

Albert-Ludwig-University Freiburg  
Faculty of Environment and Natural Resources

**Measuring What Matters: Towards A Comprehensive Approach to Measuring Students‘ Sustainability Competencies by Integrating Efficacy Beliefs**

Master-thesis submitted in partial fulfilment of the requirements for the degree of   
Master of Science Environmental Governance

by

**Daniela Naveena Gargya**

5107568



First supervisor: Prof. Dr. Astrid Carrapatoso, Pädagogische Hochschule Freiburg  
Second supervisor: Prof. Dr. Uwe Schmidt, Albert-Ludwigs-Universität Freiburg

Freiburg i.Br. / Germany

Date of submission:  
02. August 2024

**Summary**

Contents

[Introduction 8](#_Toc173164284)

[Theoretical and conceptual context 8](#_Toc173164285)

[Goals of ESD 9](#_Toc173164286)

[Operationalisation ESD outcomes 10](#_Toc173164287)

[Measurement approaches 12](#_Toc173164288)

[Potential role of efficacy beliefs in sustainability competencies 13](#_Toc173164289)

[Objectives and research questions 16](#_Toc173164290)

[Research hypotheses and predictions 16](#_Toc173164291)

[Methods 18](#_Toc173164292)

[Project description KlimaRatSchule 18](#_Toc173164293)

[Study design and data collection 18](#_Toc173164294)

[Survey tools 20](#_Toc173164295)

[Data processing 23](#_Toc173164296)

[Data analysis 23](#_Toc173164297)

[Quality Criteria testing 24](#_Toc173164298)

[Results 27](#_Toc173164299)

[Sustainability competencies as sustainability attitudes and sustainability behaviours (TPB-based) (Research Question 1) 27](#_Toc173164300)

[Sustainability attitudes/ sustainability behaviours (TPB-based) and efficacy beliefs (Research Question 2) 28](#_Toc173164301)

[Sustainability competencies as efficacy beliefs (Research Question 3) 28](#_Toc173164302)

[Personal and collective efficacy beliefs 28](#_Toc173164303)

[Aim- and action-focussed efficacy beliefs 29](#_Toc173164304)

[Discussion 31](#_Toc173164305)

[Overview 31](#_Toc173164306)

[Sustainability competencies as sustainability attitudes and sustainability behaviours (TPB-based) (Research Question 1) 33](#_Toc173164307)

[Sustainability attitudes/ sustainability behaviours (TPB-based) and efficacy beliefs (Research Question 2) 34](#_Toc173164308)

[Sustainability competencies as efficacy beliefs (Research Question 3) 35](#_Toc173164309)

[Collective efficacy beliefs as outcomes 35](#_Toc173164310)

[Aim-focussed efficacy beliefs as outcomes 36](#_Toc173164311)

[Study limitations 37](#_Toc173164312)

[Future directions 37](#_Toc173164313)

[Conclusions 39](#_Toc173164314)

[Appendices 40](#_Toc173164315)

[Declaration of consent by parents/ info etc 40](#_Toc173164316)

[Declaration of submission of the Master thesis 41](#_Toc173164317)

**Acknowledgements**

I would like to thank Dr. Eva-Maria Waltner and Prof. Dr. Astrid Carrapatoso for their constructive and helpful supervision. From getting to know the project, finding the research scope, consultation about data collection and data analysis, I always found open ears and reassuring plan Bs. I would also like to thank Fabia Spörckmann and Prof. Dr. Uwe Schmidt for the second supervision, their interest in the project and the uncomplicated and easy communication.

I also want to thank Rolf and Anette from Solare Zukunft e.V. for giving information on the project and providing the contact information to the schools. At the Angell School, I would like to thank the teaching staff for their support in collecting the data.

Also, I want to thank my parents and friends for constant support and proofreading the draft.

**List of abbreviations**

ESD Education for sustainable development

RQ Research question

SA Sustainability attitudes

SB Sustainability behaviours

SC Sustainability competencies

TPB Theory of planned behaviour

**Overview of tables and figures**

# Introduction

Globally, we as humankind are facing many ecological and social crises at the same time. Despite extensive and continuous efforts to transition to a more sustainable society, environmental and social challenges persist or have even worsened in various aspects (IPCC, 2023). In transition towards a more sustainable and just society, Education for Sustainable Development (ESD) is seen as a key aspect (BMBF, 2020). ESD is a holistic approach to education, that focuses on the development of sustainability competencies that enable the learners to contribute to sustainable development through their competencies of knowledge, skills, motivation, attitudes, and values (Rieckmann & Barth, 2022). Following the Brundtland Report and the Agenda 21 Conference in Rio, numerous programmes on Education for Sustainable Development have been initiated (Hoffmann, 2020). However, the relevance, impact, and effectiveness of ESD are often questioned (Ssossé et al., 2021). While ESD is widely accepted as a theoretical concept and goal dimension, it remains without a universally agreed upon definition. In addition, the complexity of the educational process makes it difficult to define outcomes and measure success accurately. Arguably, the effective development of ESD requires evidence-based recommendations, to make a real contribution much needed changes in society. When looking at the empirical data on the impact and outcomes of ESD interventions in educational settings, there is very little. Disentangling the goals, methods and outcomes of ESD and establishing effective education without instrumentalising education is an urgent issue (REF). Currently, we have a limited quantitative understanding of how ESD interventions influence the development of sustainability competencies over time at the level of personal characteristics of learners, such as knowledge, skills, attitudes and values. Despite the call for a more comprehensive approach to measuring and operationalising ESD research, namely a shift from an input to an outcome orientation (Waltner et al., 2022), research has mostly focused on traditional learning methods (BUGEN, REF MONI) or on uncomprehensive study designs (MONI). In this analysis, I focus on the implementation of ESD measures at school level. Although the students currently being educated in our schools are not the decision-makers of today, they will ultimately be the decision-makers of the future. Recent theoretical advances in sustainability competencies, such as the Triple A framework of efficacy beliefs, together with the possibility of collecting long-term empirical data (Pauli) in a quasi-experimental design, allow us to quantitatively test the effects of innovative ESD interventions on students' sustainability competencies. The measurement of sustainability competencies linked to effective ESD education can provide the information and educational design guidance needed for better educational policy making in the light of our rapidly changing Anthropocene.

## Theoretical and conceptual context

Education for Sustainable Development (ESD) aims to empower individuals to contribute effectively to sustainable development by fostering critical engagement with societal complexities and contradictions (REF?). This educational approach, endorsed by UNESCO, is divided into two strands: ESD 1, an instrumental approach that promotes specific sustainable behaviours as defined by experts, and ESD 2, an emancipatory approach that equips learners with the tools to critically evaluate and identify sustainable practices themselves (Vare & Scott, 2007; Wals, 2011). While ESD 1 focuses on measurable changes in behaviour, ESD 2 emphasises the development of sustainability competencies necessary to navigate and influence complex, ambiguous issues. Both approaches can complement each other, especially in educational settings, where the behavioural focus of ESD 1 is justified by the shared educational and parental responsibilities in schools (KMK, 2018). For older learners, however, the focus should be on ESD 2 in order to cultivate critical thinking and autonomous action (Rieckmann, 2021b). This shift is in line with the broader educational trend towards competency-based learning, which prioritises skills and abilities over specific content knowledge, thereby increasing the likelihood that students will internalise and apply sustainable practices (de Haan, 2006; Pant, 2016). ESD 2 approaches are even more complex to measure because of the issues surrounding the definition of relevant sustainability competencies to be measured. This complexity is evident in the ongoing debate on ESD and other transformative educational approaches (REFS SINGER). ESD, as an intentional educational intervention, has the potential to challenge existing social, economic and political systems that perpetuate injustice and inequality, aligning it with critical pedagogy and interpretivist or constructivist epistemological paradigms (Liddy, 2011). While these paradigms would lend themselves well to mixed methods approaches, which allow for the existence of multiple realities and experiences that can be viewed differently (Moustakas 1994) and complex, nuanced outcomes (Dumas and Anderson 2014), much research is still based on predominantly positivist epistemologies (O'flaherty), with quantitative measures such as pre/post surveys commonly used to evaluate educational interventions. Due to resource constraints, I also chose to adopt a purely quantitative approach, reflecting a positivist epistemology. There are many complex considerations that need to be taken into account when aiming to improve ESD - however, effective ESD interventions can benefit from being designed around measurable sustainability competencies and their continuous development.

### Goals of ESD

Sustainability competencies need targets against which they can be measured. While there may be project-specific objectives for each ESD intervention, the question remains as to the overarching understanding of the dimensions of the goals. At the international level, goals are set, for example, by Agenda 21 (REF). They provide a normative framework, but there is no operationalisation of ESD outcomes (REF). At the local level, there are regional education plans, such as the one in Baden-Württemberg: "Education for Sustainable Development enables learners to make informed decisions and act responsibly for the protection of the environment, a functioning economy and a just world society for current and future generations" (Ministry of Education Baden Württemberg, 2016), but they also remain without operationalisation. Educational goals, which are personal characteristics to be promoted in learners, should include a normative test and an empirical test (UHL). Critical analyses of prominent ESD goal recommendations have proposed sustainability competencies as "cognitive abilities and skills as well as the associated motivational, volitional and social readiness needed to be able to solve sustainability-relevant problems and shape sustainable development in private, social and institutional contexts. (WEINERT)". Although some ESD learning goals exist, there is still a lack of operationalisation of ESD outcomes (RIESS/MISCHO) to be translated into measurement models and tools. The use of empirically validated measurement tools and approaches from related disciplines makes it possible to integrate operationalised facets of competencies in the context of ESD. As such, ESD learning goals need to be structured and interrelated. One prominent framework is the tripartite framework model for sustainability competencies (RIESS), which includes cognitive, affective motivational (= attitudinal), and behavioural aspects (e.g., Rieckmann, 2018; Waltner et al., 2019) related to sustainability. The cognitive dimension of sustainability competencies ranges from knowledge of basic concepts of sustainable development and the SDGs to an understanding of physical, ecological, social, cultural, economic and political systems related to sustainability. The attitudinal dimension includes values, needs and motivational competencies, such as acceptance of intergenerational equity and attitudes towards consumption and mobility. The behavioural dimension includes promoting sustainable practices and helping learners to translate their understanding into concrete actions. The framework has the advantages of including the behavioural dimensions (Lambrecht et al), being adaptable to different contexts (REF), and countering criticism of the dominance of the cognitive dimensions (REF). If the aim is to measure sustainability competencies, it is useful to benchmark against the specific goal dimensions of sustainability knowledge, sustainability behaviour (SB) and sustainability attitudes (SA).

### Operationalisation ESD outcomes

There have been numerous attempts to operationalise the measurement of sustainability competencies through the use of survey scales. For empirical measurement, these competencies need to be precisely defined. Such specificity allows for the appropriate operationalisation of the competency in question (see Klieme & Hartig, 2007). Despite existing efforts, there is still a need to develop adequate measurement instruments for the different dimensions of sustainability competenies (SC). Integrating ESD research with well-established measurement approaches facilitates the inclusion of already operationalised facets (e.g. environmental attitudes) in the broader construct of SC. Examples of measurement approaches exist in related disciplines for each of the three target dimensions. For the cognitive dimension, tools are available to assess specific aspects of sustainability knowledge (e.g., environmental knowledge) (e.g., Frick et al., 2004; Maloney & Ward, 1973; McBeth et al., 2011; Roczen et al., 2014). For the attitudinal domain, tools such as the Greenpeace Sustainability Barometer (REF) are used, and for the behavioural dimension, the General Ecological Behaviour Scale (REF) is a notable example. Recent research has attempted to combine these scales and adapt them to the school context (BUGEN). At the same time, it remains uncertain to what extent indicators can capture the ultimate normative goal of ESD: the actual development of a more sustainable society through sustainable behaviour. Theories of action in psychology suggest that various forms of knowledge and motivational factors (including subjective and social norms and attributions of responsibility) interact to form behavioural intentions, which then lead to sustainability-relevant behaviour (Ajzen, 1991). In addition, external conditions (e.g. behavioural opportunities, situational contexts, social norms and the lifestyle of the social environment) also influence sustainability-relevant behaviour. Promoting knowledge and motivational orientations can therefore encourage desired behaviours. However, research has consistently shown a significant gap between knowledge, motivational orientations and actual sustainability-promoting behaviour (Carmi & Arnon, 2014). Specifically, in the cognitive goal dimension, sustainability knowledge is positively related to sustainability attitudes (Arcury, 1990), but the prediction of actual behaviour is quite low (Frick et al., 2004). Knowledge serves as an indirect predictor of behaviour, providing a foundation but lacking the necessary motivational factors. This often-cited gap between knowledge and behaviour is evident in several studies. Kagawa notes that "there are multiple factors which influence the process of behaviour change, and the dissonance between students' perceptions of sustainability and their individual actions needs to be further explored" (Kagawa, 2007). This is consistent with research on the attitude-behaviour gap (Olsson & Gericke, Kollmuss & Agyeman, Terlau & Hirsch) and cognitive dissonance (Kumar & Metin). For the affective-motivational and behavioural dimensions, numerous studies have shown a very close relationship. Affective goal commitment, or a positively valued sequence of actions, is central to motivation. An action will not be performed if the perceived costs outweigh the affective goal commitment. The attitude-cost relationship is modelled in the Campbell paradigm (Kaiser et al., 2010), which suggests a strong link between an individual's attitudes and behaviour. Consequently, within the Campbell paradigm, behavioural self-reports are used as indicators of a person's attitudes. When operationalising the sustainability competencies that are most relevant to real-world behaviour, it is most effective to use scales that capture sustainability attitudes and sustainability behaviours.

The operationalisation of behavioural sustainability competencies can benefit from exploring the underlying theories of behaviour. Empirically, some of the most commonly used models of sustainability-relevant behaviour are the Theory of Planned Behaviour (TPB) and the Norm Activation Model (Bamberg & Möser, 2007; Sopha, 2011; Klöckner, 2013). The TPB is one of the most prominent from-within approaches (Ertz, 2016). The from-within approach focuses on individual decision making. As this approach is more accessible to education, it will be pursued further in this paper. In the TPB, intention and perceived behavioural control (PBC) are seen as direct determinants of behaviour (Ajzen, 1991, p. 184). Intention, in turn, is influenced by PBC, but also by two other factors: firstly, attitudes towards the behaviour in question, and secondly, subjective norms that influence behavioural intention (Ajzen). Behind these determinants are beliefs or convictions that relate to behaviour, norms and control (Ajzen). The TPB, as a rational choice theory, focuses on the self-interest of the person weighing the consequences of their actions (Kaiser et al., 2006, p. 2151). Kaiser et al. assessed the explanatory power of the TPB for pro-environmental behaviour (2006, p. 2160). According to their study, the three determinants - attitude, subjective norms and perceived behavioural control - explain 76% of the variance in behavioural intention, while these in turn explain 95% of the variance in an individual's pro-environmental behaviour (ibid.). However, in a meta-analysis by Armitage and Conner, this average is significantly lower at 27% variance explanation (Armitage & Conner, 2001, p. 471). On the downside, the directions of effect between attitudes, subjective norms and perceived behavioural control remain incompletely identified in the TPB (Kaiser et al., 2006, p. 2165). Furthermore, it has been criticised for focusing on internal factors such as knowledge and attitudes and for failing to consider structural barriers, power relations and inequalities that pose systemic barriers to individuals (Bamberg, 2021), as well as moral and normative considerations (Kaiser, 2006). For operationalisation in the context of this study, the TPB was used due to existing data. Attitudes and subjective norms can be assigned to the affective dimensions, while PBC and intentions represent motivational aspects (Weber, 2008). These four variables at the first and second causal levels are summarised here as sustainability attitude dimensions, alongside the sustainability behaviour dimension, which is derived directly from self-reported behaviour. Therefore, sustainability competencies and the sustainability attitude and sustainability behaviour dimensions can be operationalised with the components of the Theory of Planned Behaviour.

### Measurement approaches

The operationalisation of sustainability competencies and their attribution to an ESD intervention can be assessed using different methods. Quantitative research projects can make a significant contribution to the normative debate by providing empirical insights and a valuable information base for ESD stakeholders, including teachers (cf. DIPF, 2007; Oekes, 1989). It is important not to rely solely on plausible and normative considerations and assumptions. Some effects of educational measures, such as teaching or the whole institutional approach, may only be empirically verifiable in the long term, or may not be clearly attributable to a specific measure due to the complexity of interacting variables that influence, for example, sustainability awareness. These considerations highlight that shifting the focus from input to output orientation of ESD policies may require longer-term assessments and different methodological approaches to effectively assess impacts. At present, empirical data on the long- and short-term impacts of ESD initiatives in educational settings are limited, indicating the need for further research and comprehensive evaluation strategies. The measurement of sustainability competencies benefits from looking at long-term empirical data.

When measuring sustainability competencies, the choice of measurement method also depends on the teaching and learning approach used. There is a strong need for alternative and innovative approaches to ESD (REF). Fostering personal attributes with high affective-motivational components, such as attitudes towards sustainable development and intergenerational equity, requires methods that effectively address complex human characteristics. Despite the inherent stability and resistance to change of attitudes, the integration of principles from general motivation and social psychology may prove beneficial. Techniques such as inducing cognitive conflict, role-playing, simulation games and experiential learning activities, including sustainability projects and internships, have shown potential to increase motivation and change attitudes (REF). These methods facilitate a deeper engagement with different perspectives and arguments, which is essential for promoting sustainable behaviour. In addition, supporting students' intrinsic motivation by providing experiences of competence, autonomy and social connectedness - such as choosing between different types of food in a school canteen and participating in group learning - can significantly enhance the effectiveness of these educational strategies (Mischo, 2004; Deci & Ryan, 2008). Promoting sustainable behaviour requires a multifaceted approach that addresses both internal and external factors that influence individual actions. Psychological theories of action suggest that a mix of knowledge, motivational factors, subjective and social norms, and attributions of responsibility can lead to the formation of behavioural intentions and ultimately drive actual behaviour. However, despite the availability of knowledge and motivational orientations, there is often a significant gap between these elements and the manifestation of behaviour that promotes sustainability. To bridge this gap, methods from volitional psychology are recommended. These methods include making individuals aware of and challenging their guiding assumptions and beliefs, forming concrete action plans, encouraging self-commitment, visualising the results of actions and using reminders. In educational settings, teachers play a crucial role in promoting behaviour change by raising awareness of students' subjective theories and fostering efficacy, thereby enabling students to see the impact of their actions and motivating them to commit to sustainable practices. However, the availability of empirical data is limited to certain types of ESD learning/teaching methods. The call for alternative and innovative methods of delivering ESD is not empirically supported. To date, innovative ESD intervention studies mostly include results in the form of self-reports, self-assessments and expert surveys (REF MONI), which are non-experimental research methods. In contrast, quasi-experimental studies can be used to test hypotheses and provide evidence of the effectiveness of methods. To date, these robust scientific methods have mainly been applied to ESD intervention methods with a high degree of teacher guidance. In addition, the validity of quasi-experiments can be enhanced with experimental controls (e.g. pre-post or control group test design) (38, Handley). Due to the lack of empirical data, it is not yet possible to make conclusive statements about the 'best' teaching/learning methods, although there are recommendations from empirical education and teaching research. Given the mismatch between evidence-based learning recommendations and the availability of sound research methods, there is a need to measure innovative ESD methods using quasi-experimental designs.

The measurement of sustainability competencies requires validity criteria to ensure that the measurement tools accurately capture the achievement of the objectives - in this case, the competencies that enable learners (among others) to contribute to real-world change. Ideally, one would obtain observational data to see if actual behaviour has changed. In reality, self-reports are more commonly used as they are easier to obtain, especially in large numbers, and also allow for a broader assessment of different behaviours (REF EWM). The challenge is that there is a gap between self-reported and objective behaviour. To overcome this, measurement indicators need to be validated with real behavioural outcomes. One reason for this gap could be the tendency of respondents to give socially desirable answers (REF). There are several ways to carry out scale validation. One option would be to perform an ad hoc scale validation, comparing the items used in this study with well-established measures that have been tested for their predictive power of real-life behaviour. Another option would be to validate against a criterion outside of the measurement process (Whitley, 1977) that is considered a valid proxy for actual behaviour and is assessed using a simple dichotomous approach (e.g. Kaiser et al., 2003; Kormos & Gifford, 2014). The newly developed Triple A framework of efficacy beliefs complements the TPB-based scales in terms of its predictive power for behaviour and behavioural intentions (REF). One study found that the TPB could benefit from incorporating aspects of the Triple A framework to increase its explanatory power (Bamberg et al., 2015, see also Huijts et al., 2013). The Triple A framework, with its high degree of differentiation, expands the possibility of assessing different behavioural outcomes. Therefore, an interesting option is to compare the TPB-based scale with one based on efficacy beliefs. Administering both a TPB-based scale and one based on efficacy beliefs provides an opportunity to validate each scale against the other, thus ensuring a higher quality of measurement.

### Potential role of efficacy beliefs in sustainability competencies

Measuring sustainability competencies can benefit from a high degree of differentiation based on the new theoretical development of the Triple-A framework of efficacy beliefs. The Triple-A framework provides a structured approach to enhance the somewhat fragmented field of efficacy research, particularly in the context of collective social and environmental goals. This framework is essential for addressing social and environmental crises more effectively by understanding how individuals and collectives can experience greater efficacy. The framework builds on foundational theories of efficacy, including Bandura, Ellen Skinner, and Turner et al, and is structured around three core components - agents, actions, and aims. The self-categorised agent can be an individual (personal self) or a collective group (an ingroup), where the group is part of an individual's self-concept influenced by social and emotional group membership. This distinction allows for the exploration of both personal and collective efficacy. Individuals can move flexibly from categorising themselves as individuals to members of groups (Cocking, Fritsche, Tajfel), reflecting different underlying social identities. The agentic aspect signals the possibility of being able to effect change. The second aspect of the framework, intentional action, emphasises that efficacy involves deliberate, measurable actions aimed at achieving specific outcomes. These actions are defined not only by their intentionality, but also by their self-determined nature, level of abstraction and content. This component emphasises the proactive nature of efficacy in influencing and changing environmental and social conditions. Finally, the Triple-A framework focuses on aims, which are cognitive representations of desired outcomes, whether personal or collective. Aims are essential because they focus attention, motivate, and promote persistence and skill development (Locke & Latham, 2002). The framework emphasises the bidirectional influence between efficacy and aims, with perceptions of efficacy influencing the choice of aims and, conversely, the nature of those aims influencing the development and maintenance of efficacy beliefs. This also highlights the importance of aim desirability. Overall, the Triple-A framework not only clarifies the structure of efficacy beliefs, but also enriches theoretical discourse by distinguishing the three efficacy links (agent-action, agent-action-aim, agent-aim) and explicitly applying them to both personal and collective contexts. This approach promises to advance the practical application of efficacy theory in addressing broader social and environmental challenges by emphasising the interconnectedness of beliefs, actions, and aims in achieving sustainable change. Given the advantages of the efficacy framework, it represents an interesting complement to the measurement of sustainability competencies.

Sustainability competencies do not only depend on the individual. As early as 1997, Albert Bandura described a collective sense of powerlessness in an increasingly interdependent world and argued that in the face of multiple crises, there is a growing need for collective as well as individual action. Many people do not act collectively against climate change or social inequality because they feel that they or their group cannot make a difference. Understanding how people come to feel that they can make a difference (a sense of efficacy) is therefore crucial to motivating people to act together for a better world. The nature of the environmental and social crisis also requires collective action. Correlational studies have found that personal efficacy predicts individual behaviour and collective efficacy predicts activist behaviour (Hamann & Reese, 2020; Morton et al., 2011). Although the pathways and linkages between personal and collective efficacy beliefs are not yet fully understood, research can contribute to practical interventions based on this distinction. Some research suggests a pathway from collective to personal efficacy to private behaviour (Jugert et al., 2016), although this is not yet causally replicable. Experimental studies suggest that ingroup efficacy interventions outperform personal efficacy interventions in influencing behaviour (Jugert et al., 2016). Research could focus on which ingroup efficacy agents are more or less important in specific social and environmental crises, drawing on social identity theory (Tajfel, 1978), which characterises different ingroup efficacy agents. Collective efficacy has been found to be a better predictor of pro-environmental behaviour than personal efficacy (M-F Chen, 2015). Furthermore, specific collective and personal efficacy perceptions have been used to measure sustainable intentions, even when controlling for attitudes and social norms, suggesting that collective efficacy beliefs are particularly relevant for achieving environmental goals (Reese & Junge). Measuring sustainability competencies may benefit from including efficacy beliefs as an outcome target at both the personal and collective levels.

Sustainability competencies are not only dependent on internal factors. The distinction between action- and aim-focused efficacy beliefs is interesting and relevant because it provides a clearer understanding of how beliefs about personal and collective capabilities translate into behaviour and goal attainment. Changes in sustainability competencies, particularly efficacy beliefs, are dependent on feedback from the environment. The Triple-A framework allows for a distinction between action and aim-focused efficacy linkages, with agent-action efficacy more closely linked to actual behavioural costs, socio-economic circumstances and impactful behaviours, whereas agent-aim efficacy is more closely linked to attitudes, goals, visions and intentional behaviours (Bain et al., 2013; Bamberg & Rees, 2015). Perceived behavioural control in the theory of planned behaviour (Ajzen, 1991) predicts intention and moderates intention-outcome relationships, suggesting that action-focused efficacy may function similarly by capturing constraints such as time, money, or social resources that may prevent the implementation of intentions. In contrast, aim-focused efficacy is more involved in intention formation and less influenced by these constraints. The field of collective social and environmental aims is particularly prone to an aim-focused understanding of efficacy due to the complex nature of collective crises (Zomeren, 2019). For collective aims, the challenge is not in performing actions (e.g., going to a protest), but in achieving social change through these actions. Given that many actions to achieve an environmental or social aim are possible, but also potentially unsuccessful, acknowledging the looser action-aim contingencies highlights the importance of distinguishing between action- and aim-focused efficacy. Individuals have limited control over collective outcomes, with many obstacles stemming from external factors and the actions of others (Hornsey et al., 2021; Jugert et al., 2016). Feedback is more difficult to obtain, as aims are often distal (for example, the impact of an awareness campaign on public opinion). Hornsey et al. (2006) found that efficacy predicted action intentions differently for members and non-members of protest groups depending on the content of the aim. Studies that have included different aims in their efficacy measures have shown that ignoring aim content can lead to inconsistent results and obscure underlying principles (Koletsou & Mancy, 2011). The desirability of a aim is crucial; if a aim is not desirable, efficacy measures may elicit defensiveness and reflect only the value of the aim (Castiglione, 2021). The relationship between action-focused and aim-focused efficacy is complex and interdependent. Action-focused efficacy refers to beliefs in one's ability to perform specific actions, whereas aim-focused efficacy refers to beliefs 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. For example, someone with high action-focused efficacy for an environmental behaviour (such as using public transport) may also develop high aim-focused efficacy for achieving broader environmental goals (such as reducing their carbon footprint). The Triple-A framework integrates these concepts, suggesting that efficacy beliefs should consider both actions and aims to provide a comprehensive understanding of how individuals and groups approach and solve complex social and environmental problems. Measuring changes in sustainability competencies at both the action and aim levels allows for a better understanding of the influence of external and structural factors on behaviour.

## Objectives and research questions

My aim is to quantify how ESD interventions influence students’ sustainability competencies. My aim is to quantify how ESD interventions influence students' sustainability competencies. Specifically, I will focus on sustainability competencies, operationalised as sustainability attitudes and sustainability behaviours, measured by scales based on the Theory of Planned Behaviour. In addition, I will assess efficacy beliefs through scales operationalised based on the newly developed Triple A framework. To achieve this, I will investigate the influence of an innovative ESD intervention in one school, in a quasi-experimental longitudinal study with a treatment group control group comparison, by analysing 206 self-reported surveys from three different measurement points of sustainability competencies indicators, addressing the following research questions (RQ):

1. How do students' sustainability attitudes and sustainability behaviours change over time (including one year after the end of an ESD intervention)?
2. How are sustainability attitudes and sustainability behaviours influenced by efficacy beliefs?
3. How do (personal and collective) efficacy beliefs (aim- and action-focussed) respond to the level of student involvement?

## Research hypotheses and predictions

I hypothesize that sustainability attitudes and sustainability behaviours will increase over time. The magnitude of sustainability competency gains increases for the involved group. When looking at the relationship between sustainability attitudes/ sustainability behaviours and efficacy beliefs, I hypothesize a strong correlation. I hypothesize that efficacy beliefs will vary within and between the involved and control group, in terms of personal and collective, and aim- and action-focussed efficacy beliefs. These hypotheses will be tested against the null hypotheses of no increase in SA and SB over time, no relationship between SA/SB and efficacy beliefs, and no differences in efficacy beliefs within and between groups.

I predict greater higher SA and SB for the involved students, as innovative learning methods seem promising to achieve SA and SB. I predict the relationship between SA/SB and efficacy beliefs to be strong, as they capture the same latent construct. I predict higher collective and aim-focussed efficacy beliefs for the involved group, as they differ in the mastery experience and therefore in their learning (Figure 1).

If I find support for my null hypotheses, it will indicate that the scales used cannot adequately capture the sustainability competencies resulting from an innovative ESD intervention, but also question whether the current approaches to ESD interventions are adequate. If the results are in line with my alternative hypotheses, this will demonstrate the importance of innovative ESD interventions in fostering human agency.

1. Predictions

A diagram of a graph

Description automatically generated with medium confidence

1. Workflow

A screenshot of a computer

Description automatically generated

Figure 1: **Influence of an innovative ESD intervention on changes in sustainability competencies.** I analysed how an innovative ESD intervention can influence the sustainability attitudes, sustainability behaviours, and efficacy beliefs of high school students at three different time points of the intervention: pre-intervention (to), immediately post-intervention (t1) and one year post-intervention (t2), comparing an involved group with a control group. **A** Conceptual diagram of my predictions, outlined in terms of sustainability attitudes, sustainability behaviours, and efficacy beliefs. **B** Analytical workflow.

# Methods

In this study, I analysed how an innovative ESD intervention influenced students' sustainability attitudes, sustainability behaviours and efficacy beliefs as components of sustainability competencies over time. To quantitatively test the effect of the ESD intervention, I assessed long-term effects by measuring data at three different time points (pre-intervention, immediately post-intervention and one year post-intervention) using the same self-reported survey on sustainability behaviours and sustainability attitudes. I also complemented the data collection with a survey on efficacy beliefs at the third measurement point. I used the data collected by Pauli (2023) for measurement points one and two. In total, I analysed 206 self-reported surveys.

## Project description KlimaRatSchule

To analyse the changes in sustainability competencies over time, I chose the ESD intervention KlimaRatSchule (KRS). The KRS project aims to promote and establish an active climate protection culture in schools through democratic participation processes. Students investigate the carbon footprint of their own school and carry out a democratic participation process, the results of which are used to develop the school's own climate protection concept (KRS project description, 2024). The project was piloted in two schools in Freiburg in January 2023. The project was designed by Solare Zukunft e.V., IZT, Energie- und Umweltzentrum am Deister e.V. and Green City e.V.. The project can be divided into 3 phases: in the first phase, a group of volunteers (involved group) was formed and collected data on the school's emissions. In the second phase, a micro-citizen report for schools was conducted with randomly selected students from the school, and in the third phase, the project partners developed a roadmap with concrete goals and steps for the school to reduce its carbon footprint based on the students' work in phases one and two. Although the project included the very interesting aspect of participation, unfortunately, due to severe data limitations, the group that was part of the micro-citizen report had to be excluded from the analysis and therefore the analysis of the participatory aspect of the project also had to be excluded. The ESD intervention aimed to achieve a real impact in terms of CO2 emission reduction, while at the same time achieving climate and democratic education.

The ESD intervention can be classified as follows in terms of learning and teaching methods and procedures. The involved group, which was voluntarily participating, assessed different dimensions of the school's CO2 emissions in a self-regulated and self-directed manner. They applied their knowledge and problem-solving skills in a real-world, group-based, solution-oriented environment (REF KRS). Considering all these learning methods and procedures, this ESD intervention can be considered as an innovative ESD intervention (as opposed to more traditional teaching-learning formats) (REF MONI).

## Study design and data collection

The research was designed as an empirical, long-term, quasi-experimental study. To collect the data, a digital survey was created that mapped the constructs of the Theory of Planned Behaviour, which was the same as that used by Pauli (2023) for the first two measurement points. Accordingly, attitudes, subjective norms, perceived behavioural control, intentions and behaviours with regard to climate protection were measured. In addition, the survey was complemented by a scale measuring efficacy construct based on the Triple A framework. Accordingly, personal and collective efficacy beliefs as well as action and aim-focused efficacy beliefs were assessed.

The survey was conducted at a secondary school in the Freiburg area, the ANGELL school, a Montessori centre in the city, which is state-recognised and run as an independent non-profit organisation. Following a quasi-experimental design, two comparison groups were formed for the study. The involved group was part of the voluntary KRS expert group that carried out the CO2 measurements in the school, and the control group did not participate at all. They were classified into one of the groups by being asked in the survey whether they participated in the KRS school group. Both groups were surveyed at three different measurement points related to the progress of the intervention, see Table XX. At all three data collection points, the same self-report survey was administered, with the idea of interviewing the same students at three different points in time. This combination of a longitudinal study and a treatment group-control group comparison promises to be particularly informative about the relationship between project participation and impact (N. Döring & Bortz, 2016, p. 1023; Ssossé et al., 2021, p. 9). (It was not possible to include the potential third group, which was part of the micro-citizen report, because the sample size of that group was too small for data collection. I have therefore excluded this group from all analyses).

There were some limitations in collecting data at the third measurement point. Due to government regulations, the survey could only be conducted in one of the schools participating in the KRS project. Conducting the survey in more than one school outside the same school authority area would have required permission from the school supervisory authority (REF VERBINF FDB), which was not possible due to time constraints. In the school where the survey was carried out, it was important to minimise disruption to the students. Given the challenge of trying to capture the same students as for the previous two data points, it was decided to survey students in grades 10 and 11. At the first two measurement points, at least 85% of the students were in Years 9 and 10 (REF PAULI), thus maximising the likelihood of interviewing the same students a year later, see Table XX for overview of data collection.

Table XX: Overview data collection at the school at three different measurement points.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **# Measurement points** | **Intervention phase** | **Data of data collection** | **Classes surveyed** | **Data collector** | **Presence of investigator** | **N data collected (control group/involved group)** |
| First measurement point | Project start | 09.02.2023 | 7-11 | Data collected by Pauli | Yes | 76 (56/20) |
| Second measurement point | End of the active phase of the project | 10.05.2023 | 7-11 | Data collected by Pauli | Yes | 81 (63/18) |
| Third measurement point | Approximately one year after the intervention | 13.05.2024 – 06.06.2024 | 10-11 | Data collected by Gargya | No | 49 (42/7) |

### Survey tools

All information below refers to the third measurement point. The survey was conducted in German, see the German questions used in Appendix XX. The items of a scale were asked together in one block. The items were all in the form of statements to be answered on a 4-point Likert scale (0 = strongly disagree, 1 = somewhat disagree, 2 = somewhat agree, 3 = strongly agree). An 'I don't know/No response' option was provided for each item. The completed survey was entered into the Sco-Scie online survey application and completed online by the students. The suggested test time was 10-15 minutes. In accordance with the state's educational research and data policies, parental and school principal consent was obtained prior to the assessment (see Appendix XX). Participation was voluntary, i.e. students received no credit or financial reward. Participants were assured of complete confidentiality and anonymity. Unfortunately, the researcher could not be present at the third measurement point. Therefore, a slide was prepared with the essential information (see Appendix XX).

#### Sustainability attitudes and sustainability behaviours

In order to ensure the comparability of the data with the first two measurement points and to allow for long-term assessment, the exact same questions were used at the third measurement point, based on Pauli (2023). The scale was designed on the basis of the Theory of Planned Behaviour and included elements of attitudes, subjective norms, perceived behavioural control, intentions and behaviour, see Table XX. Pauli (2023) collected the questions from various literature sources and partly adapted them to a high school context.

Table XX: The scales used to measure sustainability attitudes and sustainability behaviours based on the Theory of Planned Behaviour. Attitudes, subjective norms, perceived behavioural control and intentions were combined to form sustainability attitudes, and self-reported behaviours were taken as sustainability behaviours. The scales have been copied one to one from Pauli (2023) (and only translated back to English here for consistency).

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| # | Scale | Item | Theoretical classification | Question (drawn from original sources in English) | Source |
| 1 | Attitudes towards climate-protecting behaviour | AT1 | Sustainability attitude | The environment in Germany is in danger because of global climate change. | Masud et al. 2016 |
| 2 | AT2 | Current global warming is a natural not manmade phenomenon. |
| 3 | AT3 | Climate change damages the natural environment and wildlife in Germany. |
| 4 | AT4 | I'm willing to pay a certain amount to reduce the im-pact of climate change. |
| 5 | Subjective Norms | SN1 | My family often discusses climate change or global warming. | Lin 2013 |
| 6 | SN2 | My peers often discuss climate change or global warming. |
| 7 | SN3 | My classmates might criticize me if I don't take action to protect the climate. |
| 8 | Perceived Behavioural Control | PBC1 | I believe I can contribute to mitigating the effects of climate change. | Pouya und Niyaz 2022 |
| 9 | PBC2 | I can help reduce carbon emissions through the actions I take in my daily life. |
| 10 | Intentions | IN1 | It is my responsibility to encourage my neighbours to notice climate change. | Pouya und Niyaz 2022 |
| 11 | IN2 | I am willing to adopt and apply eco-friendly practices in my daily life. |
| 12 | IN3 | I am ready to do anything to reduce the impact of climate change. |
| 13 | Behaviours | B01 | Sustainability behaviour | I have reduced my intake of meat over the last month. | Lin 2013 |
| 14 | B02 | I buy fruit produced in Germany and avoid buying imported fruit (e.g. bananas, kiwis). |
| 15 | B03 | I often store food in my refrigerator that is past its use-by date. |
| 16 | B04 | I am more concerned about price when I purchase electrical appliances. |
| 17 | B05 | I purchase electrical appliances that have energy saving labels. |
| 18 | B06 | I turn lights and water taps off as much as possible. |
| 19 | B07 | I unplug appliances that are temporarily not in use. |
| 20 | B08 | I mainly drive or am driven by a car or a motor-scooter. |
| 21 | B09 | I use elevators and rarely use stairs. |
| 22 | B10 | I support an increase of fuel tax to reduce the use of fossil fuels. |

#### Efficacy beliefs

The efficacy scales were based on Hamann et al.'s (2024) questions. Further adaptations for operationalisation in this context were based on their recommendations. For all questions, I included the marker words 'if I/we want to' to avoid the risk of capturing the desirability of the aim instead of their efficacy beliefs. I considered two agents as social identities central to the context of the study. In addition to the individual, I also considered 'we as students' to be a relevant social identity in this context, given the nature of the project to shape the school environment, their primary occupation as students, and the educational goal of the intervention. In order to be able to compare personal and collective efficacy beliefs, I chose three questions and repeated them with the same wording, except that I changed the agent from "I" to "we, as students". In the questions, I included both action and aim-focussed links (see Table XX). I partially adapted some of the questions to the school context by specifying decision-makers or organisations as "school principals or schools" (for the original questions, see Hamann et al. (2024)). Due to time constraints, I could not conduct pilot studies to test the desirability of actions and aims, so I kept the questions as suggested by Hamann et al. (2024), see Table XX for all questions.

Table XX: The scales used for efficacy beliefs are based on the Triple-A framework (Agent-Action-Aim) by Hamann et al. (2024). The scales allow differentiation between personal and collective efficacy beliefs and between aim and action focus.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **#** | **Scale** | **Item** | **Theoretical classification question** | **Question** | **Source** |
| 1 | Personal efficacy beliefs | SW01\_01 | Aim-focussed | I believe that my own actions can make a contribution to climate protection if I want to. | Hamann et al. |
| 2 | SW01\_02 | Aim-focussed | I believe that I can promote climate protection by educating people around me about climate change if I want to. |
| 3 | SW01\_03 | Action-focussed | I don't think I'm in a position to stand up for climate protection. |
| 4 | SW01\_04 | Aim-focussed | I believe that I am in a position to convince others to do more for climate protection if I want to. |
| 5 | SW01\_05 | Aim-focussed | I don't believe that I have any way of influencing climate change. |
| 6 | SW01\_06 | Action-focussed | I believe that I can influence how my school directors or my school acts in relation to climate protection if I want to. |
| 7 | SW01\_07 | Aim-focussed | I believe that I can support my school directors or school in campaigning for climate protection if I want to. |
| 8 | SW01\_08 | Aim-focussed | I believe that I can make a meaningful contribution to climate protection in collaboration with others if I want to. |
| 9 | Collective efficacy beliefs | CS01\_01 | Aim-focussed | We, as students, can make a contribution to climate protection through our actions if we want to. |
| 10 | CS01\_02 | Action-focussed | We, as students, can influence how our school directors or school acts in relation to climate protection if we want to. |
| 11 | CS01\_03 | Aim-focussed | We, as students, are in a position to convince others to do more for climate protection if we want to. |

## Data processing

I conducted all data processing and analyses in R v. 4.0.2. (REF R Core Team, 2021).

Data cleaning and transformation involved several steps, see also Appendix XX. As my analysis was based on (cleaned and transformed) data from Pauli (2023), I carried out very similar steps for my data. I excluded incomplete surveys. I also excluded surveys with more than 25% 'I don't know/No answer' responses. With a total of 32 items, this meant deleting cases with 8 or more unspecified responses. This ensured that each scale could be calculated from as many items as possible. According to Döring and Bortz (2016), individual missing values in the dataset are not a problem for further analysis. To avoid collecting data where students simply clicked through the survey as quickly as possible without answering the questions conscientiously, I excluded surveys with a timestamp of less than 120 seconds. Likert scale responses were coded into numerical values for each item: 0 = strongly disagree, 1 = somewhat disagree, 2 = somewhat agree, 3 = strongly agree, so that in the end a low score indicated a low level of SA/SB/efficacy beliefs and a high score indicated a correspondingly high level. To ensure that this was the case for all questions, items AT2, B03, B04, B08, B09, SW01\_03, SW01\_05 had to be inverted, as they contained negative formulations. Following Döring and Bortz (2016, p. 269), the measurement level of the Likert-scale items was defined as metric, which allowed the calculation of mean values. After reviewing the quality criteria for quantitative research, described below, the scales reflecting the five components of the TPB and personal and collective efficacy beliefs were calculated from the individual items. For all scales, the means of the corresponding items were calculated.

## Data analysis

In order to quantify the relationship between the components of sustainability competencies across different levels of involvement and over time, I used different methods of statistical analysis. In the area of inductive statistics, methods from dependency analysis were used. As this is an explanatory study, the hypothesis test with classical significance test was applied for the inferential statistical evaluation (N. Döring & Bortz, 2016). The significance level α = 0.05 was used for the significance test (Döring and Bortz). If a p-value was below this threshold, it was marked with one asterisk as 'significant', with two asterisks as 'very significant' or with three asterisks as 'highly significant' (cf. Krüger et al., 2014, p. 291), see Table XX.

Table XX: Overview labelling of the p-values in the significance tests

|  |  |  |
| --- | --- | --- |
| **Labelling** | **Designation** | **Threshold values** |
| \* | significant | 0.05 |
| \*\* | very significant | 0.01 |
| \*\*\* | highly significant | 0.001 |

In order to choose the appropriate statistical analysis methods, I first tested the distribution of my data. This was done using the Shapiro-Wilk test, whose null hypothesis is that there is a normal distribution. Almost all scales were below the significance threshold of 0.05, see Appendix XX. This means that the null hypothesis had to be rejected for most scales. The data were therefore largely free of distribution. In order to ensure uniformity and thus comparability, non-parametric tests were used for all scales despite metric scaling, as required by a distribution-free data base (University of Zurich, 2024).

In order to answer some of my research questions (RQ1 and RQ3) by comparing the scale means between groups, I conducted the Wilcoxon test. This test is suitable for distribution-free, interval-scaled dependent samples and determines whether the central tendencies of two dependent samples are different (Uni Zürich, 2024). Given the lack of normal distribution in my data and the need to compare tendencies between groups, I considered the Wilcoxon test appropriate to provide statistical answers to the question of whether the central tendencies of the groups differ. This approach was applied to all scales relevant to my research questions: sustainability attitudes (comprising scales of attitude, intention, perceived behavioural control and subjective norms), sustainability behaviours (all based on the Theory of Planned Behaviour), collective and personal efficacy beliefs, and action and aim-focused efficacy beliefs.

To answer my research question (RQ1) on whether sustainability attitudes and sustainability behaviours differ over time, I conducted the Kruskal-Wallis test. This test is suitable for distribution-free, ordinal-scaled independent samples and determines whether the central tendencies of more than two independent samples are different (University of Zurich). Given the lack of normal distribution in my data and the need to compare three different time points for each SA and SB in each group, I considered this test sufficient to provide statistical answers. I decided not to perform any post-hoc tests as the Kruskal-Wallis test showed no statistical differences between the groups.

To answer my research question (RQ2) regarding the relationship between SA/SB and efficacy beliefs, I conducted the Spearman rank correlation test. This test is suitable for distribution-free, ordinal-scaled samples and calculates the linear correlation between two variables (University of Zurich). The rank correlation can range from -1 to +1, where -1 indicates a perfect negative correlation and +1 indicates a perfect positive correlation. If rs=0 , there is no correlation (Tachtsoglou & König, 2017).

## Quality Criteria testing

In the case of a fully standardised questionnaire survey, it must be subjected to a number of checks to ensure the reliable quality of the results. In order to draw conclusions about the population in the sense of inductive statistics, it must be ensured that the sample is representative of the population. This is an aspect of external validity (N. Döring & Bortz, 2016, p. 104). The KRS school groups met at the school itself and thus became the invovled group in the study, consisting of about 20 students, which fluctuated somewhat during the school year. All the classmates of the students in the KRS group were defined as the basic population of the control group, as they are exposed to similar other conditions and thus comparability is ensured. No information is available on the exact size of this population, so an estimate is given in Table XX. The sample from this control group was determined by the willingness of the students approached to participate. Calculations of ideal sample sizes, based on a 90% confidence interval and a 10% margin of error (REF Qualtrics), show that for measurement point 3, neither the control group nor the involved group was representative of the population, which limits the conclusions that can be drawn about the population, see Table XX.

Table XX: Comparison of ideal and actual sample sizes in relation to the population.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Measurement point 1** | | **Measurement point 2** | | **Measurement point 3** | |
| **Group** | Control group | Involved group | Control group | Involved group | Control group | Involved group |
| **Total population** | 225 | 20 | 225 | 20 | 150 | 20 |
| **Ideal N (90% confidence interval + 10% margin of error)** | > 53 | > 16 | > 53 | > 16 | > 47 | > 16 |
| **Actual N** | 56 | 20 | 63 | 18 | 42 | 7 |

For reliability analysis, Cronbach's alpha, the most common measure of scale reliability (Field, 2012 REF EWM DEV), was calculated to assess the internal consistency of the subscales. Reliabilities for measurement points 1 and 2 were calculated and confirmed by Pauli (2023). The results for measurement point 3 are shown in Table XX below.

Table XX: Reliabilities (Cronbach’s α) for sustainability competency scales at measurement point 3.

|  |  |  |
| --- | --- | --- |
| **Scale** | **Number of items** | **Cronbach’s α** |
| Attitude | 4 | 0.65 |
| Subjective Norms | 3 | **0.41** |
| Perceived behaviour control | 2 | 0.77 |
| Intention | 3 | 0.86 |
| Behaviour1 | 8 | 0.78 |
| Personal efficacy beliefs | 8 | 0.90 |
| Collective efficacy beliefs | 3 | 0.90 |

1 To calculate the cronbach’s α the variables B03 and B04 were omitted, because there were not considered fitting and essential. The content validity should not be jeopardised by this.

In the literature, Cronbach's alpha values above 0.6 are considered good for short scales (≤ 6 items) (Krüger et al., 2014). For longer scales (8 items), values above 0.7 can be considered good (N. Schmitt, 1996). Almost all scales achieved satisfactory results in this test, with the exception of the subjective norms scale. This lower reliability should be taken into account in further analyses, although the literature also points out that these thresholds should not be given too much significance (SchMIDT REF EWM DEV).

The full R script and the link to my GitHub repository can be found in the appendices.

# Results

## 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 sustainability behaviours did not increase for either the involved or the control group (Figure XX) across the 206 surveys (see Figure X). I found no statistical differences between time points within any group (see Appendix XX for statistical results). In line with my predictions, I found that the involved group (n=7) reported (some) statistically higher sustainability attitudes and behaviours (cumulative SA/SB and SB) than the control group one year after the intervention (see Figure X and Table X for statistical outputs). For all scales, I observed a peak at the second measurement point (immediately after the ESD intervention) for the involved group, with scores significantly higher than those of the control group (Figure XX). The involved group reported higher mean scores than the control group for all sustainability competencies at all time points. Before the start of the intervention and one year after the intervention, the reported behaviour, intention and cumulative TPB scores differed significantly between the groups. Surprisingly, for the involved group, all reported sustainability competencies peaked immediately after the ESD intervention (time point 2) but decreased again one year after the intervention (time point 3).

FIGURE: CHANGE GROUP NAMES, INCLUDE Ns, USE NICER COLOURS, SORT ACCORDING TO SA AND SB, RENAME CUMULATED TO COMBINED SA/SB

A screenshot of a graph

Description automatically generated

Figure XX: The components of sustainability competencies differed between the groups of XX (nt1= XX, nt2= XX, nt3= XX) students surveyed using the Wilcoxon test for significance (see Table XX for full test results). Points represent calculated means from self-reported surveys. Error bars indicate 95% confidence intervals. Stars indicate the level of significance obtained from the Wilcoxon signed rank test comparing the groups. See Table XX for the meaning of the number of stars. No significant differences were found between the time points using the Kruskal-Wallis test.

## 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 sustainability behaviours operationalised with the theory of planned behaviour and the scales operationalised with the construct of efficacy beliefs (*r* (47) = .79, p < .001, Figure X, see Table X for statistical results).

FIGURE: CHANGE AXIS LABELS, P VALUE LABEL

A graph of a line with green dots

Description automatically generated

Figure XX: The Theory of Planned Behaviour Sustainability Attitudes and Behaviours scales and the Efficacy Beliefs scales capture the same underlying construct and show a strong positive relationship among the XX students surveyed, tested with Spearman's rank correlation (see Table X for full statistical results). Points represent raw data. The line represents the calculated Spearman's rank correlation.

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

### Personal and collective efficacy beliefs

Contrary to my predictions, I found no statistical differences between personal and collective efficacy beliefs within and between the involved and control groups (Figure XX) among the XX students surveyed at time point 3. The reported mean scores of the involved group were higher than those of the control group for both collective and personal efficacy beliefs. Overall, I did not find any statistical differences and the means were all very close to each other. There were some differences between each pair of questions within the groups (see Figure XX in Appendix X).

FIGURE: LEAVE TITLE OUT, CHANGE COLOURS, CHANGE GROUP NAMES, CHANGE LABELLING DATA COLLECTIVE AND PERSONAL EFFICACY,

A graph with different colored bars

Description automatically generated

Figure XX: Reported personal and collective efficacy beliefs were not significantly different within and between the involved and control groups (see Table XX for full statistical test results) using the Wilcoxon sign ranked test. Bars represent calculated means from the self-report survey. Error bars indicate 95% confidence intervals.

### Aim- and action-focussed efficacy beliefs

In line with my predictions, I found that the involved group reported significantly higher scores for aim-focus than the control group (W = 8, p < .05, Figure X, see Table X for statistical results). Although the involved group also scored higher on the action-focus, there were no statistical differences (W = 1, p > .05). I found no differences between the action and aim scores for the involved group (W = 1, p > .05) or the control group (W = 1, p > .05).

FIGURE: LEAVE TITLE OUT, CHANGE COLOURS, CHANGE GROUP NAMES, CHANGE LABELLING DATA AIM-AND ACTION-FOCUSSED, ADD ERROR BARS, ADD STAR FOR SIGNIFICANT DIFFERENCE

A graph of different colored bars

Description automatically generated

Figure XX: The aim-focused aspect of efficacy beliefs was significantly higher for the involved group than the control group across the XX students surveyed using the Wilcoxon signed rank test. The action-focused aspect and the within-group comparison were not significantly different. Bars represent calculated means from the self-report survey. Error bars indicate 95% confidence intervals.

# Discussion

## Overview

My analysis of 208 self-reported surveys reveals complex heterogeneity of the influence of an innovative ESD intervention on changes in sustainability competencies. On the one hand, contrary to my predictions, I found that one year after the end of the intervention, sustainability attitudes and behaviours did not increase for either the involved or the control group (Figure XX), 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 behaviour than the control group one year after the intervention (Figure XX), indicating positive effects of an innovative ESD intervention on changes in sustainability competencies. Yet, the differences between the groups were already present at the beginning of the intervention (Figure XX). 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 the third point of measurement. 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 XX), 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 XX), 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 XX), 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. 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. 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 hegemony in research is to state one's own positionality and to contextualise the research, as I have tried to do hereIn 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 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 (REF). 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). 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.

A key criticism of the measurement of sustainability competencies is the challenge of adequately operationalising ESD and its 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, 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. 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 & Chapman, 2010; Cattaneo et al., 2014; Drury et al., 2005; Zimmerman, 1995; Louis, La Macchia, et al., 2016). 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 XX). 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 after the intervention (Figure XX), possibly indicating positive effects of an innovative ESD intervention. For all scales, I observed a peak at the second measurement point (immediately after the ESD intervention) for the involved group, with scores significantly higher than those of the control group (Figure XX). At the same time, the cumulative SA/SB and SB were reported to be even significantly higher at the first measurement point (at the beginning of the ESD intervention), which raises the question of attributing sustainability competencies to the ESD intervention (Figure XX). 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 at the beginning of the 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, BUGEN), 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 [71], reconsideration of preconceptions [78], improved ability to solve complex problems related to the environment [75], increased likelihood of identifying environmental issues as personal concerns [73], and relative maintenance of newly acquired positive practices [76] (SOSSE). 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) (REF SOSSE?). There is no quantitative empirical research on the effectiveness of innovative ESD interventions such as the KRS project (MONI). 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 (BUGEN), 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, 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 (REF) of the response options and the influence of the media, which may have a greater influence than ESD interventions, especially in this age group (REF). 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 (REF). 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 (BUGEN). 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 the peak at the second time point to a level similar to the first time point at the third time point, this could also be due to frustration based on feedback from the environment. Research has shown that when basic psychological needs for relatedness (belonging), competence (efficacy), and autonomy (self-determination) are not met and need frustration occurs, this could lead to a decrease in pro-environmental behaviour (WUllenkord). 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 (REF) 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 XX). 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 RQ1). 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. Given that in this study the group was formed but had not yet begun its work at the first measurement point, 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 XX). Although all the mean scores, including those of the control group, were very similar, interestingly I found that for the involved group the personal efficacy beliefs were higher than the collective efficacy beliefs (although the differences were not statistically significant). 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 [31,32]). Indeed, Jugert et al [12] were able to show that 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. Similarly, using a qualitative research approach, Cocking and Drury [11] 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). 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. 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 CO2 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 (REF). 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 XX). 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. 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. 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 (Fishbach and Ferguson, 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 (REF). 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.

# Conclusions

My analysis revealed complex relationships of students’ sustainability competencies and innovate ESD interventions with the .

STILL MISSING

# Appendices

STILL MISSING

## Declaration of consent by parents/ info etc

# Declaration of submission of the Master thesis

I hereby declare that I have written the submitted Master thesis independently, have not used any sources and aids other than those specified and have labelled all content taken from other works as such. The Master thesis submitted is not or was not the subject of another examination procedure, neither in its entirety nor in significant parts.

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Place, date Signature