so-called all-comers, statistical significance is reached. As the effect sizes for mean group differences are clearly higher in the supplementary analyses, this drastic change does not seem attributable to the sample sizes. A mediator variable, such as motivation, probably explain these significant differences between groups when tested in a quasi-experimental design. When two groups of motivated students are compared, there is no difference in terms of academic performance or engagement with supplementary material (or, more precisely, for a significant difference to emerge, it needs to be driven by "super-motivated" students, here the ones attending more than 5 sessions). However, once a comparison is made between the motivated students who attended the scheme and the rest of the cohort who did not display initial interest, these two differences become statistically significant. The place of motivation ought to be studied in more detail on further work to truly understand its role in academic achievement for peer-learning students (Fredricks et al., 2004).

## 6. Conclusion

Literature on SI-PASS now includes two full-fledged randomized trials: Dekker et al.'s (2023) and the present study. While Dekker et al. (2023) confirm the positive effect on average grades, consistent with many previous quasi-experimental investigations, the present research cannot account for the positive effect previously found. Both studies identify a critical threshold in attendance. It is reasonable to assume that thresholds and results in general vary based on targeted courses, attendee profiles (for instance the mean age is different between the two randomized controlled trials) and dynamics that unfold within SI-PASS sessions in terms of interactions and mastery learning. Further qualitative, observational, and controlled research is welcome to delve deeper into these aspects. Such explorations are essential for refining the understanding of SI-PASS, as well as other forms of academic support, and for putting into perspective their efficacy, efficiency, and personalization.

## CRediT authorship contribution statement

**Ninon Puttaert:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Dylan Dachet:** Writing – review & editing, Validation, Methodology, Formal analysis, Data curation, Conceptualization. **Ariane Baye:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization. **Etienne Quertemont:** Writing – review & editing, Methodology, Conceptualization. **Laurent Leduc:** Writing – review & editing. **Anne-Sophie Nyssen:** Writing – review & editing, Supervision. **Dominique Verpoorten:** Writing – review & editing, Validation, Supervision, Methodology, Conceptualization.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.learninstruc.2024.102042.

## References

Abegglen, S., & Morris, A. M. (2015). Peer-led learning: Challenges and opportunities. *Investigations in University Teaching and Learning, 10*, 77–81. https://repository.londonmet.ac.uk/id/eprint/1228.

Armentor, M. M. (2019). A quasi-experimental study of the effect of supplemental instruction on course completion and persistence at a two-year college. *Doctoral dissertation, Northcentral University)*, 123 pages. ISBN: 978-1-3922-0686-8 https://eric.ed.gov/?id=ED609606.

Ashwin, P. (2003). Peer support: Relations between the context, process and outcomes for the students who are supported. *Instructional Science, 31*, 159–173. https://doi.org/10.1023/A:1023227532029

Aspelmeier, J. E., Love, M. M., McGill, L. A., Elliott, A. N., & Pierce, T. W. (2012). Self-esteem, locus of control, college adjustment, and GPA among first- and continuing-generation students: A moderator model of generational status. *Research in Higher Education, 53*(7), 755–781. https://doi.org/10.1007/s11162-011-9252-1

Barron, B. (2004). Learning ecologies for technological fluency: Gender and experience differences. *Journal of Educational Computing Research, 31*(1), 1–36. https://doi.org/10.2190/1n20-vv12-4rb5-33va

Bengesai, A. V., Amusa, L. B., & Dhunpath, R. (2023). A meta-analysis on the effect of formal peer learning approaches on course performance in higher education. *Cogent Education, 10*(1), Article 2203990. https://doi.org/10.1080/2331186x.2023.2203990

Bolker, B. (2024). lme4 (version 1.1-35.3). Retrieved from https://rdocumentation.org/packages/lme4/versions/1.1-35.3.

Bronstein, S. B. (2007). *Supplemental instruction: Supporting persistence in barrier courses.* Amherst: University of Massachusetts. https://doi.org/10.7275/18739695

Brooman, S., & Darwent, s. (2014). Measuring the beginning. A quantitative study of the transition to higher education. *Studies in Higher Education, 39*(9), 1523–1541. https://doi.org/10.1080/03075079.2013.801428

Brunet, S., Dujardin, C., Louis, V., O'Dorchai, S., et al.Dehon, C. (2021). Les réformes de l'enseignement supérieur en Fédération Wallonie-Bruxelles à l'heure de l'évaluation [Higher education reforms in the Wallonia-Brussels Federation at the time of evaluation]. *Dynamiques Régionales - IWEPS, 11*, 1–90.

Carayon, S., & Gilles, P. Y. (2005). Développement du questionnaire d'adaptation des étudiants à l'université (QAEU). *L'Orientation Scolaire et Professionnelle, 34*(2), 165–189. https://doi.org/10.4000/osp.463 [Development of the student adaptation to college questionnaire].

Cheng, D., & Walters, M. (2009). Peer-assisted learning in mathematics: An observational study of student success. *Journal of Peer Learning, 2*(1), 23–39. https://ro.uow.edu.au/ajpl/vol2/iss1/3.

Cheung, A. C., & Slavin, R. E. (2016). How methodological features affect effect sizes in education. *Educational Researcher, 45*(5), 283–292. https://doi.org/10.3102/0013189X16656615

Cole, J. S., & Spence, S. W. (2012). Using continuous assessment to promote student engagement in a large class. *European Journal of Engineering Education, 37*(5), 508–525. https://doi.org/10.1080/03043797.2012.719002

Credé, M., Roch, S. G., & Kieszczynka, U. M. (2010). Class attendance in college: A meta-analytic review of the relationship of class attendance with grades and student characteristics. *Review of Educational Research, 80*(2), 272–295. https://doi.org/10.3102/0034654310362998

Dahmus, S., Bernardin, H. J., & Bernardin, K. (1992). Student adaptation to college questionnaire. *Measurement and Evaluation in Counseling and Development, 25*(3), 139–142. https://psycnet.apa.org/record/1993-12063-001.

Dancer, D., Morrison, K., & Smith, M. (2007). Measuring the impact of a peer assisted learning program on students' academic performance in econometrics. In *The quantitative analysis of teaching and learning in higher education: Forum proceedings* (pp. 19–42). Melbourne: Victoria, Australia: University of Melbourne. https://doi.org/10.1080/03075079.2014.916671.

Dawson, P., van der Meer, J., Skalicky, J., & Cowley, K. (2014). On the effectiveness of supplemental instruction: A systematic review of supplemental instruction and peer-assisted study sessions literature between 2001 and 2010. *Review of Educational Research, 84*(4), 609–639. https://doi.org/10.3102/0034654314540007

De Clercq, M., Roland, N., Dangoisse, F., & Frenay, M. (2023). *La transition vers l'enseignement supérieur: Comprendre pour mieux agir sur l'adaptation des étudiants en première année [The transition to higher education: Understanding to better act on the adaptation of students in the first year]*. Lausanne: Peter Lang. https://doi.org/10.3726/b20782

Dekker, I., Koerhuis-Pasanisi, M.-J., & Koek, M. (2024). When and how do peers stimulate engaging in desirable difficulties: Student perspectives on the effectiveness of supplemental instruction. *Active Learning in Higher Education, 0*(0). https://doi.org/10.1177/14697874241249130

Dekker, I., Luberti, M., & Stam, J. (2023). Effects of supplemental instruction on grades, mental well-being, and belonging: A field experiment. *Learning and Instruction, 87*, 1–9. https://doi.org/10.31219/osf.io/nw8j9

Delnoij, L. E., Dirkx, K. J., Janssen, J. P., & Martens, R. L. (2020). Predicting and resolving non-completion in higher (online) education–A literature review. *Educational Research Review, 29*, Article 100313. https://doi.org/10.1016/j.edurev.2020.100313

Dobbie, M., & Joyce, S. (2008). Peer-assisted learning in accounting: A qualitative assessment. *Asian Social Science, 4*(3), 18–25. https://doi.org/10.5539/ass.v4n3p18

Fayowski, V., & MacMillan, P. D. (2008). An evaluation of the supplemental instruction programme in a first-year calculus course. *International Journal of Mathematical Education in Science & Technology, 39*(7), 843–855. https://doi.org/10.1080/00207390802054433

The first year of college. In Feldman, R. S. (Ed.), *The first year of College: Research, theory, and practice on improving the student experience and increasing retention*, (pp. i–ii). (2018) (pp. i–ii). Cambridge University Press.

Finch, W. H., Bolin, J. E., & Kelley, K. (2019). *Multilevel modeling using R* (2nd ed.). CRC Press (Taylor & Francis Group) https://api.taylorfrancis.com/content/books/mono/download?identifierName=doi&identifierValue=10.1201/9781351062268&type=googlepdf.

Francotte, E., Baudoin, N., Coertjens, L., Galand, B., Crépin, F., Quittre, V., Baye, A., Monseur, C., & Lafontaine, D. (2023). *Enquête multidimensionnelle et systémique relative au bien-être à l'école et au climat* scolaire *(OASE 7) – Rapport final* [Multidimensional and systemic investigation of well-being at school and school climate (OASE 7) – Final report]. *ULiège & UCLouvain*. http://hdl.handle.net/2078.1/274271.

Fredricks, J. A., Blumenfeld, P. C., et al.Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59–109. https://doi.org/10.3102/00346543074001059

Galand, B., & Bourgeois, E. (Eds.). (2006). *Se) Motiver à apprendre [Motivate (yourself) to learn]*. Presses universitaires de France.

Galand, B., Neuville, S., & Frenay, M. (2005). L'échec à l'université en Communauté française de Belgique: Comprendre pour mieux prévenir [Failure at university in the French Community of Belgium: Understanding to better prevent]. *Les Cahiers de Recherche en Education et Formation, 39*, 5–17. http://hdl.handle.net/2078.1/93664.

Gattis, K. W. (2002). Responding to self-selection bias in assessments of academic support programs: A motivational control study of supplemental instruction. *Learning Assistance Review, 7*(2), 26–36. https://eric.ed.gov/?id=EJ659738.

Guarcello, M. A., Levine, R. A., Beemer, J., Frazee, J., Laumakis, M., & Schellenberg, S. (2017). Balancing student success: Assessing supplemental instruction through coarsened exact matching. *Technology, Knowledge and Learning, 22*(3), 335–352. https://doi.org/10.1007/s10758-017-9317-0

Hafer, G. R. (2001). Ideas in practice: Supplemental instruction in freshman composition. *Journal of Developmental Education, 24*(3), 30–37. https://www.proquest.com/scholarly-journals/ideas-practice-supplemental-instruction-freshman/docview/228489488/se-2.

Hager, A. E. (2018). *How does peer-assisted learning effect math anxiety, help-seeking behavior, and performance in undergraduate mathematics?* Plymouth State University. https://hdl.handle.net/20.500.12774/319.

Hayes, C., & Fulton, J. (2019). A participatory action research study on the impact of peer assisted student support (PASS) and supplemental instruction (SI) by international PhD students. *Journal of Learning Development in Higher Education, 14*, 1–17. https://doi.org/10.47408/jldhe.v0i14.477

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*(3), 172–177. https://doi.org/10.1111/j.1750-8606.2008.00061.x

Hockings, C. (2009). Reaching the students that student-centered learning cannot reach. *British Educational Research Journal, 35*(1), 83–98. https://doi.org/10.1080/01411920802041640

Hodges, R., Dochen, C., & Joy, D. (2001a). Increasing students' success: When supplemental instruction becomes mandatory. *Journal of College Reading and Learning, 31*(2), 143–156. https://doi.org/10.1080/10790195.2001.10850111

Hodges, R., & White, W. G. (2001b). Encouraging high-risk student participation in tutoring and supplemental instruction. *Journal of Developmental Education, 24*(3), 2–43. http://ies.ed.gov/ncee/wwc/study/81665.

Howard, E., Meehan, M., & Parnell, A. (2019). An exploration of the relationship between continuous assessment and resource use in a service mathematics module. In *Eleventh congress of the European society for research in mathematics education (No. 23)*. Freudenthal Group. https://hal.science/hal-02422644.

International Labour Office. (2012). *International standard classification of occupations 2008 (ISCO-08): Structure, group definitions and correspondence tables*. International Labour Office. https://isco-ilo.netlify.app/en/isco-08/.

Jacquemart, J., Clercq, De, Mikaël, & Galand, B. (2023). *Mieux comprendre les pratiques enseignantes en classe dans l'enseignement supérieur : Proposition d'un cadre de référence* [Better understanding teaching practices in the classroom in higher education: Proposal of a reference framework]. *Formation et profession: Revue scientifique internationale en éducation, 31*(3), 1–19. https://dial.uclouvain.be/downloader/downloader.php?pid=boreal%3A287360&datastream=PDF_01&disclaimer=6dc9f4791175427b1aa54d915cd424c7a0ced60731581e7fe177484d4735b1f2.

James Madison University. (2023). Implementation fidelity. https://www.jmu.edu/assessment/sass/ac-step-four.shtml.

Jiang, B., XU, X., Garcia, A., & Lewis, J. E. (2010). Comparing two tests of formal reasoning in a college chemistry context. *Journal of Chemical Education, 87*(12), 1430–1437. https://doi.org/10.1021/ed100222v.

Johnston, B. (2010). *The first year at university: Teaching students in transition*. UK: McGraw-Hill Education.

Kochenour, E., Jolley, D., Kaup, J., Patrick, D., Roach, K., & Wenzler, L. (1997). Supplemental instruction: An effective component of student affairs programming. *Journal of College Student Development, 38*(6), 577–586. https://eric.ed.gov/?id=EJ557886.

Kovač, V. B. (2015). Transition: A conceptual analysis and integrative model. In D. L. Caeron, & et R. Thygesen (Eds.), *Transitions in the field of special education: Theoretical perspectives and implications for practice* (pp. 19–33). Waxmann.

Kraft, M. A. (2020). Interpreting effect sizes of education interventions. *Educational Researcher, 49*(4), 241–253. https://doi.org/10.3102/0013189X20912798

Larose, S., Cyrenne, D., Garceau, O., Harvey, M., Guay, F., Godin, F., et al. (2011). Academic mentoring and dropout prevention for students in math, science and technology. *Mentoring & Tutoring: Partnership in Learning, 19*(4), 419–439. https://doi.org/10.1080/13611267.2011.622078

Leclercq, D., & Glowacki, J. (2005). TOPACMER: Indices de participation d'étudiants à un cours dans l'Enseignement Supérieur [TOPACMER: Student participation indices in a course in Higher Education]. *Paper presented at the XXIIe Colloque International de l'Association Internationale Universitaire (AIPU)*. Université de Genève, Genève, Switzerland. https://orbi.uliege.be/bitstream/2268/2571/1/TOPACMER/20texte/203/20octobre/202005/20pour/20AIPU/20Genève.pdf.

Lieury, A., & Fenouillet, F. (2013). *Motivation et réussite scolaire (3ème édition) [Motivation and academic success* (3rd ed.). Dunod.

Lowe, T., & Wright, S. (2024). Mapping the student experience: A framework for assessing student support, success, community and voice. *Student Success, 15*(1), 92–98. https://doi.org/10.5204/ssj.2866

Malm, J. (2021). A study on learning activities in supplemental instruction. In *Supplemental instruction: Volume 2: Student learning processes* (pp. 25–46). Waxmann Verlag.

Malm, J., Carey, W., Dahlberg, L., Mörner, L. L., & Ody, M. (2024). Status report for European SI/PASS/PAL-programmes: Post-pandemic/Second edition. https://www.si-pass.lu.se/sites/si-pass.lu.se/files/2024-01/Status/20report/20European_post/20pandemic_v2.pdf.

Martin, D. C., & Arendale, D. R. (1992). Supplemental instruction: Improving first-year student success in high-risk courses. *National Resource Center for The First Year Experience*. https://hdl.handle.net/11299/200471.

Martin, D., & Arendale, D. (1993). Supplemental instruction: Improving first-year student success in high-risk courses. *Columbia: National resource center for the first year experience and students in transition* (2nd ed.). University of South Carolina https://hdl.handle.net/11299/200471.

Mayhew, M. J., Rockenbach, A. N., Bowman, N. A., Seifert, T. A., Wolniak, G. C., Pascarella, E. T., & Terenzini, P. T. (2016). *How college affects students: Findings from the 21st century* (Vol. 3). John Wiley et Sons.

McCarthy, A., Smuts, B., & Cosser, M. (1997). Assessing the effectiveness of supplemental instruction: A critique and a case study. *Studies in Higher Education, 22*(2), 221–231. https://doi.org/10.1080/03075079712331381054

Meehan, M., & McCallig, J. (2015). Patterns of student engagement with a variety of learning resources in a large first year module. *Paper presented at the inaugural European conference on the scholarship of teaching and learning (EUROSoTL)*. Dublin, Ireland https://www.ehu.eus/documents/7269444/7573413/EuroSoTLBookofAbstracts.pdf.

Miller, V., Oldfield, E., & Bulmer, M. (2004). Peer assisted study sessions (PASS) in first year chemistry and statistics courses: Insights and evaluations. In *Proceedings of the scholarly inquiry in flexible science teaching and learning symposium* (pp. 30–51). The University of Queensland. https://openjournals.library.sydney.edu.au/IISME/article/view/6493.

Moore, R., & LeDee, O. (2006). Supplemental instruction and the performance of developmental education students in an introductory biology course. *Journal of College Reading and Learning, 36*(2), 9–20. https://doi.org/10.1080/10790195.2006.10850184

Noël, B., & Parmentier, P. (1998). De l'élève à l'étudiant [From pupil to student]. In M. Frenay, B. Noël, P. Parmentier, & M. Romainville (Eds.), *L'étudiant apprenant*. Bruxelles: De Boeck.

Normak, P., Pata, K., & Kaipainen, M. (2012). An ecological approach to learning dynamics. *Educational Technology & Society, 15*(3), 262–274. https://www.jstor.org/stable/jeductechsoci.15.3.262.

Paloyo, A. R., Rogan, S., & Siminski, P. (2016). The effect of supplemental instruction on academic performance: An encouragement design experiment. *Economics of Education Review, 55*, 57–69. https://doi.org/10.1016/j.econedurev.2016.08.005

Parkinson, M. (2009). The effect of peer assisted learning support (PALS) on performance in mathematics and chemistry. *Innovations in Education & Teaching International, 46*(4), 381–392. https://doi.org/10.1080/14703290903301784

Peterfreund, A. R., Rath, K. A., Xenos, S. P., & Bayliss, F. (2008). The impact of supplemental instruction on students in STEM courses: Results from San Francisco State University. *Journal of College Student Retention: Research, Theory and Practice, 9*(4), Article 487503. https://doi.org/10.2190/cs.9.4.e

Phan, H. P. (2009). Exploring students' reflective thinking practice, deep processing strategies, effort, and achievement goal orientations. *Educational Psychology, 29*(3), 297–313. https://doi.org/10.1080/01443410902877988

Roozen, I., Goeman, K., & De Grez, L. (2024). Motivation and performance of first-year students. *Higher Education Studies, 14*(1), 109–125. https://doi.org/10.5539/hes.v14n1p109

Siemens, G. (2003). Learning ecology, communities, and networks : Extending the classroom. *Elearnspace. Everything eLearning*.

Slavin, R. E. (2007). *Statement of robert E. Slavin, director center for data-driven reform in education. Committee on appropriations subcommittee on labor, health and human services, education, and related activities*. Hearings on Implementation of No Child Left Behind. Retrieved March 16, 2007, from http://www.ednews.org/articles/8996/1/Statement-of-Robert-E-Slavin-Director-Center-for-Data-Driven-Reform-in-Education/Page1.html. (Accessed 14 March 2007).

Stansbury, S. (2001). Accelerated learning groups enhance supplemental instruction for at-risk students. *Journal of Developmental Education, 24*(3), 20–40. https://www.proquest.com/scholarly-journals/accelerated-learning-groups-enhance-supplemental/docview/228473917/se-2.

Strømmen-Bakhtiar, A., Helde, R., & Susen, E. (2021). *Supplemental instruction: Volume 2: Student learning processes*. Münster, Germany: Waxmann Verlag. https://doi.org/10.31244/9783830993254

Szal, R. J., & Kennelly, K. R. (2017). The effects of supplemental instruction on student grades in a blended learning context. In *Developments in business simulation and experiential learning: Proceedings of the annual ABSEL conference (vol. 44)*. https://absel-ojs-ttu.tdl.org/absel/article/view/3095.

Tampakis, A., & Vitoratos, E. (2009). Estimation of students' workload: Correlation of teaching and learning methods with examination results. A case study. In *Internationalization and the Role of university networks Proceedings of the 2009 EMUNI Conference on higher Education and research (1-20)*. Portorož, Slovenia: EMUNI University.

Thomas, L. (2012). *Building student engagement and belonging in higher Education at a time of change: A summary of findings and recommendations from the what works? Student retention & Success programme*. York: Higher Education Academy.

Tinto, V. (1993). *Leaving college: Rethinking the causes and cures of student attrition*. University of Chicago press. https://doi.org/10.7208/chicago/9780226922461.001.0001

Tobin, K. G., & Capie, W. (1981). The development and validation of a group test of logical thinking. *Educational and Psychological Measurement, 41*(2), 413–423. https://doi.org/10.1177/001316448104100220

Trautwein, C., & Bosse, E. (2017). The first year in higher education: Critical requirements from the student perspective. *Higher Education, 73*(3), 371–387. https://doi.org/10.1007/s10734-016-0098-5

Trevino, N. N., & DeFreitas, S. C. (2014). The relationship between intrinsic motivation and academic achievement for first generation Latino college students. *Social Psychology of Education, 17*, 293–306.

Upcraft, M. L., Gardner, J. N., & Barefoot, B. O. (2005). *Challenging and supporting the first-year student: A handbook for improving the first year of college* (Vol. 254). San Francisco, CA: Jossey-Bass.

Van der Zanden, P. J., Denessen, E., Cillessen, A. H., & Meijer, P. C. (2018). Domains and predictors of first-year student success: A systematic review. *Educational Research Review, 23*, 57–77. https://doi.org/10.1016/j.edurev.2018.01.001

Van Nuland, H. J. C., Dusseldorp, E., Martens, R. L., et al.Boekaerts, M. (2010). Exploring the motivation jungle: Predicting performance on a novel task by investigating constructs from different motivation perspectives in tandem. *International Journal of Psychology, 45*(4), 250–259. https://doi.org/10.1080/00207591003774493

Van Rooij, E. C. M., Jansen, E. P. W. A., & van de Grift, W. J. C. M. (2018). First-year university students' academic success: The importance of academic adjustment. *European Journal of Psychology of Education, 33*(4), 749–767. https://doi.org/10.1007/s10212-017-0347-8

Vanthournout, G., Gijbels, D., Coertjens, L., Donche, V., & Van Petegem, P. (2012). Students' persistence and academic success in a first-year professional bachelor program: The influence of students' learning strategies and academic motivation. *Educational Research International, 2012*(1), Article 152747. https://doi.org/10.1155/2012/152747

Vázquez, S. M., & de Anglat, H. D. (2009). Academic achievement and formal thought in engineering students. *Electronic Journal of Research in Educational Psychology, 7*(2), 653–672. https://doi.org/10.25115/ejrep

Verpoorten, D., & Jérôme, F. (2022). Observation par les pairs entre leaders SI-PASS: Qualité du feedback et bénéfices perçus [Peer observation between SI-PASS leaders: Quality of feedback and perceived benefits]. *Paper presented at the 2nd didactifen conference, Liège, belgique*. https://didactifen2022.sciencesconf.org/data/pages/Livret_des_re_sume_s_DIDACTIfen2022.pdf#page=4.

Verpoorten, D., Parlascino, E., & Colaux, C. (2021). SI-PASS in a Belgian university: A pilot showcase. In A. Strømmen-Bakhtiar, R. Helde, & E. Suzen (Eds.), *Supplemental instruction. Volume 2: Student learning processes* (pp. 123–139). https://doi.org/10.31244/9783830993254. Münster, Germany: Waxman.

Viau, R. (1994). *La motivation en contexte scolaire [Motivation in a school context]*. Bruxelles: De Boeck.

Vu, T. V., Scharmer, A. L., van Triest, E., van Atteveldt, N., & Meeter, M. (2024). The reciprocity between various motivation constructs and academic achievement: A systematic review and multilevel meta-analysis of longitudinal studies. *Educational Psychology, 44*(2), 136–170. https://doi.org/10.1080/01443410.2024.2307960

Wasylkiw, L. (2016). Students' perspectives on pathways to university readiness and adjustment. *Journal of Education and Training Studies, 3*(4), 28–39. https://doi.org/10.11114/jets.v4i3.1197

Zha, S., Estes, M. D., & Xu, L. (2019). A meta-analysis on the effect of duration, task, and training in peer-led learning. *Journal of Peer Learning, 12*(2), 5–28. https://files.eric.ed.gov/fulltext/EJ1219651.pdf.