

# Discussion

## Overview

My analysis of 206 self-reported surveys reveals complex heterogeneity of the influence of an innovative ESD intervention on changes in students' sustainability competencies. On the one hand, contrary to my predictions, I found that one year post-intervention, sustainability attitudes and behaviours did not increase for either the involved or the control group (Figure 2), suggesting a rejection of my alternative hypothesis of a positive relationship. On the other hand, in line with my predictions, I was able to show that the involved group ( $n=7$ ) reported overall higher sustainability attitudes and behaviours than the control group one year post-intervention (Figure 2), indicating positive effects of an innovative ESD intervention on changes in sustainability competencies. Yet, the differences between the groups were already present pre-intervention (Figure 2). This highlights the importance of long-term empirical data collection when analysing the impact of ESD interventions. Careful consideration should be given to the generality of the findings due to the very small sample size ( $n=7$ ) of involved students at measurement point 3. I found a strong positive relationship between the sustainability attitudes and behaviours scales operationalised with the Theory of Planned Behaviour and those operationalised with the efficacy beliefs construct (Figure 3), suggesting mutual validation of the two scales to capture the same latent constructs of sustainability competencies and an indication of validation through prediction of an impact-relevant behaviour. I found no differences between personal and collective efficacy beliefs within and between the involved and control groups (Figure 4) at the third measurement point, highlighting the need to include collective efficacy as a goal dimension in ESD intervention designs. In line with my predictions, I found that the involved students reported higher aim-focused efficacy beliefs (Figure 5), indicating the stabilisation of the formation of sustainability intentions, while also highlighting the importance of considering external barriers. The lack of sufficient data prevented me from conducting more statistically robust tests and leads to severe limitations on the generalisability of the results, highlighting the challenges of collecting comprehensive data in school contexts. Measuring the outcome of ESD interventions requires many considerations and trade-offs - by using a quantitative, longitudinal, treatment-control group, outcome-focused approach, I uncovered heterogeneous responses of students' sustainability competencies, aiming to contribute to ongoing improvements in the measurability of sustainability competencies and pointing to the potential of innovative ESD interventions.

The measurement of sustainability competencies has been criticised for various reasons, which have important implications for the interpretation of the results. A major discussion within the field of ESD measurement and research is context sensitivity (Waltner et al., 2019). Behavioural outcomes can vary significantly between different contexts and cultures. An intervention that promotes sustainable behaviour in one setting may not have the same effect in another due to different social norms, economic conditions and cultural values. This variability calls into question the universality of behavioural outcomes (Sossé et al., 2021). In the use of indicators, there is an inherent trade-off between the desire to achieve a global scope to allow comparisons to be made, and the need for context specificity. In thinking globally, I have not researched beyond the Western perspective, given resource limitations and the scope of this study. All of the underlying theories come from a Western context, as does most of the empirical research of the studies cited. Given the dominance of Western research, one should be cautious about the global claim of the indicators and their relevance in different (including non-Western) contexts. A very first step, as I see it, in trying to overcome the Western predominance in research is to state one's own positionality and to contextualise the research, as I have tried to do here. In addition to the limited empirical and theoretical basis of this study in the Western literature, this study is also based on only one school in Freiburg, which calls into question the generalisability of the findings. To allow for long-term data collection, I used a scale developed by Pauli (2023) for her master's thesis, operationalised through the Theory of Planned Behaviour, to capture sustainability attitudes and behaviours. Although Pauli (2023) tried to adapt some of the scales to students, I would argue that they have limited suitability for young people and students, as many of the questions used are neither age appropriate nor necessarily in the hands of students, especially the behavioural intentions questions (e.g. I support an increase in fuel taxes to reduce fossil fuel consumption OR I mainly drive or am driven by a car or motor scooter). In terms of using indicators for young people and students, there are better alternatives, for example in the ProBiKlima project (ProBiKlima, 2024). Furthermore, the context

and therefore the project-specific objectives need to be stated. The objectives of the KRS project were only partly aimed at improving students' competencies in terms of sustainability attitudes and behaviours. Rather, the focus was also on improving democratic education (KRS website, 2024). As the scales I used did not capture these dimensions at all, the analysis is limited. While I consider my critical reflections on the usefulness of the indicators used in this study as a contribution to SC measurement research, the empirical data from this study is very limited and not very generalisable.

Another key challenge of the measurement of sustainability competencies is the adequate operationalising of ESD outcomes, including related goals. The method and tool for operationalising sustainability competencies aim at equipping students with core competencies for shaping a sustainable future. However, this raises questions about evaluation, definition and the need for openness in the ESD concept to remain adaptable to future sustainability challenges. As Wals et al. conclude, "the main point is that there is no single model of education and learning for environmental sustainability, nor should there be" (Wals & Benavot, 2017). Maintaining an adaptive and flexible approach to ESD should not hinder empirical research efforts to verify the effectiveness of ESD programmes. Evaluating and improving ESD interventions is necessary to ensure that they enable learners to shape a sustainable future. This compatibility between ESD and empirical research is crucial, especially when distinguishing between ESD1 (instrumental) and ESD2 (emancipatory) approaches. While ESD1 focuses on specific behavioural outcomes, ESD2 emphasises the process of learning and critical engagement rather than predetermined outcomes. Similarly, the debate continues as to whether behavioural change should be the outcome of studies, given the complexity of behaviour, its multiple influences, including factors outside the individual, especially for young people (Getzin & Singer-Brodowski, 2017). To try to address these critiques, I applied the Triple-A framework of efficacy beliefs, which offers a promising approach to navigating the instrumental vs. emancipatory debate by focusing on whether agents believe they can achieve their self-selected goals, rather than prescribing specific behaviours. This framework inherently incorporates external factors, with higher efficacy occurring when individuals believe they can effect change. Empowerment research further challenges scholars to consider changes in actual power and agency, not just self-reported outcomes, and highlights the importance of real-world impacts on collective social and environmental goals (Cattaneo et al., 2014). The inclusion of personal efficacy beliefs and collective efficacy as outcome indicators can address some criticisms of non-instrumental ESD by capturing the complexity of human motivations and broader educational goals. While behavioural change remains an important indicator, it should be complemented by measures that promote critical thinking, empowerment and intrinsic motivation. This comprehensive approach ensures that interventions not only change behaviour, but also cultivate the underlying values and beliefs necessary for sustained and meaningful engagement with sustainability issues. Although my research on sustainability competencies is limited by data availability, it can contribute to the debate on appropriate indicators of sustainability competencies by including students' efficacy beliefs. Assessing these efficacy beliefs over time and in relation to project development would provide valuable insights. Despite the cautious interpretation of my findings due to the limited data, this research contributes to the understanding of sustainability competencies, their continuous development and ways to measure them, thus helping to identify effective strategies for ESD interventions.

## **Sustainability competencies as sustainability attitudes and sustainability behaviours (TPB-based) (Research Question 1)**

Contrary to my prediction, I found that one year after the ESD intervention, sustainability attitudes and behaviours did not increase for either the involved or the control group (Figure 2). On the other hand, in line with my predictions, I demonstrated that the involved group ( $n=7$ ) reported statistically higher sustainability attitudes and behaviours (cumulative SA/SB and SB) than the control group one year post-intervention (Figure 2), possibly indicating positive effects of an innovative ESD intervention. For all scales, I observed a peak at the second measurement point (immediately post-intervention) for the involved group, with scores significantly higher than those of the control group (Figure 2). At the same time, the cumulative SA/SB and SB were reported to be even significantly higher at the first measurement point (pre-intervention), which raises the question of attributing sustainability competencies to the ESD intervention (Figure 2). This doubt is reinforced by the fact that SA and SB peaked for the involved group, but then fell back to levels similar to those pre-intervention. There could be several reasons for the observed results. Firstly, the path

from an educational intervention to changes in SA and SB remains complex and difficult to predict. Research suggests that sustainability attitudes decrease as students get older (Krettenauer, 2017; Waltner et al., 2021), which could potentially counteract the effects of the intervention. Looking at the effectiveness of other ESD interventions, a meta-study from 2021 found that ESD interventions led to increased environmental sensitivity, reconsideration of preconceptions, improved ability to solve complex problems related to the environment, increased likelihood of identifying environmental issues as personal concerns, and relative maintenance of newly acquired positive practices (Ssossé et al., 2021). Most of these studies focused on small groups and used different operationalisations, making comparisons difficult. There is a consensus in the literature on the need for long-term, longitudinal impact studies that take into account other types of concrete ESD outcomes that can be realised in a sometimes more distant horizon than what current studies can cover (especially activism) (Ssossé et al., 2021). There is no quantitative empirical research on the effectiveness of innovative ESD interventions such as the KRS project (Riess et al., 2022). My findings highlight the importance of conducting sustainability measurement based on long-term empirical and quasi-experimental designs when assessing the effectiveness of (innovative) ESD interventions.

The observed findings could also be due to factors that influenced students' SA and SB more than the ESD intervention itself. A recent long-term study (Waltner et al., 2021), based on the ESD goals in local education plans (and not on a specific ESD intervention), found significant predictors of SA and SB to be average school grade, sustainability-related attitudes at the beginning of the school year, participation in Fridays for Future activities, knowledge of the concept of sustainability (only predictive of SA) and grade level (only predictive of SA), which I did not take into account. Other factors found to influence SA and SB were the social desirability (Armitage & Conner, 2001) of the response options and the influence of the media, which may have a greater influence than ESD interventions, especially in this age group (Waltner et al., 2021). Extracurricular learning environments (e.g. friends, family, social media) have a strong influence during adolescence, so the effectiveness of school-based ESD may be limited. Furthermore, the role of the teacher or, in this case, the project conductor, could influence sustainability competencies. Research has shown that there seems to be a negative relationship between teachers' attitudes towards sustainability and environmental awareness, leading to less SB among students. It can also only be assumed that if the teacher makes too pointed statements about his or her own environmental and sustainability awareness, this could possibly lead to reactance in the students' own attitudes (Waltner et al., 2021). All these factors could explain why there was no increase in SA and SB over time for either the involved or the control group. To explain the decline of the involved group from immediately post-intervention to a level similar to pre-intervention at one year post-intervention, this could also be due to frustration based on feedback from the environment. Research has shown that when basic psychological needs for belonging, competence, and autonomy are not met and need frustration occurs, this could lead to a decrease in pro-environmental behaviour (Wullenkord et al., 2021). As my study did not investigate the extent to which the school implemented the roadmap, the lack of implementation could have led to negative feelings among the students involved. My findings highlight the urgent need for whole-institution approaches (FMER, 2017) to ESD and the design of ESD interventions.

## **Sustainability attitudes/ sustainability behaviours (TPB-based) and efficacy beliefs (Research Question 2)**

I found a strong positive relationship between the scales of sustainability attitudes and behaviours operationalised with the Theory of Planned Behaviour and the scales operationalised with the construct of efficacy beliefs (Figure 3). This suggests that they capture the same latent construct of sustainability competencies and can be used as a first step in a mutual validation of the scales. In a possible second step of validation, the (voluntary and self-determined) participation of the surveyed students in the expert group (which makes them the involved group in this study) could possibly be used to draw conclusions about environmental activism resulting from a high level of environmental attitudes. Participation in this group can be interpreted as a behavioural manifestation aimed at promoting a more sustainable school, which consequently serves to achieve ESD goals. Participation in the involved group was recorded at the first measurement point of the project with a simple question about this activity. In this case, self-reported behaviour was considered a valid proxy for actual behaviour. The data showed that students in the involved group had higher SA/SB than

students who were not involved (Figure 2). Normally, validation by predicting impact-relevant behaviour is done by checking SA and then, at a second measurement point, seeing if the higher SA translates into actual behaviour (Waltner et al., 2022). Given that in this study the group was formed but had not yet begun its work at the first measurement point (pre-intervention), it could be argued that their high SA led to actual behaviour by the time they joined the group. The question remains to what extent socially desirable responses played a role, as being part of the group and perceiving oneself as such, even if the work had not yet started, could have influenced their results. With caution, participation in the expert group could be seen as validation by predicting impact-relevant behaviour. Nevertheless, by validating the scales, the competence differences assessed by this instrument may indicate meaningful differences between students that may have a real impact on their future behaviour. To the best of my knowledge, I found no validation procedures in the literature for either of these two scales. My findings highlight the importance of ongoing and critical validation of the indicators and scales used to determine whether the measurement is meeting its objectives and is able to predict actual behaviour.

## **Sustainability competencies as efficacy beliefs (Research Question 3)**

### **Collective efficacy beliefs as outcomes**

Contrary to my predictions, I found no differences between personal and collective efficacy beliefs within and between the involved and control groups (Figure 4). All the mean scores were very similar. There could be several reasons for the observed results. On the one hand, both personal and collective efficacy beliefs could have been mutually reinforcing, resulting in no detectable difference between them. Studies have found that individuals can derive personal benefits (e.g. efficacy beliefs) from social groups because groups can make them feel personally capable and in control. Collective efficacy made individuals feel in control of their outcomes: People's intention to act was enhanced by a sense of efficacy transferred from the group to the self (Jugert et al., 2016). Similarly, using a qualitative research approach, research (Cocking & Drury, 2004) found that collective efficacy led to a sense of personal empowerment. Thus, collective and personal efficacy are strong and closely intertwined predictors of pro-environmental behaviour and are mutually reinforcing. On the other hand, from a theoretical perspective, the participatory and innovative design of the ESD intervention may have led to an increase in the collective efficacy beliefs of the group involved. Bandura (1997) suggests that efficacy is enhanced when individuals acquire the specific skills necessary for pro-environmental behaviour and are verbally encouraged about their ability to perform such actions. These sources of efficacy are effective when individuals evaluate the pro-environmental effectiveness of their own groups. Consequently, when a group successfully completes a task and receives positive reinforcement for its pro-environmental efforts, its members are likely to have a stronger belief in their ability to achieve the desired outcomes. Commitment to collective efforts is essential to fostering a belief that their actions can impact their environment. This increased belief, in turn, can lead to a greater willingness to engage in pro-environmental behaviour (Chen, 2015). Given the findings that the involved group did not express more collective efficacy beliefs than the control group, questions arise about the relevance of the chosen group, the desirability of the aims asked in the survey, but also the success of their efforts and the extent of positive external reinforcement. The advantage of differentiation, i.e. the possibility to self-categorise the agent, is undermined by the fact that in a quantitative survey without a pilot study the relevant self-categorisations could not be found out. Empirically, there is very little research that distinguishes between personal and collective efficacy. The very few studies that have made this distinction between personal and collective efficacy beliefs have found that collective efficacy is significantly stronger when the task difficulty is moderate - rather than easy or difficult (Reese & Junge, 2017). Behaviours that are easy to perform tend to have a lower environmental impact. People may perceive that actions that are too easy (e.g. avoiding plastic bags) are unlikely to make a significant difference to environmental problems, even if they are widely practiced. In other words, when actions are too simple, the potential success may not lead to a sense of collective efficacy. This could be an indication that the aim of reducing the school's CO<sub>2</sub> emissions, led by a student initiative, was considered too difficult a task by the students and/or would have required more support and positive reinforcement from their environment. Furthermore, participative efficacy may play a role as a moderator between personal and collective efficacy and is concerned with how well a group can

achieve its goal together, influenced by group size and group cohesion (van Zomeren et al., 2013). It also raises the question of the extent to which collective efficacy beliefs were the aim of this particular project and were therefore supported and reinforced by the school and the project partners. My findings highlight the relevance of collective efficacy as an outcome indicator of ESD interventions, both in the design of ESD interventions, but also as a relevant indicator for measuring sustainability competencies.

### **Aim-focussed efficacy beliefs as outcomes**

In line with my predictions, I found that the involved students reported statistically higher aim-focused efficacy beliefs than the control group (Figure 5). This suggests that the involved group had demonstrated intention formation. At the same time, the lower action-focused efficacy beliefs scores of the involved group may also have some interesting implications. Action-focused efficacy refers to the belief in one's ability to perform certain actions, whereas aim-focused efficacy refers to the belief in one's ability to achieve desired outcomes. These two facets are interrelated; effective action often leads to aim attainment, and belief in aim attainment can motivate individuals to take necessary action (Hamann et al., 2024). Given the suggestion that action-focused efficacy is more related to the perception of actual constraints such as time, money, and social resources, this finding highlights the potential external barriers that the involved group encountered, which moderated their beliefs in their ability to carry out their actions. From a methodological point of view, it could also be that the actions chosen for the survey were not relevant to either group. This aspect could not be tested and identified in a pilot study. As the Triple-A framework allows the combination of very concrete actions with very abstract (collective) aims, the testing of these concrete actions becomes more relevant. There is little empirical research on aim-focused versus action-focused efficacy beliefs. Research (Hornsey et al., 2006) found that aim content significantly influenced how efficacy predicted action intentions for members and non-members of a protest group. This suggests that the relevance and desirability of the aim are crucial to understanding how efficacy beliefs are translated into intentions and actions (Kruglanski & Higgins, 2007). To assess the effect of the ESD intervention on aim- or action-focused efficacy beliefs, it would have been necessary to measure efficacy beliefs over time. Assessing efficacy beliefs over time and in relation to the development and implementation of the project would be an interesting direction for future research. Furthermore, differentiating the links between agents, actions and aims allows better predictions about which characteristics of efficacy make it more or less predictive of relevant social and environmental outcome variables. This differentiation could also inform the design of ESD, depending on the desired outcome of more action or aim orientation. My findings highlight the importance of differentiating between action-focused and aim-focused efficacy beliefs as an outcome indicator of ESD interventions to inform the design of ESD interventions and also to understand potential external factors faced by individuals.

### **Study limitations**

Analyses of students' sustainability competencies can be limited by insufficient sample size, limited temporal and geographical scope of the data, and limited empirical validation of the theoretical frameworks used. The lack of data, which prevented me from answering one of my original research questions regarding the participatory nature of the project, highlighted the need for more comprehensive data collection. The generalisability of the empirical findings is severely limited by the small sample size of students involved at the third measurement point ( $n=7$ ). The small sample size also affects the statistical power and reliability of the findings. Larger sample sizes generally provide more accurate and generalisable results, reducing the margin of error and increasing confidence in the findings. Furthermore, I argue that the scales used to operationalise sustainability attitudes and behaviours had a limited fit with the context. However, they provided an opportunity to collect longitudinal empirical data. Unfortunately, no prior data were available for the efficacy belief scales, which would be an interesting future direction. Non-Western contexts, theories and empirical evidence were under-represented, limiting applicability to other contexts. Finally, the Triple-A framework, being newly developed, still lacks strong empirical evidence for the distinctions it makes, with potential moderators of the relationship between different efficacy links still missing (Hamann et al., 2024). Recognising heterogeneous effects, including complex behaviours, and analysing what is already practised, highlights previously overlooked research questions and helps researchers to make more strategic choices in

the study of efficacy beliefs. The analysis of students' sustainability competencies and their attribution to ESD interventions can benefit from a more comprehensive study design and scope of data, as well as broader theoretical and empirical contexts.

## **Future directions**

Measuring students' sustainability competencies and attributing them to ESD interventions requires going beyond previous studies that are based on limited study designs, focus on traditional learning methods or are limited in their relevance to measure what matters. Studying ESD interventions in a quasi-experimental design and collecting long-term empirical data will provide a more nuanced view of the effectiveness of ESD interventions. Focusing on promising innovative learning and teaching methods could provide the evidence needed to challenge existing learning approaches that have so far proved insufficient to address global challenges. A better understanding of which sustainability competencies can have a real impact, without instrumentalising students, will lead to more human agency. For example, measuring efficacy beliefs over time could be an interesting future direction. Consequently, the development of indicators and outcomes is more complex and relates to the researcher/educator's definition of development education, as discussed earlier. This focus on product outcomes misses the specificity of ESD, where learning outcomes may take the form of questioning and activism rather than immediate or short-term goals. Continuous development of relevant indicators is needed. Qualitative methods of ESD research could begin to address these challenges by integrating different research fields such as environmental psychology, environmental sociology, science education and empirical education sciences. The analysis of sustainability competencies with a comprehensive study design will allow the development of better educational policies.