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

The critical role of education, particularly Education for Sustainable Development (ESD), in addressing escalating environmental and social challenges is widely acknowledged. However, we lack long-term empirical data on how innovative ESD interventions, as a promising educational approach, influence the sustainability competencies of students over time. Despite calls for more comprehensive testing, research has predominantly focused on either traditional learning methods or has utilized inadequate experimental research designs. In this study, I quantify the impact of an innovative ESD intervention, specifically the KlimaRatSchule project, on students' sustainability competencies, measured as sustainability attitudes, behaviours, and efficacy beliefs at one school, by analysing 206 self-reported surveys across three measurement points. The analysis revealed that the treatment group (n=7) demonstrated higher sustainability competencies compared to the control group (n=42) one year post ESD intervention. Additionally, I emphasized the importance of considering efficacy beliefs as a valuable outcome of ESD interventions and as a notable indicator for measuring sustainability competencies. My findings highlight the complexity of measuring sustainability competencies that matter and support calls for analysing and enhancing innovative ESD intervention methods as an educational design that can contribute to fostering human agency.

Contents

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

KRS KlimaRatSchule

MP Measurement point

PBC Perceived behavioural control

RQ Research question

SA Sustainability attitudes

SB Sustainability behaviours

SC Sustainability competencies

TPB Theory of planned behaviour

1 Introduction

1.1 Background

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 an 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 for 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 (Getzin & Singer-Brodowski, 2017; Ssossé et al., 2021). 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 (Riess et al., 2022; Waltner et al., 2021) or on uncomprehensive study designs (Riess et al., 2022). 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 (Hamann et al., 2024), together with the possibility of collecting long-term empirical data (Pauli, 2023) 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.

1.2 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 (Rieß, 2011). This educational approach, endorsed by UNESCO (Rieckmann, 2021), 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 (Wals, 2012). 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, 2023). For older learners, however, the focus should be on ESD 2 in order to cultivate critical thinking and autonomous action (Rieckmann, 2021). 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). 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 (Getzin & Singer-Brodowski, 2017; Wals, 2006). 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 (O'Flaherty & Liddy, 2018). 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 & Anderson, 2014), much research is still based on predominantly positivist epistemologies, with quantitative measures such as pre/post surveys commonly used to evaluate educational interventions (O'Flaherty & Liddy, 2018). Due to resource constraints, I 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.

1.2.1 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 (UNCED, 1992). They provide a normative framework, but there is no operationalisation of ESD outcomes (Wiek et al., 2011). At the local level, there are regional education plans, such as the one in Baden-Württemberg, Germany: “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. Critical analyses of major ESD goal recommendations suggest that sustainability competencies encompass the cognitive abilities, skills, and the related motivational, volitional, and social preparedness necessary for addressing sustainability-related challenges and promoting sustainable development in personal, social, and institutional contexts (Rieß et al., 2018). Although some ESD learning goals exist, there is still a lack of operationalisation of ESD outcomes to be translated into measurement models and tools (Rieß et al., 2018). 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 et al., 2018), which includes cognitive, affective motivational (= attitudinal), and behavioural dimensions 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 (Lambrechts et al., 2013), being adaptable to different contexts (Waltner et al., 2019), and countering criticism of the dominance of the cognitive dimensions (Rieckmann & Barth, 2016). 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).

1.2.2 Operationalisation of 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 (Klieme & Hartig, 2008). Despite existing efforts, there is still a need to develop adequate measurement instruments for the different dimensions of sustainability competencies (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 (Frick et al., 2004)). For the attitudinal domain, tools such as the Greenpeace Sustainability Barometer (Grunenberg et al., 2012) are used, and for the behavioural dimension, the General Ecological Behaviour Scale (Kaiser et al., 2001) is a notable example. Recent research has attempted to combine these scales and adapt them to the school context (Waltner et al., 2021). At the same time, it remains uncertain as 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 (Kollmuss & Agyeman, 2002; Terlau & Hirsch, 2015). 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 and should be explored further (Kagawa, 2007). 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) (Ajzen, 1991) and the Norm Activation Model (Bamberg & Möser, 2007) .The TPB is one of the most prominent from-within approaches (Ertz et al., 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). 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. Behind these determinants are beliefs or convictions that relate to behaviour, norms and control (Ajzen, 1991). The TPB, as a rational choice theory, focuses on the self-interest of the person weighing the consequences of their actions (Kaiser et al., 2005). Research assessed the explanatory power of the TPB for pro-environmental behaviour. 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 (Kaiser et al., 2005) . However, in a meta-analysis, this average was significantly lower at 27% variance explanation (Armitage & Conner, 2001). On the downside, the directions of effect between attitudes, subjective norms and perceived behavioural control remain incompletely identified in the TPB (Kaiser et al., 2005). 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 et al., 2021), as well as moral and normative considerations (Kaiser et al., 2005). 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. 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.

1.2.3 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 (Waltner et al., 2019). 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 (Waltner et al., 2022). At present, empirical data on the long- and short-term impacts of ESD initiatives in educational settings are limited (Riess et al., 2022), 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 call for alternative and innovative approaches to ESD (Tejedor et al., 2019). 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 (Deci & Ryan, 2008). 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 (Riess et al., 2022). 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 (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 (Ajzen, 1991). 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 (Rieß, 2011). 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 (Riess et al., 2022). 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 (Riess et al., 2022). To date, innovative ESD intervention studies mostly include results in the form of self-reports, self-assessments and expert surveys (Riess et al., 2022), 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) (Handley et al., 2018). 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 (Riess et al., 2022). Given the mismatch between learning recommendations and the availability of comprehensive research design methods, there is a need to measure innovative ESD methods using quasi-experimental research 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 (Kaiser et al., 2001). 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 (Waltner et al., 2019). 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 that is considered

a valid proxy for actual behaviour and is assessed using a simple dichotomous approach (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 (Hamann et al., 2024). One study found that the TPB could benefit from incorporating aspects of the Triple A framework to increase its explanatory power (Bamberg & Rees, 2017). The Triple A framework, with its high degree of differentiation, expands the possibility of assessing different behavioural outcomes (Hamann et al., 2024). 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.

1.2.4 The 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 (Hamann et al., 2024). 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 (Bandura, 1978; Skinner, 1991), 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 & Drury, 2004), 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 (Hamann et al., 2024). 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 (Hamann et al., 2024). 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 action (Bandura, 1997). 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 (Hamann et al., 2024). 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). 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 (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, 2017). 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 un-

derstanding 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(van Zomeren et al., 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 (Hamann et al., 2024). Individuals have limited control over collective outcomes, with many obstacles stemming from external factors and the actions of others (Hornsey et al., 2022). 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 an aim is crucial; if an aim is not desirable, efficacy measures may elicit defensiveness and reflect only the value of the aim (Castiglione, 2020). 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 (Hamann et al., 2024). 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 as both aim- and action-focussed efficacy beliefs allows for a better understanding of

the influence of external and structural factors on behaviour.

1.3 Objectives and research questions

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, addressing the following research questions (RQ):

1. How do students' sustainability attitudes and sustainability behaviours change over time (including one year post 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?

1.4 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 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.

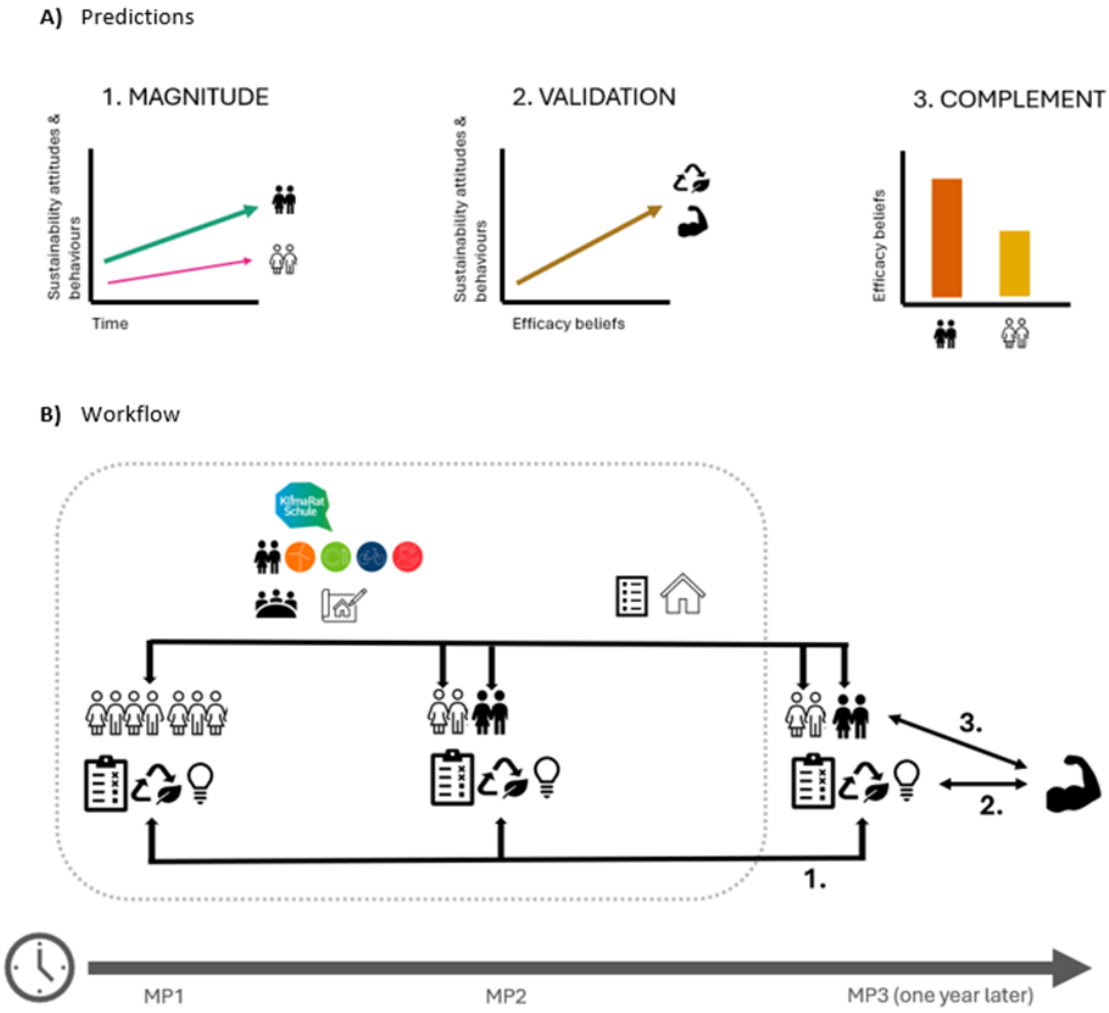

Figure 1: Influence of an innovative ESD intervention on long-term changes in sustainability competencies. I analysed how an innovative ESD intervention can influence the sustainability attitudes, sustainability behaviours, and efficacy beliefs of high school students at three different measurement points (MPs) of the intervention: pre-intervention (MP1), immediately post-intervention (MP2) and one year post-intervention (MP3), comparing an involved group with a control group. **A** Conceptual diagrams of my predictions, outlined in terms of sustainability attitudes, sustainability behaviours, and efficacy beliefs. **B** Analytical Workflow.

2 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 measurement 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.

2.1 Project description KlimaRatSchule

To analyse 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 website, 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 (KRS website, 2024). 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 (= treatment group) was voluntarily participating in the group. They assessed different dimensions of the school's CO2 emissions in a self-regulated and self-directed man-

ner. They applied their knowledge and problem-solving skills in a real-world, group-based, solution-oriented environment (KRS website, 2024). 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) (Riess et al., 2022).

2.2 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. The surveyed students 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 1. 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 (Ssossé et al., 2021).

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 (VerbundFDB, 2022), 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 (Pauli, 2023), thus

maximising the likelihood of interviewing the same students a year later, see Table 1 for overview of data collection.

Table 1: Overview data collection at the school at three different measurement points.

# Measurement points	Intervention phase	Date of data collection	Classes surveyed	Data collector	N data collected (control group/involved group)
Measurement Point 1 (MP1)	Pre-intervention (Project start)	09.02.23	7-11	Data collected by Pauli	76 (56/20)
Measurement Point 2 (MP2)	Immediately post-intervention (end of active phase of project)	10.05.23	7-11	Data collected by Pauli	81 (63/18)
Measurement Point 3 (MP3)	One year post-intervention	13.05.24 – 06.06.24	10-11	Data collected by Gargya	49 (42/7)

2.2.1 Survey tools

All information below refers to the Measurement Point 3. The survey was conducted in German, see the German questions used in Appendix A. 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 B). 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 C).

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 2. Pauli (2023) collected the questions from various literature sources and partly adapted them to a high school context.

Table 2: The scales used to measure sustainability attitudes and sustainability behaviours based on the Theory of Planned Behaviour (Pauli, 2023).

#	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 3). 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 3 for all questions.

Table 3: The scales used for efficacy beliefs based on the Triple-A framework (Hamann et al., 2024).

#	Scale	Item	Theoretical classification question	Question	Source
1	Personal efficacy beliefs	SW01_01	Aim-focussed	I believe that my own actions can contribute to climate protection if I want to.	Hamann et al. (2024)
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 contribute 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.	

2.3 Data processing

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

Data cleaning and transformation involved several steps, see also the R code in Appendix E. 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. 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 mean of the scales, reflecting the five components of the TPB and personal and collective efficacy beliefs, were calculated from the individual items.

2.4 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 (Döring & Bortz, 2016). The significance level alpha = 0.05 was used for the significance test. 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', see Table 4.

Table 4: Overview of p-value labels 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. 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 Zürich, 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 (University of 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 differed 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 Zürich, 2024). Given the lack of normal distribution in my data and the need to compare three different measurement 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 Zürich, 2024). 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.

2.5 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 (Döring & Bortz, 2016). The KRS school group met at the school

itself and thus became the involved 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 were exposed to similar other conditions and thus comparability was ensured. No information was available on the exact size of this population, so an estimate is given in Table 5. 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 (Qualtrics, 2024), show that for measurement point 3, neither the control group nor the involved group were representative of the population, which limits the conclusions that can be drawn about the population, see Table 5.

Table 5: 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 & Miles, 2010), 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 6 below.

Table 6: Reliabilities (Cronbach's alpha) for sustainability competency scales at MP3.

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
Behaviour[1]	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 (above 5 items). For longer scales (8 items), values above 0.7 can be considered good (Krüger et al., 2014).

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 (Schmitt, 1996).

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

3 Results

3.1 Sustainability competencies as sustainability attitudes and sustainability behaviours (Research Question 1)

Contrary to my prediction, I found that one year post ESD intervention, sustainability attitudes and sustainability behaviours did not increase for either the involved or the control group (Figure 2, Table 9) across the 206 surveys. 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 post intervention (Figure 2, Table 10). The involved group reported higher mean scores than the control group for all sustainability competencies at all measurement points. Pre-intervention and one year post-intervention, the reported behaviour, intention and cumulative TPB scores differed significantly between the groups (Table 10). Surprisingly, for the involved group, all reported sustainability competencies peaked immediately post-intervention (MP2) but decreased again one year after the intervention (MP3) (Figure 2).

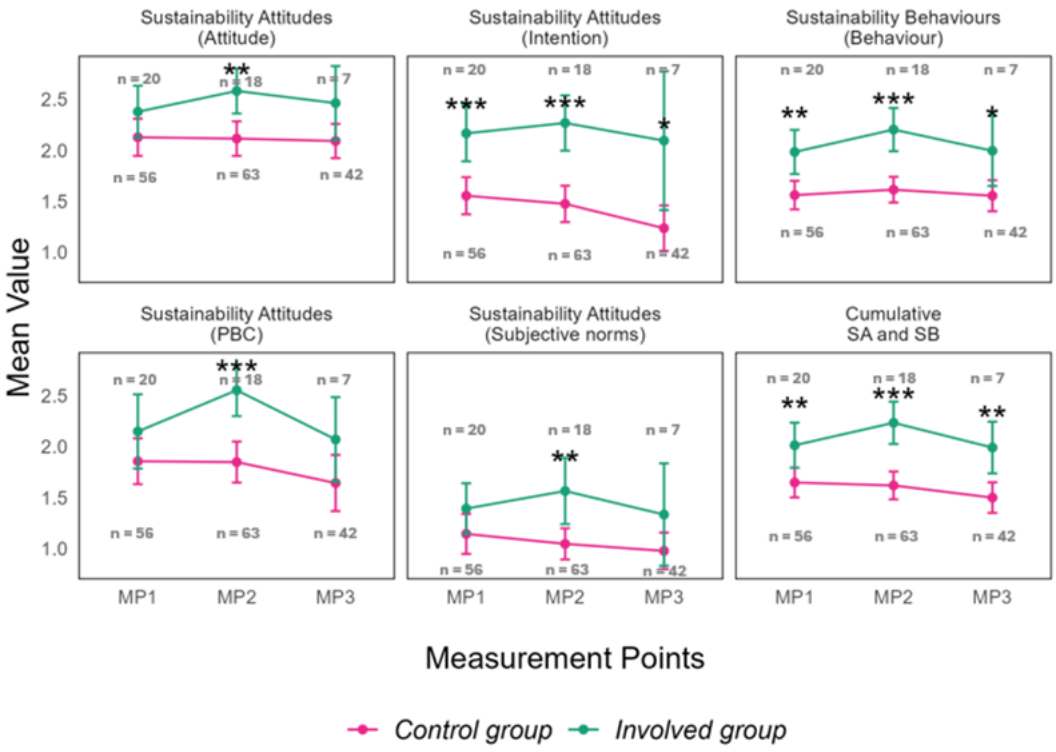

Figure 2: The components of sustainability competencies differed between the groups of 206 students surveyed using the Wilcoxon test for significance (Table 10). Points represent calculated means from self-reported surveys. Error bars indicate 95% confidence intervals. Data labels n indicate sample size. Stars indicate the level of significance obtained from the Wilcoxon signed rank test comparing the groups. See Table 4 for the meaning of the number of stars. No significant differences were found between the measurement points using the Kruskal-Wallis test (Table 9).

3.2 Sustainability attitudes/ sustainability behaviours 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 3).

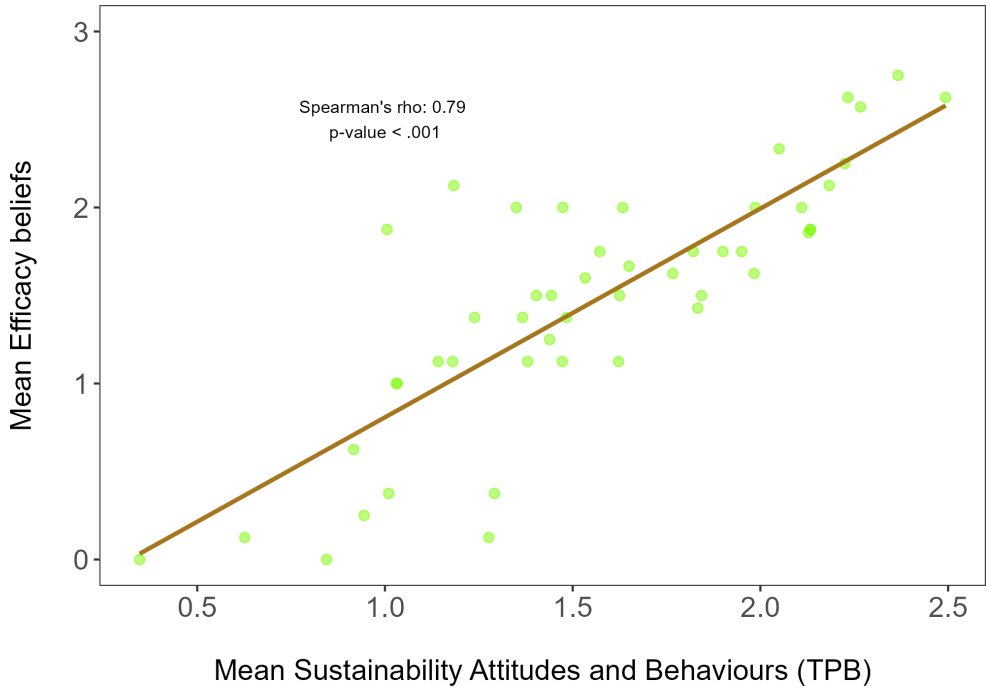

Figure 3: The (TPB-based) Sustainability Attitudes/ Sustainabilitiy Behaviours scales and the (Triple-A-based) Efficacy Beliefs scales capture the same underlying construct and show a strong positive relationship among the 49 students surveyed, tested with Spearman's rank correlation. Points represent raw data. The line represents the calculated Spearman's rank correlation.

3.3 Sustainability competencies as efficacy beliefs (Research Question 3)

3.3.1 Personal and collective efficacy beliefs

Contrary to my predictions, I found no statistical differences between personal ($W = 88.5$, $p > .05$, see Table 12) and collective ($W = 123$, $p > .05$, see Table 12) efficacy beliefs between the involved and the control group among the 49 students surveyed at measurement point 3 (Figure 4). I found no differences between the collective and personal efficacy beliefs within the involved group ($W = 0$, $p > .05$, see Table 11) or within the control group ($W = 1$, $p > .05$, see Table 11).

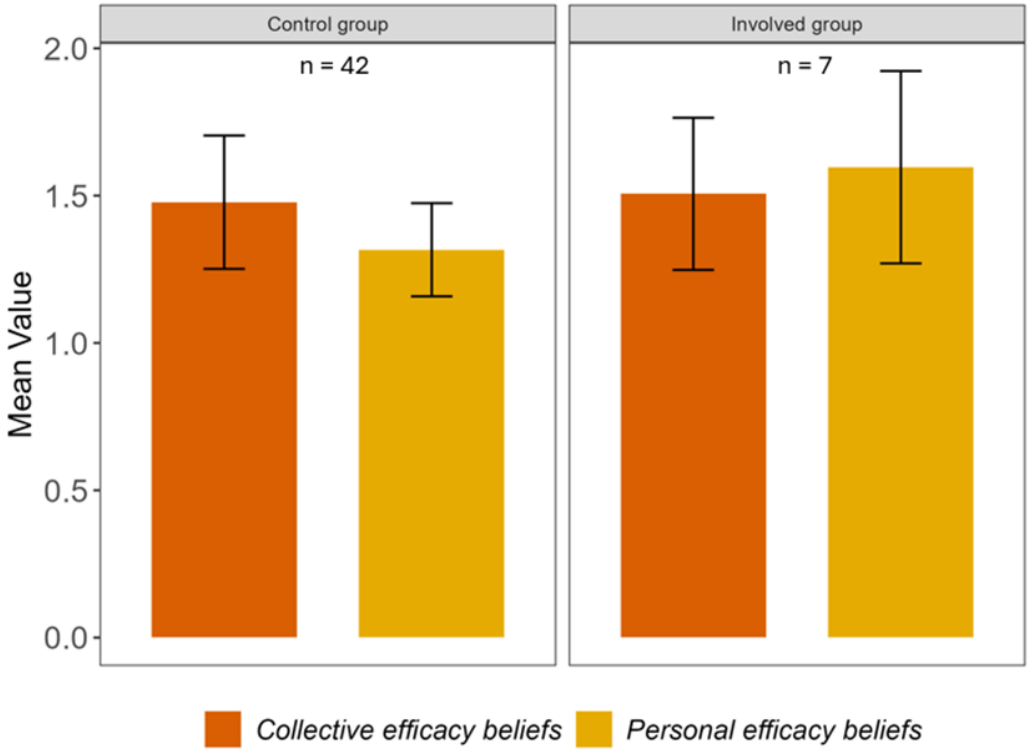

Figure 4: Reported personal and collective efficacy beliefs were not significantly different within and between the involved and control groups (see Tables 11 and 12) using the Wilcoxon sign ranked test at measurement point 3. Bars represent calculated means from the self-report survey. Error bars indicate 95% confidence intervals. Data labels n indicate sample size.

3.3.2 Aim- and action-focussed efficacy beliefs

In line with my predictions, I found that the involved group reported significantly higher scores for aim-focussed efficacy beliefs than the control group ($W = 8$, $p < .05$, see Table 14) among the 49 students surveyed at measurement point 3 (Figure 5). Although the involved group also scored higher on the action-focussed efficacy beliefs, there were no statistical differences ($W = 1$, $p > .05$, see Table 13). I found no differences between the action and aim-focussed scores for within the involved group ($W = 1$, $p > .05$, see Table 13) or within the control group ($W = 1$, $p > .05$, see Table 13).

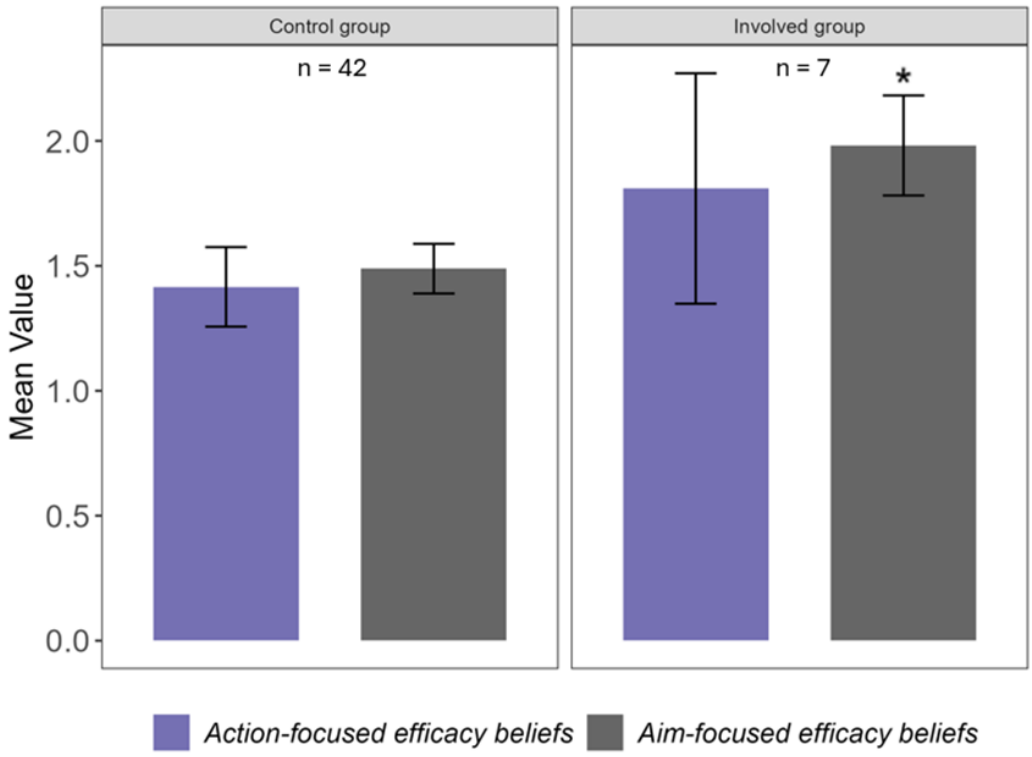

Figure 5: The aim-focused aspect of efficacy beliefs was significantly higher for the involved group than the control group across the 49 students surveyed using the Wilcoxon signed rank test at measurement point 3 (Table 14). The action-focused efficacy beliefs and the within-group comparisons were not significantly different (Tables 13 and 14). Bars represent calculated means from the self-report survey. Error bars indicate 95% confidence intervals. Data labels n indicate sample size. Stars indicate the level of significance obtained from the Wilcoxon signed rank test comparing the groups. See Table 4 for the meaning of the number of stars.

4 Discussion

4.1 Overview

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

The measurement of sustainability competencies has been criticised for various reasons, which have important implications for the interpretation of the results. A major discussion within the field of ESD

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

Another key challenge of the measurement of sustainability competencies is the adequate operationalising of ESD outcomes, including related goals. The method and tool for operationalising sustainability competencies aim at equipping students with core competencies for shaping a sustainable future. However, this raises questions about evaluation, definition and the need for openness in the ESD con-

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

4.2 Sustainability competencies as sustainability attitudes and sustainability behaviours (Research Question 1)

Contrary to my prediction, I found that one year after the ESD intervention, sustainability attitudes and behaviours did not increase for either the involved or the control group (Figure 2). On the other hand, in line with my predictions, I demonstrated that the involved group (n=7) reported statistically higher sustainability attitudes and behaviours (cumulative SA/SB and SB) than the control group one year post-intervention (Figure 2), possibly indicating positive effects of an innovative ESD intervention. For all scales, I observed a peak at the second measurement point (immediately post-intervention) for the involved group, with scores significantly higher than those of the control group (Figure 2). At the same time, the cumulative SA/SB and SB were reported to be even significantly higher at the first measurement point (pre-intervention), which raises the question of attributing sustainability competencies to the ESD intervention (Figure 2). This doubt is reinforced by the fact that SA and SB peaked for the involved group, but then fell back to levels similar to those pre-intervention. There could be several reasons for the observed results. Firstly, the path from an educational intervention to changes in SA and SB remains complex and difficult to predict. Research suggests that sustainability attitudes decrease as students get older (Krettenauer, 2017; Waltner et al., 2021), which could potentially counteract the effects of the intervention. Looking at the effectiveness of other ESD interventions, a meta-study from 2021 found that ESD interventions led to increased environmental sensitivity, reconsideration of preconceptions, improved ability to solve complex problems related to the environment, increased likelihood of identifying environmental issues as personal concerns, and relative maintenance of newly acquired positive practices (Ssossé et al., 2021). Most of these studies focused on small groups and used different operationalisations, making comparisons difficult. There is a consensus in the literature on the need for long-term, longitudinal impact studies that take into account other types of concrete ESD outcomes that can be realised in a sometimes more distant horizon than what current studies can cover (especially activism) (Ssossé et al., 2021). There is no quantitative empirical research on the effectiveness of innovative ESD interventions such as the KRS project (Riess et al., 2022). My findings highlight the importance of conducting sustainability measurement based on long-term empirical and quasi-experimental designs when assessing the effectiveness of (innovative) ESD interventions.

The observed findings could also be due to factors that influenced students' SA and SB more than the ESD intervention itself. A recent long-term study (Waltner et al., 2021), based on the ESD goals in local education plans (and not on a specific ESD intervention), found significant predictors of SA and SB to be average school grade, sustainability-related attitudes at the beginning of the school year,

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

4.3 Sustainability attitudes/ sustainability behaviours and efficacy beliefs (Research Question 2)

I found a strong positive relationship between the scales of sustainability attitudes and behaviours operationalised with the Theory of Planned Behaviour and the scales operationalised with the construct of efficacy beliefs (Figure 3). This suggests that they capture the same latent construct of sustainability competencies and can be used as a first step in a mutual validation of the scales. In a possible second step of validation, the (voluntary and self-determined) participation of the surveyed students in the expert group (which makes them the involved group in this study) could possibly be used to draw conclusions about environmental activism resulting from a high level of environmental attitudes. Participation in this group can be interpreted as a behavioural manifestation aimed at promoting a more sustainable school, which consequently serves to achieve ESD goals. Participation in the involved

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

4.4 Sustainability competencies as efficacy beliefs (Research Question 3)

4.4.1 Collective efficacy beliefs as outcomes

Contrary to my predictions, I found no differences between personal and collective efficacy beliefs within and between the involved and control groups (Figure 4). All the mean scores were very similar. There could be several reasons for the observed results. On the one hand, both personal and collective efficacy beliefs could have been mutually reinforcing, resulting in no detectable difference between them. Studies have found that individuals can derive personal benefits (e.g. efficacy beliefs) from social groups because groups can make them feel personally capable and in control. Collective efficacy made individuals feel in control of their outcomes: People's intention to act was enhanced by a sense of efficacy transferred from the group to the self (Jugert et al., 2016). Similarly, using a qualitative research approach, research (Cocking & Drury, 2004) found that collective efficacy led to a sense of personal empowerment. Thus, collective and personal efficacy are strong and closely intertwined predictors of pro-environmental behaviour and are mutually reinforcing. On the other hand, from a theoretical perspective, the participatory and innovative design of the ESD intervention may have led to an increase in the collective efficacy beliefs of the group involved. Bandura (1997) suggests that efficacy

is enhanced when individuals acquire the specific skills necessary for pro-environmental behaviour and are verbally encouraged about their ability to perform such actions. These sources of efficacy are effective when individuals evaluate the pro-environmental effectiveness of their own groups. Consequently, when a group successfully completes a task and receives positive reinforcement for its pro-environmental efforts, its members are likely to have a stronger belief in their ability to achieve the desired outcomes. Commitment to collective efforts is essential to fostering a belief that their actions can impact their environment. This increased belief, in turn, can lead to a greater willingness to engage in pro-environmental behaviour (Chen, 2015). Given the findings that the involved group did not express more collective efficacy beliefs than the control group, questions arise about the relevance of the chosen group, the desirability of the aims asked in the survey, but also the success of their efforts and the extent of positive external reinforcement. The advantage of differentiation, i.e. the possibility to self-categorise the agent, is undermined by the fact that in a quantitative survey without a pilot study the relevant self-categorisations could not be found out. Empirically, there is very little research that distinguishes between personal and collective efficacy. The very few studies that have made this distinction between personal and collective efficacy beliefs have found that collective efficacy is significantly stronger when the task difficulty is moderate - rather than easy or difficult (Reese & Junge, 2017). Behaviours that are easy to perform tend to have a lower environmental impact. People may perceive that actions that are too easy (e.g. avoiding plastic bags) are unlikely to make a significant difference to environmental problems, even if they are widely practiced. In other words, when actions are too simple, the potential success may not lead to a sense of collective efficacy. This could be an indication that the aim of reducing the school's 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 (van Zomeren et al., 2013). It also raises the question of the extent to which collective efficacy beliefs were the aim of this particular project and were therefore supported and reinforced by the school and the project partners. My findings highlight the relevance of collective efficacy as an outcome indicator of ESD interventions, both in the design of ESD interventions, but also as a relevant indicator for measuring sustainability competencies.

4.4.2 Differentiation of aim- and action-focussed efficacy beliefs as outcomes

In line with my predictions, I found that the involved students reported statistically higher aim-focused efficacy beliefs than the control group (Figure 5). This suggests that the involved group had demon-

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

4.5 Study limitations

Analyses of students' sustainability competencies can be limited by insufficient sample size, limited temporal and geographical scope of the data, and limited empirical validation of the theoretical frameworks used. The lack of data, which prevented me from answering one of my original research questions regarding the participatory nature of the project, highlighted the need for more comprehensive data collection. The generalisability of the empirical findings is severely limited by the small sample

size of students involved at the third measurement point (n=7). The small sample size also affects the statistical power and reliability of the findings. Larger sample sizes generally provide more accurate and generalisable results, reducing the margin of error and increasing confidence in the findings. Furthermore, I argue that the scales used to operationalise sustainability attitudes and behaviours had a limited fit with the context. However, they provided an opportunity to collect longitudinal empirical data. Unfortunately, no prior data were available for the efficacy belief scales, which would be an interesting future direction. Non-Western contexts, theories and empirical evidence were under-represented, limiting applicability to other contexts. Finally, the Triple-A framework, being newly developed, still lacks strong empirical evidence for the distinctions it makes, with potential moderators of the relationship between different efficacy links still missing (Hamann et al., 2024). Recognising heterogeneous effects, including complex behaviours, and analysing what is already practised, highlights previously overlooked research questions and helps researchers to make more strategic choices in the study of efficacy beliefs. The analysis of students' sustainability competencies and their attribution to ESD interventions can benefit from a more comprehensive study design and scope of data, as well as broader theoretical and empirical contexts.

4.6 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. Including 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.

5 Conclusions

My analysis revealed complex relationships between students' sustainability competencies and innovative ESD interventions over time, with the involved group demonstrating higher sustainability attitudes and behaviours one year post-intervention compared to the control group, while also having displayed differences even pre- intervention. These findings highlight the potential of innovative learning and teaching methods within ESD interventions, such as self-regulated and self-directed learning of applicable knowledge and problem-solving skills. I support the latest research advocating for the importance of using long-term empirical data (Waltner et al., 2022) and more robust experimental research designs (Ssossé et al., 2021) to uncover longer-term dynamics that reveal more complex learning outcomes. I further underscored the added value of incorporating efficacy beliefs into the objectives of ESD interventions, as well as in the measurement of sustainability competencies. To fully understand learning outcomes, including real-world behavioural changes, it is insufficient to rely solely on self-reported behaviour intentions and attitudes based on an instrumental understanding of ESD and educational approaches. More nuanced research that distinguishes the capacity for collective agency, as well as the desirability of self-chosen aims and potential external barriers, is necessary. This continued development and empirical research can potentially expand upon the Triple A framework of efficacy beliefs (Hamann et al., 2024). My findings underscore the complexity of measuring sustainability competencies and the heterogeneity of students' responses to them. Consequently, I support calls for more innovative ESD design as a promising avenue for fostering human agency. Accurately quantifying sustainability competencies in relation to ESD interventions will enhance predictions for more effective educational policymaking in the context of our rapidly changing Anthropocene.

6 References

Ajzen, I. (1991). The theory of planned behavior. Organizational Behavior and Human Decision Processes, 50(2), 179–211. https://doi.org/10.1016/0749-5978(91)90020-T

Arcury, T. (1990). Environmental Attitude and Environmental Knowledge. Human Organization, 49(4), 300–304. https://doi.org/10.17730/humo.49.4.y6135676n433r880

Armitage, C. J., & Conner, M. (2001). Efficacy of the Theory of Planned Behaviour: A meta-analytic review. British Journal of Social Psychology, 40(4), 471–499. https://doi.org/10.1348/014466601164939

Bain, P. G., Hornsey, M. J., Bongiorno, R., Kashima, Y., & Crimston, C. R. (2013). Collective Futures: How Projections About the Future of Society Are Related to Actions and Attitudes Supporting Social Change. Personality and Social Psychology Bulletin, 39(4), 523–539. https://doi.org/10.1177/0146167213478200

Bamberg, S., Fischer, D., & Geiger, S. M. (2021). Editorial: The Role of the Individual in the Great Transformation Toward Sustainability. Frontiers in Psychology, 12. https://doi.org/10.3389/fpsyg.2021.710897

Bamberg, S., & Möser, G. (2007). Twenty years after Hines, Hungerford, and Tomera: A new meta-analysis of psycho-social determinants of pro-environmental behaviour. Journal of Environmental Psychology, 27(1), 14–25. https://doi.org/10.1016/j.jenvp.2006.12.002

Bamberg, S., & Rees, J. (2015). Environmental Attitudes and Behavior: Measurement. In J. D. Wright (Ed.), International Encyclopedia of the Social & Behavioral Sciences (Second Edition) (pp. 699–705). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.91066-3

Bamberg, S., & Rees, J. (2017). The impact of voluntary travel behavior change measures – A meta-analytical comparison of quasi-experimental and experimental evidence. Transportation Research Part A: Policy and Practice, 100, 16–26. https://doi.org/10.1016/j.tra.2017.04.004

Bandura, A. (1978). Self-efficacy: Toward a unifying theory of behavioral change. Advances in Behaviour Research and Therapy, 1(4), 139–161. https://doi.org/10.1016/0146-6402(78)90002-4

Bandura, A. (1997). Self-efficacy: The exercise of control. Freeman & Company.

BMBF. (2020, November 17). Die UNESCO veröffentlicht ihre ‘BNE 2030’ Roadmap—BNE-Portal Kampagne. BNE-Portal - BNE-Portal Kampagne. https://www.bne-portal.de/bne/de/news/die-unesco-veroeffentlicht-ihre-bne-2030-roadmap.html

Castiglione, A. (2020). Climate activism: What we know and what more we need to learn. A psychological and sociological review. OSF. https://doi.org/10.31234/osf.io/7brz3

Cattaneo, L. B., Calton, J. M., & Brodsky, A. E. (2014). Status Quo Versus Status Quake: Putting the Power Back in Empowerment. Journal of Community Psychology, 42(4), 433–446. https://doi.org/10.1002/jcop.21619

Chen, M.-F. (2015). Self-efficacy or collective efficacy within the cognitive theory of stress model: Which more effectively explains people's self-reported proenvironmental behavior? Journal of Environmental Psychology, 42, 66–75. https://doi.org/10.1016/j.jenvp.2015.02.002

Cocking, C., & Drury, J. (2004). Generalization of Efficacy as a Function of Collective Action and Intergroup Relations: Involvement in an Anti-Roads Struggle. Journal of Applied Social Psychology, 34(2), 417–444. https://doi.org/10.1111/j.1559-1816.2004.tb02555.x

de Haan, G. (2006). The BLK '21' programme in Germany: A 'Gestaltungskompetenz'□based model for Education for Sustainable Development. Environmental Education Research, 12(1), 19–32. https://doi.org/10.1080/13504620500526362

Deci, E. L., & Ryan, R. M. (2008). Self-determination theory: A macrotheory of human motivation, development, and health. Canadian Psychology / Psychologie Canadienne, 49(3), 182–185. https://doi.org/10.1037/a0012801

Döring, N., & Bortz, J. (2016). Forschungsmethoden und Evaluation in den Sozial- und Humanwissenschaften (5. Aufl.). https://doi.org/10.1007/978-3-642-41089-5

Dumas, M., & Anderson, G. L. (2014). Qualitative research as policy knowledge: Framing policy problems and transforming education from the ground up. Education Policy Analysis Archives, 22, 11–11. https://doi.org/10.14507/epaa.v22n11.2014

Ertz, M., Karakas, F., & Sarigöllü, E. (2016). Exploring pro-environmental behaviors of consumers: An analysis of contextual factors, attitude, and behaviors. Journal of Business Research, 69(10), 3971–3980. https://doi.org/10.1016/j.jbusres.2016.06.010

Field, A., & Miles, J. (2010). Discovering Statistics Using SAS. 1–752.

FMER. (2017). National Action Plan on Education for Sustainable Development—The German contribution to the UNESCO Global Action Programme. Federal Ministry of Education and Research. https://www.unesco.de/sites/default/files/2020-04/NAP_BNE_EN_2019.pdf

Frick, J., Kaiser, F. G., & Wilson, M. (2004). Environmental knowledge and conservation behavior:

Exploring prevalence and structure in a representative sample. Personality and Individual Differences, 37(8), 1597–1613. https://doi.org/10.1016/j.paid.2004.02.015

Getzin, S., & Singer-Brodowski, M. (2017). Transformatives Lernen in einer Degrowth-Gesellschaft. SOCIENCE - Journal of Science-Society Interfaces, 1, 33–46.

Grunenberg, H., Küster, K., & Rode, H. (2012). Greenpeace Nachhaltigkeitsbarometer – Was bewegt die Jugend?

Hamann, K. R. S., & Reese, G. (2020). My Influence on the World (of Others): Goal Efficacy Beliefs and Efficacy Affect Predict Private, Public, and Activist Pro-environmental Behavior. Journal of Social Issues, 76(1), 35–53. https://doi.org/10.1111/josi.12369

Hamann, K. R. S., Wullenkord, M. C., Reese, G., & van Zomeren, M. (2024). Believing That We Can Change Our World for the Better: A Triple-A (Agent-Action-Aim) Framework of Self-Efficacy Beliefs in the Context of Collective Social and Ecological Aims. Personality and Social Psychology Review: An Official Journal of the Society for Personality and Social Psychology, Inc, 28(1), 11–53. https://doi.org/10.1177/10888683231178056

Handley, M. A., Lyles, C. R., McCulloch, C., & Cattamanchi, A. (2018). Selecting and Improving Quasi-Experimental Designs in Effectiveness and Implementation Research. Annual Review of Public Health, 39(Volume 39, 2018), 5–25. https://doi.org/10.1146/annurev-publhealth-040617-014128

Hoffmann, D. T. (2020). Bildung für nachhaltige Entwicklung – Begriff, Merkmale, Aufgaben. Bildung für nachhaltige Entwicklung.

Hornsey, M. J., Blackwood, L., Louis, W., Fielding, K., Mavor, K., Morton, T., O'Brien, A., Paasonen, K.-E., Smith, J., & White, K. M. (2006). Why Do People Engage in Collective Action? Revisiting the Role of Perceived Effectiveness. Journal of Applied Social Psychology, 36(7), 1701–1722. https://doi.org/10.1111/j.0021-9029.2006.00077.x

Hornsey, M. J., Chapman, C. M., & Oelrichs, D. M. (2022). Why it is so hard to teach people they can make a difference: Climate change efficacy as a non-analytic form of reasoning. Thinking & Reasoning, 28(3), 327–345. https://doi.org/10.1080/13546783.2021.1893222

IPCC. (2023). AR6 Synthesis Report: Climate Change 2023 — IPCC. https://www.ipcc.ch/report/sixth-assessment-report-cycle/

Jugert, P., Greenaway, K. H., Barth, M., Büchner, R., Eisentraut, S., & Fritsche, I. (2016). Collective efficacy increases pro-environmental intentions through increasing self-efficacy. Journal of Environmental Psychology, 48, 12–23. https://doi.org/10.1016/j.jenvp.2016.08.003

Kagawa, F. (2007). Dissonance in students' perceptions of sustainable development and sustainability: Implications for curriculum change. International Journal of Sustainability in Higher Education, 8(3), 317–338. https://doi.org/10.1108/14676370710817174

Kaiser, F. G., Byrka, K., & Hartig, T. (2010). Reviving Campbell's Paradigm for Attitude Research. Personality and Social Psychology Review, 14(4), 351–367. https://doi.org/10.1177/1088868310366452

Kaiser, F. G., Doka, G., Hofstetter, P., & Ranney, M. A. (2003). Ecological behavior and its environmental consequences: A life cycle assessment of a self-report measure. Journal of Environmental Psychology, 23(1), 11–20. https://doi.org/10.1016/S0272-4944(02)00075-0

Kaiser, F. G., Frick, J., & Stoll-Kleemann, S. (2001). Zur Angemessenheit selbstberichteten Verhaltens: Eine Validitätsuntersuchung der Skala Allgemeinen Ökologischen Verhaltens. Diagnostica, 47(2), 88–95. https://doi.org/10.1026//0012-1924.47.2.88

Kaiser, F. G., Hübner, G., & Bogner, F. X. (2005). Contrasting the Theory of Planned Behavior With the Value-Belief-Norm Model in Explaining Conservation Behavior. Journal of Applied Social Psychology, 35(10), 2150–2170. https://doi.org/10.1111/j.1559-1816.2005.tb02213.x

Klieme, E., & Hartig, J. (2008). Kompetenzkonzepte in den Sozialwissenschaften und im erziehungswissenschaftlichen Diskurs. In M. Prenzel, I. Gogolin, & H.-H. Krüger (Eds.), Kompetenzdiagnostik: Zeitschrift für Erziehungswissenschaft (pp. 11–29). VS Verlag für Sozialwissenschaften. https://doi.org/10.1007/978-3-531-90865-6_2

KMK. (2023). Informationen zum deutschen Bildungssystem. https://www.kmk.org/dokumentation-statistik/informationen-zum-deutschen-bildungssystem.html

Koletsou, A., & Mancy, R. (2011). Which efficacy constructs for large-scale social dilemma problems? Individual and collective forms of efficacy and outcome expectancies in the context of climate change mitigation. Risk Management, 13(4), 184–208. https://doi.org/10.1057/rm.2011.12

Kollmuss, A., & Agyeman, J. (2002). Mind the Gap: Why do people act environmentally and what are the barriers to pro-environmental behavior? Environmental Education Research, 8(3), 239–260. https://doi.org/10.1080/13504620220145401

Kormos, C., & Gifford, R. (2014). The validity of self-report measures of proenvironmental behavior: A meta-analytic review. Journal of Environmental Psychology, 40, 359–371. https://doi.org/10.1016/j.jenvp.2014.09.003

Krettenauer, T. (2017). Pro-Environmental Behavior and Adolescent Moral Development. Journal of Research on Adolescence, 27(3), 581–593. https://doi.org/10.1111/jora.12300

KRS website. (2024). KRS Project description. 1643125883s Webseite! https://www.klimaratschule.de/das-projekt/

Krüger, D., Parchmann, I., & Schecker, H. (2014). Methoden in der naturwissenschaftsdidaktischen Forschung. https://doi.org/10.1007/978-3-642-37827-0

Kruglanski, A., & Higgins, E. (2007). The Goal Construct in Social Psychology. In Social Psychology of Education (Vol. 1, pp. 334–352). https://doi.org/10.1007/BF02333407

Lambrechts, W., Mulà, I., Ceulemans, K., Molderez, I., & Gaeremynck, V. (2013). The integration of competences for sustainable development in higher education: An analysis of bachelor programs in management. Journal of Cleaner Production, 48, 65–73. https://doi.org/10.1016/j.jclepro.2011.12.034

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. American Psychologist, 57(9), 705–717. https://doi.org/10.1037/0003-066X.57.9.705

Ministry of Education Baden Württemberg. (2016). Bildung und Kultur – Statistisches Landesamt Baden-Württemberg. https://www.statistik-bw.de/BildungKultur/

Moustakas, C. (1994). Phenomenological research methods. SAGE Publications, Inc. https://doi.org/10.4135/9781412995658

O'Flaherty, J., & Liddy, M. (2018). The impact of development education and education for sustainable development interventions: A synthesis of the research. Environmental Education Research, 24(7), 1031–1049. https://doi.org/10.1080/13504622.2017.1392484

Pauli, L. S. (2023). Wirkungsevaluation des BNE-Projekts „KlimaRatSchule": Eine quantitative Analyse mithilfe der Theory of Planned Behavior. (unpublished Master thesis). Katholische Universität Eichstätt-Ingolstadt.

ProBiKlima. (2024). https://www.ph-freiburg.de/recce/projekte/forschungsprojekte/probiklima.html

Qualtrics. (2024). Sample Size Calculator (Stichprobenrechner) I Qualtrics. Qualtrics. https://www.qualtrics.com/de/erlebnismanagement/marktforschung/sample-size-calculator/

R Core Team. (2021). R: The R Project for Statistical Computing. R: A Language and Enviroment for Statistical Computing. https://www.r-project.org/

Reese, G., & Junge, E. A. (2017). Keep on Rockin' in a (Plastic-)Free World: Collective Efficacy and Pro-Environmental Intentions as a Function of Task Difficulty. Sustainability, 9(2), Article 2. https://doi.org/10.3390/su9020200

Rieckmann, M. (2021). Reflexion einer Bildung für nachhaltige Entwicklung aus bildungstheoretischer Perspektive. Religionspädagogische Beiträge, 44(2), Article 2. https://doi.org/10.20377/rpb-153

Rieckmann, M., & Barth, M. (2016). Kompetenzentwicklungsprozesse in der Bildung für nachhaltige Entwicklung erfassen – Überblick über ein heterogenes Forschungsfeld.

Rieckmann, M., & Barth, M. (2022). Educators' Competence Frameworks in Education for Sustainable Development (pp. 19–26). https://doi.org/10.1007/978-3-030-91055-6_3

Rieß, W. (2011). Bildung für nachhaltige Entwicklung: Theoretische Analysen und empirische Studien. Waxmann Verlag.

Riess, W., Martin, M., Mischo, C., Kotthoff, H.-G., & Waltner, E.-M. (2022). How Can Education for Sustainable Development (ESD) Be Effectively Implemented in Teaching and Learning? An Analysis of Educational Science Recommendations of Methods and Procedures to Promote ESD Goals. Sustainability, 14(7), Article 7. https://doi.org/10.3390/su14073708

Rieß, W., Mischo, C., & Waltner, E.-M. (2018). Ziele einer Bildung für nachhaltige Entwicklung in Schule und Hochschule: Auf dem Weg zu empirisch überprüfbaren Kompetenzen. GAIA - Ecological Perspectives for Science and Society, 27(3), 298–305. https://doi.org/10.14512/gaia.27.3.10

Schmitt, N. (1996). Uses and abuses of coefficient alpha. Psychological Assessment, 8(4), 350–353. https://doi.org/10.1037/1040-3590.8.4.350

Skinner, E. A. (1991). Development and perceived control: A dynamic model of action in context. In Self processes and development (pp. 167–216). Lawrence Erlbaum Associates, Inc.

Ssossé, Q., Wagner, J., & Hopper, C. (2021). Assessing the Impact of ESD: Methods, Challenges, Results. Sustainability, 13, 2854. https://doi.org/10.3390/su13052854

Tajfel, H. (1978). Differentiation Between Social Groups: Studies in the Social Psychology of Intergroup Relations. European Association of Experimental Social Psychology by Academic Press.

Tejedor, G., Segalàs, J., Barrón, Á., Fernández-Morilla, M., Fuertes, M. T., Ruiz-Morales, J., Gutiérrez, I., García-González, E., Aramburuzabala, P., & Hernández, À. (2019). Didactic Strategies to Promote Competencies in Sustainability. Sustainability, 11(7), Article 7. https://doi.org/10.3390/su11072086

Terlau, W., & Hirsch, D. (2015). Sustainable Consumption and the Attitude-Behaviour-Gap Phenomenon—Causes and Measurements towards a Sustainable Development. Proceedings in Food System Dynamics, 0, Article 0. https://doi.org/10.18461/pfsd.2015.1516

UNCED. (1992). United Nations Conference on Environment & Development Rio de Janerio, Brazil,

3 to 14 June 1992, Agenda 21. 1992. https://sustainabledevelopment.un.org/content/documents/Agenda21.pdf

University of Zürich. (2024). Datenanalyse mit SPSS | Methodenberatung | UZH. https://www.methodenberatung.uzh.ch/de/datenanalyse_spss.html

van Zomeren, M., Pauls, I. L., & Cohen-Chen, S. (2019). Is hope good for motivating collective action in the context of climate change? Differentiating hope's emotion- and problem-focused coping functions. Global Environmental Change, 58, 101915. https://doi.org/10.1016/j.gloenvcha.2019.04.003

van Zomeren, M., Saguy, T., & Schellhaas, F. M. H. (2013). Believing in "making a difference" to collective efforts: Participative efficacy beliefs as a unique predictor of collective action. Group Processes & Intergroup Relations, 16(5), 618–634. https://doi.org/10.1177/1368430212467476

VerbundFDB. (2022). Genehmigungen für Erhebungen an Schulen – eine länderspezifische Übersicht. Verbund Forschungsdaten Bildung. https://www.forschungsdaten-bildung.de/genehmigungen

Wals, A. E. J. (2006). The end of ESD… The beginning of transformative learning. Emphasizing the 'E' in ESD. In Proceedings of the Seminar on Education for Sustainable Development (pp. 42–59). https://library.wur.nl/WebQuery/wurpubs/353568

Wals, A. E. J. (2012). Learning Our Way Out of Unsustainability: The Role of Environmental Education. In S. D. Clayton (Ed.), The Oxford Handbook of Environmental and Conservation Psychology (p. 0). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199733026.013.0032

Wals, A. E. J., & Benavot, A. (2017). Can we meet the sustainability challenges? The role of education and lifelong learning. European Journal of Education, 52(4), 404–413. https://doi.org/10.1111/ejed.12250

Waltner, E.-M., Overbeck, A., & Rieß, W. (2022). Application-Oriented Development of Outcome Indicators for Measuring Students' Sustainability Competencies: Turning from Input Focus to Outcome Orientation. In G. Karaarslan-Semiz (Ed.), Education for Sustainable Development in Primary and Secondary Schools: Pedagogical and Practical Approaches for Teachers (pp. 205–219). Springer International Publishing. https://doi.org/10.1007/978-3-031-09112-4_15

Waltner, E.-M., Rieß, W., & Mischo, C. (2019). Development and Validation of an Instrument for Measuring Student Sustainability Competencies. Sustainability, 11, 1717. https://doi.org/10.3390/su11061717

Waltner, E.-M., Rieß, W., Mischo, C., Hörsch, C., & Scharenberg, K. (2021). Abschlussbericht: Bildung

für nachhaltige Entwicklung – Umsetzung eines neuen Leitprinzips und seine Effekte auf Schüler/-innenseite. https://phfr.bsz-bw.de/frontdoor/index/index/docId/877

Wiek, A., Withycombe, L., & Redman, C. L. (2011). Key competencies in sustainability: A reference framework for academic program development. Sustainability Science, 6(2), 203–218. https://doi.org/10.1007/s11625-011-0132-6

Wullenkord, M. C., Tröger, J., Hamann, K. R. S., Loy, L. S., & Reese, G. (2021). Anxiety and climate change: A validation of the Climate Anxiety Scale in a German-speaking quota sample and an investigation of psychological correlates. Climatic Change, 168(3), 20. https://doi.org/10.1007/s10584-021-03234-6

7 Appendices

7.1 Appendix A: German scales used

Table 7: The German scales used to measure sustainability attitudes and sustainability behaviours based on the Theory of Planned Behaviour.

#	Skala	Item	Theoretische Einordnung	Verwendete Frage auf Deutsch	Quelle
1	**Einstellungen zu klimaschützendem Verhalten**	AT1	Nachhaltigkeits-Einstellungen	Die Umwelt in Deutschland ist durch den globalen Klimawandel gefährdet.	Masud et al. 2016
2		AT2		Die derzeitige globale Erwärmung ist NICHT vom Menschen verursacht, sondern ein natürlicher Vorgang.	
3		AT3		Der Klimawandel schadet der natürlichen Umwelt und der Tierwelt in Deutschland.	
4		AT4		Ich bin bereit, einen gewissen Betrag zu bezahlen, um die Auswirkungen des Klimawandels zu verringern.	
5	**Subjektive Normen**	SN1		In meiner Familie wird oft über den Klimawandel oder die globale Erwärmung diskutiert.	Lin 2013
6		SN2		Meine Mitschülerinnen und Mitschüler diskutieren oft über den Klimawandel oder die globale Erwärmung.	
7		SN3		Meine Mitschülerinnen und Mitschüler könnten mich kritisieren, wenn ich keine Maßnahmen zum Klimaschutz ergreife.	
8	**Wahrgenommene Verhaltens-kontrolle**	PBC 1		Ich glaube, dass ich dazu beitragen kann, die Auswirkungen des Klimawandels abzuschwächen.	Pouya und Niyaz 2022
9		PBC 2		Durch mein Handeln im Alltag kann ich zur Verringerung von CO2-Ausstoß beitragen.	
10	**Intentionen**	IN1		Es liegt in meiner Verantwortung, meine Mitbürgerinnen und Mitbürger zu ermutigen, den Klimawandel zu beachten.	Pouya und Niyaz 2022
11		IN2		Ich bin bereit dazu, mich in meinem täglichen Leben umweltfreundlicher zu verhalten.	
12		IN3		Ich bin bereit, alles zu tun, um die Auswirkungen des Klimawandels zu mindern.	
13	**Verhalten**	B01	Nachhaltigkeits-Verhalten	Ich habe meinen Fleischkonsum in den letzten Monaten bewusst reduziert.	Lin 2013
14		B02		Ich kaufe in Deutschland produziertes Obst und vermeide den Kauf von importiertem Obst (z. B. Bananen, Kiwis).	
15		B03		In meinem Kühlschrank lagere ich oft Lebensmittel, die das Haltbarkeitsdatum überschritten haben.	
16		B04		Beim Kauf von Elektrogeräten achte ich am meisten auf den Preis der Geräte.	
17		B05		Ich kaufe Elektrogeräte, die ein Energiesparlabel haben.	
18		B06		Ich schalte Lichter und Wasserhähne so oft wie möglich aus.	
19		B07		Ich ziehe den Stecker von Geräten, die vorübergehend nicht in Gebrauch sind.	
20		B08		Ich fahre hauptsächlich mit einem Auto oder einem Motorroller, beziehungsweise werde gefahren.	
21		B09		Ich nutze Aufzüge und steige selten Treppen.	
22		B10		Ich unterstütze eine Erhöhung der Besteuerung von Kraftstoffen, um den Verbrauch fossiler Kraftstoffe zu reduzieren.	

Table 8: The German scales used to measure efficacy beliefs based on the Triple-A framework by Hamann et al. (2024).

#	Skala	Theoretische Einstufung	Item	Verwendete Frage auf Deutsch	Quelle
1	Persönliche Selbstwirksamkeits-überzeugungen	Aim-focussed	SW01_01	Ich glaube, dass meine eigenen Handlungen einen Beitrag zum Klimaschutz leisten können, wenn ich das möchte.	Hamann et al. 2024
2		Aim-focussed	SW01_02	Ich glaube, dass ich den Klimaschutz vorantreiben kann, indem ich in meinem Umfeld über den Klimawandel aufkläre, wenn ich das möchte.	
3		Action-focussed	SW01_03_inverse	Ich glaube nicht, dass ich in der Lage bin, mich für den Klimaschutz einzusetzen.	
4		Aim-focussed	SW01_04	Ich glaube, dass ich dazu in der Lage bin, andere davon zu überzeugen, sich für mehr Klimaschutz einzusetzen, wenn ich das möchte.	
5		Aim-focussed	SW01_05_inverse	Ich glaube nicht, dass ich Möglichkeiten habe, einen Einfluss auf den Klimawandel zu nehmen.	
6		Action-focussed	SW01_06	Ich glaube, dass ich beeinflussen kann, wie meine Schulleitung oder meine Schule bezogen auf den Klimaschutz handelt, wenn ich das möchte.	
7		Aim-focussed	SW01_07	Ich glaube, dass ich meine Schulleitung oder Schule dabei unterstützen kann, sich für Klimaschutz einzusetzen, wenn ich das möchte.	
8		Aim-focussed	SW01_08	Ich glaube, dass ich mich in Zusammenarbeit mit anderen sinnvoll für den Klimaschutz engagieren kann, wenn ich das möchte.	
9	Kollektive Wirksamkeits-überzeugungen	Aim-focussed	CS01_01	Wir, als Schülerinnen und Schüler, können durch unsere Handlungen einen Beitrag zum Klimaschutz leisten, wenn wir das möchten.	
10		Action-focussed	CS01_02	Wir, als Schülerinnen und Schüler, können beeinflussen, wie unsere Schulleitung oder Schule bezogen auf den Klimaschutz handelt, wenn wir das möchten.	
11		Aim-focussed	CS01_03	Wir, als Schülerinnen und Schüler, sind in der Lage, andere davon zu überzeugen, sich für mehr Klimaschutz einzusetzen, wenn wir das möchten.	

7.2 Appendix B: Form Declaration of consent by parents

Einverständniserklärung der Eltern

Teilnahme an einer Untersuchung zum Thema „Nachhaltigkeitskompetenzen im Projekt KlimaRatSchule“

Liebe Eltern,

im Rahmen meiner **Masterarbeit** an der Universität Freiburg im Studiengang „Environmental Governance“ begleite ich, Daniela Gargya, das **Projekt „KlimaRatSchule“** (KRS), welches letztes Jahr an der Schule durchgeführt wurde. Mein Ziel ist es nun die längerfristige Wirkung des Projekts, über das Ende hinausgehend, zu untersuchen.

Inwiefern Schülerinnen und Schüler, welche an dem Projekt teilgenommen haben, veränderte Einstellungen und Selbstwirksamkeitsüberzeugungen, also die Überzeugung Vorgänge verstehen und beeinflussen zu können, aufweisen ist ein wesentlicher Teil der Arbeit.

Hierzu füllen die Schülerinnen und Schüler, welche am Projekt teilgenommen haben, sowie eine „Kontrollgruppe“, einen identischen Fragebogen online aus. In diesem werden Fragen zu Einstellungen, Verhalten und Selbstwirksamkeit abgefragt. So kann der Einfluss des Projektes auf Nachhaltigkeitskompetenzen untersucht werden.

Das Ausfüllen des Fragebogens erfordert ca. 10-15 Minuten und erfolgt in Anwesenheit einer Lehrperson.

Selbstverständlich **erfolgt eine Anonymisierung aller personenbezogenen Daten**, sodass keine Ergebnisse auf einzelne Personen rückführbar sind.

Falls Sie noch Fragen haben, können Sie sich sehr gerne mit mir in Verbindung setzen. Ich danke Ihnen für Ihre Unterstützung.

Mit freundliche Grüßen,

Daniela Gargya

Universität Freiburg
E-Mail: daniela@gargya.de

Albert-Ludwigs-Universität Freiburg

Fakultät für Umwelt und natürliche Ressourcen

Teilnahme an einer Untersuchung zum Thema
„Nachhaltigkeitskompetenzen im Projekt KlimaRatSchule“

Einverständniserklärung der Eltern

Mir ist das Ziel und Wesen der Forschungsprojektes bewusst. Ich hatte die Möglichkeit Fragen zu stellen und habe zurzeit keine weiteren Fragen mehr.

Ich weiß, dass die Teilnahme meines Kindes daran freiwillig ist. Ich weiß, dass ich jederzeit und ohne Angabe von Gründen diese Zustimmung widerrufen kann, ohne dass sich dieser Entschluss nachteilig auf mein Kind auswirken wird. Ich bin damit einverstanden, dass in diesem Forschungsprojekt Daten von meinem Kind in Form von Befragungen erhoben werden. Mir ist bekannt, dass die erhobenen Daten anonym gespeichert und ausschließlich für wissenschaftliche Zwecke verwendet werden.

Hiermit erkläre ich mich einverstanden, dass

Meine Tochter/mein Sohn _____

freiwillig an einer Untersuchung zum Thema „Nachhaltigkeitskompetenzen im Projekt KlimaRatSchule“ teilnimmt.

_____ _____

Ort, Datum Unterschrift Erziehungsberechtigte:r

7.3 Appendix C: Information slide given to students at measurement point 3

7.4 Appendix D: Statistical test results

Table 9: Kruskal-Wallis test results comparing scores across 3 measurement points within groups (RQ1).

Scale	Group	P-Value	Kruskal-Wallis H
AT_Mean	Control group	0.7653188	0.5349257
AT_Mean	Involved group	0.3578105	2.0555033
B_Mean	Control group	0.7167712	0.6659973
B_Mean	Involved group	0.2900163	2.4756363
INT_Mean	Control group	0.1033114	4.5400142
INT_Mean	Involved group	0.9685065	0.0640002
PBC_Mean	Control group	0.4085675	1.7901962
PBC_Mean	Involved group	0.0862698	4.9005516
SN_Mean	Control group	0.7083967	0.6895022
SN_Mean	Involved group	0.6809252	0.7686057
TPB_Mean	Control group	0.2351661	2.8949266

Scale	Group	P-Value	Kruskal-Wallis H
TPB_Mean	Involved group	0.1762751	3.4714192

Table 10: Wilcoxon test results comparing groups across scales (RQ1).

Scale	Measurement Point	P-Value	W Statistic	Significance symbols
AT_Mean	Measurement point 1	0.1203383	429.0	
AT_Mean	Measurement point 2	0.0024719	302.5	**
AT_Mean	Measurement point 3	0.0884757	87.5	
B_Mean	Measurement point 1	0.0021725	300.0	**
B_Mean	Measurement point 2	0.0000515	210.5	***
B_Mean	Measurement point 3	0.0319491	71.5	*
INT_Mean	Measurement point 1	0.0009521	283.0	***
INT_Mean	Measurement point 2	0.0000551	214.5	***
INT_Mean	Measurement point 3	0.0145942	62.0	*
PBC_Mean	Measurement point 1	0.1694962	446.0	
PBC_Mean	Measurement point 2	0.0005760	270.5	***
PBC_Mean	Measurement point 3	0.2163233	104.5	
SN_Mean	Measurement point 1	0.1392966	435.5	
SN_Mean	Measurement point 2	0.0041714	317.0	**
SN_Mean	Measurement point 3	0.1142683	92.0	
TPB_Mean	Measurement point 1	0.0043080	317.5	**
TPB_Mean	Measurement point 2	0.0000257	196.0	***
TPB_Mean	Measurement point 3	0.0082193	54.0	**

Table 11: Wilcoxon test results comparing personal/collective efficacy within groups (RQ3).

Group	W Statistic	P-Value
Control group	1	1
Involved group	0	1

Table 12: Wilcoxon test results comparing personal/collective efficacy between groups (RQ3).

Scale	P-Value	W Statistic
CS_Mean	0.6103249	123.0
SW_Mean	0.0968496	88.5

Table 13: Wilcoxon test results comparing aim/action focused efficacy within groups (RQ3).

Group	W Statistic	P-Value
Control group	1	1
Involved group	1	1

Table 14: Wilcoxon test results comparing aim/action focused efficacy between groups (RQ3).

Theoretical classification	P-Value	W Statistic	Significance symbol
action	0.2000000	1	
aim	0.0135185	8	*

7.5 Appendix E: R Code

R Code, including outputs, can be accessed through GitHub https://github.com/DaniGargya/MEG_thesis.git.

Data formatting and cleaning

```
# Full R script analysing sustainability competencies of students
# Master thesis project 2024, University of Freiburg
# Daniela Gargya
# July 2024

# loading libraries ----
library(tidyverse)
library(ggplot2)
```

```
library(dplyr)
library(haven) # to transform data from sav to csv
library(readr) # to transform data from sav to csv

# data conversion Paulis data from sav to csv ----
#pre angell
lisa_data_angell_pre <- read_sav("data/data_collection/data_pauli/ANGELL_PRE_anonym.SAV")
write_csv(x=lisa_data_angell_pre, path="data/data_collection/data_pauli/angell_pre.csv")

#post angell
lisa_data_angell_post <- read_sav("data/data_collection/data_pauli/ANGELL_POST_anonym.SAV")
write_csv(x=lisa_data_angell_post, path="data/data_collection/data_pauli/angell_post.csv")



# import MP3 data and formating ----
## copied from scoscie website: GNU R-SCript für Daten-Import
ds_file = "rdata_KRSumfrage_2024-06-07_10-21.csv"
options(encoding = "UTF-8")
ds = read.delim(
  file=ds_file, encoding="UTF-8", fileEncoding="UTF-8",
  header = FALSE, sep = "\t", quote = "\"",
  dec = ".", row.names = NULL,
  col.names = c(
    "CASE","SERIAL","REF","QUESTNNR","MODE","STARTED","AT01_01","AT01_02","AT01_03",
    "AT01_04","B001_01","B001_02","B001_03","B001_04","B001_05","B001_06","B001_07",
    "B001_08","B001_09","B001_10","CS01_01","CS01_02","CS01_03","IN01_01","IN01_02",
    "IN01_03","PB01_01","PB01_02","SN01_01","SN01_02","SN01_03","SW01_01","SW01_02",
    "SW01_03","SW01_04","SW01_05","SW01_06","SW01_07","SW01_08","WD01","WD02",
    "WD02_01","WD02_02","WD02_03","TIME001","TIME002","TIME003","TIME004","TIME005",
    "TIME006","TIME007","TIME008","TIME009","TIME_SUM","MAILSENT","LASTDATA",
    "FINISHED","Q_VIEWER","LASTPAGE","MAXPAGE","MISSING","MISSREL","TIME_RSI"
  ),
  as.is = TRUE,
```

```
  colClasses = c(
    CASE="numeric", SERIAL="character", REF="character", QUESTNNR="character",
    MODE="factor", STARTED="character", AT01_01="numeric", AT01_02="numeric",
    AT01_03="numeric", AT01_04="numeric", B001_01="numeric", B001_02="numeric",
    B001_03="numeric", B001_04="numeric", B001_05="numeric", B001_06="numeric",
    B001_07="numeric", B001_08="numeric", B001_09="numeric", B001_10="numeric",
    CS01_01="numeric", CS01_02="numeric", CS01_03="numeric", IN01_01="numeric",
    IN01_02="numeric", IN01_03="numeric", PB01_01="numeric", PB01_02="numeric",
    SN01_01="numeric", SN01_02="numeric", SN01_03="numeric", SW01_01="numeric",
    SW01_02="numeric", SW01_03="numeric", SW01_04="numeric", SW01_05="numeric",
    SW01_06="numeric", SW01_07="numeric", SW01_08="numeric", WD01="numeric",
    WD02="numeric", WD02_01="logical", WD02_02="logical", WD02_03="logical",
    TIME001="integer", TIME002="integer", TIME003="integer", TIME004="integer",
    TIME005="integer", TIME006="integer", TIME007="integer", TIME008="integer",
    TIME009="integer", TIME_SUM="integer", MAILSENT="character",
    LASTDATA="character", FINISHED="logical", Q_VIEWER="logical",
    LASTPAGE="numeric", MAXPAGE="numeric", MISSING="numeric", MISSREL="numeric",
    TIME_RSI="numeric"
  ),
  skip = 1,
  check.names = TRUE, fill = TRUE,
  strip.white = FALSE, blank.lines.skip = TRUE,
  comment.char = "",
  na.strings = ""
)

row.names(ds) = ds$CASE

rm(ds_file)


### assign treatment/ control groups ----
ds <- ds %>%
```

```r
  mutate(Group = ifelse(WD02_01, "group2", ifelse(WD02_03, "group0", "group1")))



# exclude incomplete data ----
# Function to count the number of NAs in a row
count_nas <- function(row) {
  sum(is.na(row))
}

# Apply the function to each row and create a new column 'count_nas'
ds$count_nas <- apply(ds, 1, count_nas)

# Filter the data frame to exclude rows with NAs
ds_filtered <- ds[ds$count_nas <= 1, ]



# exclude data with more than 25% missing (8 answers) ----
# Function to count the number of -1s in a row
count_minus_ones <- function(row) {
  sum(row == -1, na.rm = TRUE)
}

ds_filtered$count_minus_ones <- apply(ds_filtered, 1, count_minus_ones)

# Filter the data frame to exclude rows with more than eight -1s
ds_filtered <- ds_filtered[ds_filtered$count_minus_ones <= 8, ]

# exclude data with less than  2.5 min/ 150sec! processing time (instead of 4)----
ds_filtered <- ds_filtered[ds_filtered$TIME_SUM >= 150, ]



# data transformation # scaling according to Pauli (0-3, instead of 1-4) ----
# Define a function to convert values
```

```r
convert_values <- function(x) {
  x <- ifelse(x == 4, 3,
              ifelse(x == 3, 2,
                     ifelse(x == 2, 1,
                            ifelse(x == 1, 0,
                                   ifelse(x == -1, -100,
                                          x)))))
  return(x)
}


# Apply the function to all columns of the data frame
ds_scaled <- lapply(ds_filtered, convert_values)
ds_scaled <- as.data.frame(ds_scaled)

# invert certain scales to reflect meaning (AT2, B3, B4, B8, B9, SW3, SW5) ----
# Define a function to inverse scales and add new columns
inverse_scale_and_add_columns <- function(df, columns) {
  for (col in columns) {
    new_col_name <- paste(col, "inverse", sep = "_")
    df[[new_col_name]] <- inverse_scale(df[[col]])
  }
  return(df)
}

# Apply the function to add extra columns with inverse scales
ds_scaled_in <- inverse_scale_and_add_columns(ds_scaled, c("AT01_02", "B001_03", "B001_04", "B0

# save formated dataset as csv (angell_mzp3) ----
write_csv(x=ds_scaled_in, path="data/data_collection/angell_mzp3.csv")
```

Data analysis and visualisation

```
# Data analysis and visualisation


# loading libraries ----
library(tidyverse)
library(ggplot2)
library(RColorBrewer)
library(ggthemes)
library(dplyr)
library(purrr)
library(psych) # for cronbach alpha
library(ggrepel)
library(readxl)
library(knitr)


# create colour palette ----
display.brewer.pal(n = 8, name = 'Dark2')
brewer.pal(n = 8, name = "Dark2")


# clean theme for graphs ----
theme_clean <- function(){
  theme_bw() +
    theme(axis.text.x = element_text(size = 14),
          axis.text.y = element_text(size = 14),
          axis.title.x = element_text(size = 14, face = "plain"),
          axis.title.y = element_text(size = 14, face = "plain"),
          panel.grid.major.x = element_blank(),
          panel.grid.minor.x = element_blank(),
          panel.grid.minor.y = element_blank(),
          panel.grid.major.y = element_blank(),
          plot.margin = unit(c(0.5, 0.5, 0.5, 0.5), units = , "cm"),
          plot.title = element_text(size = 15, vjust = 1, hjust = 0.5),
          legend.text = element_text(size = 12, face = "italic"),
```

```
          legend.title = element_text(size = 12, face = "bold"),
          legend.position = c(0.2, 0.8))
}

# importing relevant data ----
mzp1_clean <- read.csv("data/data_collection/data_pauli/angell_pre.csv")
mzp2_clean <- read.csv("data/data_collection/data_pauli/angell_post.csv")
mzp3_clean <- read.csv("data/data_collection/angell_mzp3.csv")


### calculate means MP3 ----
# Define a function to calculate the mean, excluding -100 and NAs
mean_exclude_negative_100 <- function(row) {
  values <- row[row != -100 & !is.na(row)]
  if (length(values) == 0) {
    return(NA)  # Return NA if all values are excluded
  }
  return(mean(values))
}

# Calculate the mean values for specific groups of columns and add them as new columns
#AT_mean
mzp3_clean$AT_mean <- apply(mzp3_clean[, c("AT01_01", "AT01_02_inverse", "AT01_03", "AT01_04")]

#B_mean
mzp3_clean$B_mean <- apply(mzp3_clean[, c("B001_01", "B001_02", "B001_03_inverse", "B001_04_inv

# CS_mean
mzp3_clean$CS_mean <- apply(mzp3_clean[, c("CS01_01", "CS01_02", "CS01_03")], 1, mean_exclude_r

# IN_mean
mzp3_clean$IN_mean <- apply(mzp3_clean[, c("IN01_01", "IN01_02", "IN01_03")], 1, mean_exclude_r
```

```r
# PB_mean
mzp3_clean$PB_mean <- apply(mzp3_clean[, c("PB01_01", "PB01_02")], 1, mean_exclude_negative_100

# SN_mean
mzp3_clean$SN_mean <- apply(mzp3_clean[, c("SN01_01", "SN01_02", "SN01_03")], 1, mean_exclude_n

# SW_mean
mzp3_clean$SW_mean <- apply(mzp3_clean[, c("SW01_01", "SW01_02", "SW01_03_inverse", "SW01_04",

# TPB_mean
mzp3_clean$TPB_mean <- apply(mzp3_clean[, c("AT_mean", "SN_mean", "PB_mean", "IN_mean", "B_mean

# SW_CS_mean
mzp3_clean$SW_CS_mean <- apply(mzp3_clean[, c("SW_mean", "CS_mean")], 1, mean_exclude_negative_


### harmonising and combining dfs ----
# Add a time point indicator to each data frame
mzp1_clean$TP <- "Measurement point 1"
mzp2_clean$TP <- "Measurement point 2"
mzp3_clean$TP <- "Measurement point 3"


# MZP3
# Select only the columns that include '_mean' in their names along with 'TP' and 'Group'
selected_columns_t3 <- mzp3_clean %>%
  select(TP, Group, contains("_mean")) %>%
  rename (Time_Point = TP)


# Convert to long format
long_df_t3 <- selected_columns_t3 %>%
  pivot_longer(
```

```
    cols = contains("_mean"),
    names_to = "Category",
    values_to = "MeanValue")%>%
  mutate(Category = case_when(
    Category == "AT_mean" ~ "AT_Mean",
    Category == "B_mean" ~ "B_Mean",
    Category == "CS_mean" ~ "CS_Mean",
    Category == "IN_mean" ~ "INT_Mean",
    Category == "PB_mean" ~ "PBC_Mean",
    Category == "SN_mean" ~ "SN_Mean",
    Category == "SW_mean" ~ "SW_Mean",
    Category == "TPB_mean" ~ "TPB_Mean",
    Category == "SW_CS_mean" ~ "SW_CS_Mean",
    TRUE ~ Category  # Keep other values unchanged
  ))

# mzp2
# Select only the columns that include '_Mean' in their names along with 'TP' and 'Group'
selected_columns_t2 <- mzp2_clean %>%
  select(TP, Gruppe, contains("_Mean")) %>%
  rename(Time_Point = TP,
         Group = Gruppe) %>%
  mutate(Group = case_when(
   Group == 0 ~ "group0",
    Group == 1 ~ "group1",
   Group == 2 ~ "group2",
   TRUE ~ as.character(Group))) %>%
  mutate(Group = as.factor(Group))

# Convert to long format
long_df_t2 <- selected_columns_t2 %>%
  pivot_longer(
    cols = contains("_Mean"),
```

```
    names_to = "Category",
    values_to = "MeanValue"
  )



# mzp 1
# Select only the columns that include '_mean' in their names along with 'TP' and 'Group'
selected_columns_t1 <- mzp1_clean %>%
  rename(TPB_Mean = TPB_Mittelwert,
         Group = Gruppe,
         Time_Point = TP) %>%
 select(Time_Point, Group, contains("_Mean")) %>%
 mutate(Group = case_when(
  Group == 0 ~ "group0",
  Group == 1 ~ "group1",
  Group == 2 ~ "group2",
  TRUE ~ as.character(Group))) %>%
 mutate(Group = as.factor(Group))

# Convert to long format
long_df_t1 <- selected_columns_t1 %>%
    pivot_longer(
      cols = contains("_Mean"),
      names_to = "Category",
      values_to = "MeanValue"
    )

# Combine the data frames
combined_df <- rbind(long_df_t1, long_df_t2, long_df_t3) #1617 obs.

# Convert Group, Category and TP to factors
combined_df$Group <- as.factor(combined_df$Group)
combined_df$Time_Point <- as.factor(combined_df$Time_Point)
```

```r
combined_df$Category <- as.factor(combined_df$Category)

### excluding group 1 from dataset ----
combined_df_g02 <- combined_df %>%
  filter(Group != "group1")

### check for normality of data ----
### MZP1
# exclude group1
selected_columns_t1 <- selected_columns_t1 %>%
  filter(Group != "group1")

# Find columns with '_Mean' in their name
mean_cols_t1 <- grep("_Mean", names(selected_columns_t1), value = TRUE)

# Function to conduct normality test and return results
normality_test_t1 <- function(col) {
  shapiro_test <- shapiro.test(selected_columns_t1[[col]])
  tibble(Column = col,
         Test = "Shapiro-Wilk",
         W = shapiro_test$statistic,
         P_Value = shapiro_test$p.value,
         Normal = shapiro_test$p.value >= 0.05)
}

# Map function over columns and combine results
normality_results_t1 <- map_dfr(mean_cols_t1, normality_test_t1)

### MZP2
# exclude group1
selected_columns_t2 <- selected_columns_t2 %>%
  filter(Group != "group1")
```

```r
# Find columns with '_Mean' in their name
mean_cols_t2 <- grep("_Mean", names(selected_columns_t2), value = TRUE)

# Function to conduct normality test and return results
normality_test_t2 <- function(col) {
  shapiro_test <- shapiro.test(selected_columns_t2[[col]])
  tibble(Column = col,
         Test = "Shapiro-Wilk",
         W = shapiro_test$statistic,
         P_Value = shapiro_test$p.value,
         Normal = shapiro_test$p.value >= 0.05)
}

# Map function over columns and combine results
normality_results_t2 <- map_dfr(mean_cols_t2, normality_test_t2)

### MZP3
# exclude group1
selected_columns_t3 <- selected_columns_t3 %>%
  filter(Group != "group1")

# Find columns with '_Mean' in their name
mean_cols_t3 <- grep("_mean", names(selected_columns_t3), value = TRUE)

# Function to conduct normality test and return results
normality_test_t3 <- function(col) {
  shapiro_test <- shapiro.test(selected_columns_t3[[col]])
  tibble(Column = col,
         Test = "Shapiro-Wilk",
         W = shapiro_test$statistic,
         P_Value = shapiro_test$p.value,
         Normal = shapiro_test$p.value >= 0.05)
}
```

```r
# Map function over columns and combine results
normality_results_t3 <- map_dfr(mean_cols_t3, normality_test_t3)



### calculating internal validity with cronbach alpha MZP3 ----
# Convert -100 to NA
mzp3_cleaner <- mzp3_clean %>%
  filter(Group != "group1") %>%
  mutate(across(everything(), ~ ifelse(. == -100, NA, .)))

# Get column names with _inverse
columns_with_inverse <- grep("_inverse$", names(mzp3_cleaner), value = TRUE)
# Remove "_inverse" from column names
columns_to_exclude_without_inverse <- gsub("_inverse", "", columns_with_inverse)

filtered_df_cronbach <- mzp3_cleaner %>%
  select(matches("_0[1-9]$|_10$|_inverse"), -matches("WD")) %>%
  select(-one_of(columns_to_exclude_without_inverse)) %>%
  select(-c("B001_03_inverse", "B001_04_inverse"))
#excluding these two items to  increase Cronbachs alpha, see MA Pauli S.70

# handle missing values by using the mean
filtered_df_cronbach <- filtered_df_cronbach %>%
  mutate(across(everything(), ~ ifelse(is.na(.), mean(., na.rm = TRUE), .)))

# Group the columns based on the first two letters of the column name
groups <- split.default(filtered_df_cronbach, sub("^(..).*", "\\1", names(filtered_df_cronbach)

# Calculate Cronbach's alpha for each group
alpha_results <- lapply(groups, function(x) {
  result <- alpha(x)
  result$total$raw_alpha  # Extracting raw alpha from the result
```

```
})



# Create a dataframe with scale names and Cronbach's alpha values
df_alpha <- data.frame(
  Scale = names(alpha_results),
  Alpha = unlist(alpha_results)
)



# Statistical analyses ----
### RQ1 Comparing SA/SB over time and between groups ----
# Kruskal-Wallis test for time points within each category and group
kw_results <- combined_df_g02 %>%
  filter(!Category %in% c("SW_Mean", "CS_Mean", "SW_CS_Mean")) %>%
  group_by(Category, Group) %>%
  summarise(
    P_value = kruskal.test(MeanValue ~ Time_Point)$p.value,
    Kruskal_Wallis_H = kruskal.test(MeanValue ~ Time_Point)$statistic,
    .groups = 'drop') #%>%
  mutate(significance_time = case_when(
      kruskal_p_time < 0.001 ~ "***",
      kruskal_p_time < 0.01 ~ "**",
      kruskal_p_time < 0.05 ~ "*",
      TRUE ~ ""))

kw_results <- kw_results %>%
  mutate(Group = as.character(Group)) %>%
    mutate(Group = case_when(
      Group == "group0" ~ "Control group",
      Group == "group2" ~ "Involved group",
      TRUE ~ Group)) %>%
    rename(Scale = Category)
```

```r
# Create the table with kable
markdown_table_kruskal <- kable(kw_results, format = "markdown",
                          col.names = c("Scale", "Group", "P-Value", "Kruskal-Wallis H"))

# Save the Markdown table to a text file
writeLines(markdown_table_kruskal, "outputs/markdown_table_kruskal.md")

# Wilcoxon rank-sum test for groups within each category and time point
wilcox_results_rq1 <- combined_df_g02 %>%
  filter(!Category %in% c("SW_Mean", "CS_Mean", "SW_CS_Mean")) %>%
  group_by(Category, Time_Point) %>%
  summarise(
    wilcox_p_group = wilcox.test(MeanValue ~ Group)$p.value,
    W_Statistic = wilcox.test(MeanValue ~ Group)$statistic,
    .groups = 'drop'
  ) %>%
  mutate(
    significance_group = case_when(
      wilcox_p_group < 0.001 ~ "***",
      wilcox_p_group < 0.01 ~ "**",
      wilcox_p_group < 0.05 ~ "*",
      TRUE ~ ""
    )
  )

wilcox_results_rq1 <- wilcox_results_rq1 %>%
  rename(Scale = Category)


# Create the table with kable
markdown_table_wilcox_rq1 <- kable(wilcox_results_rq1, format = "markdown",
```

```
                                col.names = c("Scale", "Measurement Point", "P-Value", "W Stati

# Save the Markdown table to a text file
writeLines(markdown_table_wilcox_rq1, "outputs/markdown_table_wilcox_rq1.md")

# Merge results back into the main dataframe
combined_df_g02 <- combined_df_g02 %>%
  left_join(kw_results, by = c("Category", "Group")) %>%
  left_join(wilcox_results, by = c("Category", "Time_Point"))

# Calculate the means for each Group, Competence, and TimePoint
# adding error bars
df_means <- combined_df_g02 %>%
  filter(!Category %in% c("SW_Mean", "CS_Mean", "SW_CS_Mean")) %>% #exclude irrelevant categor
  group_by(Group, Category, Time_Point) %>%
  summarise(MeanValue2 = mean(MeanValue),
            LowerCI = MeanValue2 - qt(0.975, length(MeanValue) - 1) * sd(MeanValue) / sqrt(leng
            UpperCI = MeanValue2 + qt(0.975, length(MeanValue) - 1) * sd(MeanValue) / sqrt(leng
            .groups = 'drop') %>%
  left_join(kw_results, by = c("Category", "Group")) %>%
  left_join(wilcox_results, by = c("Category", "Time_Point"))

# Define a labeller function to rename facets
facet_labeller <- labeller(Category = c(
  AT_Mean = "Sustainability Attitudes\n(Attitude)",
  INT_Mean = "Sustainability Attitudes\n(Intention)",
  PBC_Mean = "Sustainability Attitudes\n(PBC)",
  SN_Mean = "Sustainability Attitudes\n(Subjective norms)",
  B_Mean = "Sustainability Behaviours\n(Behaviour)",
  TPB_Mean = "Cumulative\nSA and SB"))

df_means1 <- df_means %>%
  mutate(Group = recode(Group, "group0" = "Control group", "group2" = "Involved group")) %>%
```

```r
  mutate(Time_Point = recode(Time_Point, "t1" = "MP1", "t2" = "MP2", "t3" = "MP3"))



# Custom colors for the groups
custom_colors <- c("#E7298A", "#1B9E77")

# Add a new factor level to control the order and size of the facets
df_means1$Category <- factor(df_means1$Category, levels = c("AT_Mean", "INT_Mean", "B_Mean", "P

# Create the graph
(rq1_graph_perfect <- ggplot(df_means1, aes(x = Time_Point, y = MeanValue2, group = Group, colo
  geom_line() +
  geom_point() +
  geom_errorbar(aes(ymin = LowerCI, ymax = UpperCI), width = 0.1) +  # Adding error bars
  facet_wrap(~ Category, scales = "fixed", labeller = facet_labeller, ncol = 3) +  # 3x2 grid l
  labs(x = "\nMeasurement Points", y = "\nMean Value") +
  scale_color_manual(values = custom_colors) +  # Custom colors for the groups
  theme_minimal() +
  theme(panel.grid.major = element_blank(),    # Remove major grid lines
        panel.grid.minor = element_blank(),    # Remove minor grid lines
        panel.border = element_rect(fill = NA, color = "black"), # Add border around each facet
        legend.position = "bottom",
        legend.text = element_text(size = 12, face = "italic"),
        legend.title = element_text(size = 12, face = "bold"),
        axis.title.x = element_text(size = 14, face = "plain"),
        axis.title.y = element_text(size = 14, face = "plain")) +
  geom_text_repel(data = df_means_grouped %>% filter(significance_group != ""),
                  aes(x = Time_Point, y = MeanValue2, label = significance_group),
                  vjust = -1.5, color = "black", size = 6, inherit.aes = FALSE, segment.color =
  geom_text_repel(data = df_means_grouped %>% filter(significance_time != ""),
                  aes(x = Time_Point, y = MeanValue2, label = significance_time),
                  vjust = -1.0, color = "blue", size = 3, inherit.aes = FALSE) +
  guides(color = guide_legend(title = NULL)))  # Remove the legend title
```

```r
ggsave(rq1_graph_perfect, file = "outputs/rq1_graph_perfect.png", width = 7, height = 5)


### RQ2 relationship Efficacy and SA/SB (TPB) with spearman correlation at MZP3 ----
# Shapiro-Wilk normality test
shapiro.test(mzp3_cleaner$TPB_mean) #normally distributed p>0.05
shapiro.test(mzp3_cleaner$SW_mean) #not normally distributed p<0.05
shapiro.test(mzp3_cleaner_cs$CS_mean) #not normally distributed p<0.05


# Spearman correlation
spearman_test <- cor.test(mzp3_cleaner$TPB_mean, mzp3_cleaner$SW_mean, method = "spearman")
print(spearman_test)
spearman_cor <- spearman_test$estimate
spearman_pval <- spearman_test$p.value
# p value rejects 0 hypothesis of no correlation -> relevant
# spearman_correlation 0.794 -> strong relationship

# Plot with annotation (SPEARMAN)
(rq2a_graph_cor_tpb_sw <-ggplot(mzp3_cleaner, aes(x = TPB_mean, y = SW_mean)) +
  geom_point(alpha = 0.5, size = 2, color = "#7CFC00") +
  geom_smooth(method = "lm", se = FALSE, color = "#A6761D") +
  ylim(0, 3) +
  theme_clean() +
    annotate("text", size = 3, x = 1, y = 2.5, label = paste("Spearman's rho:", round(spearman_
  labs(x = "\nMean Sustainability Attitudes and Behaviours (TPB)", y = "Mean Efficacy beliefs\r

ggsave(rq2a_graph_cor_tpb_sw, file = "outputs/rq2a_graph_cor_tpb_sw.png", width = 7, height = 5

### RQ3a: Comparing personal and collective efficacy between and within groups at MP3 ----
# data preparation
# Read the Excel file for classifications
```

```r
codebook_sw_cs <- read_excel("data/data_collection/codebook_sw_cs.xlsx")

# Specify the questions of interest explicitly
questions_sw <- c("CS01_01", "CS01_02", "CS01_03", "SW01_01", "SW01_02", "SW01_03_inverse", "SW

# Calculate the mean scores for the specified questions
mean_scores_sw <- colMeans(mzp3_cleaner[, questions_sw], na.rm = TRUE)

# Convert the mean scores to a data frame for plotting
mean_scores_sw_df <- data.frame(
  Question = names(mean_scores_sw),
  Mean_Score = mean_scores_sw
)

# Join the dataframes by the column 'Question'
merged_sw_cs <- left_join(mean_scores_sw_df, codebook_sw_cs, by = "Question")


### comparing between groups SW/CS ----
# check distribution of data
# Perform Shapiro-Wilk tests
shapiro_results_sw_cs <- data.frame(
  Question = questions_sw,
  p_value = sapply(questions_sw, function(question) shapiro.test(mzp3_cleaner[[question]])$p.va
)
## none is normally distributed

# filter out relevant categories
df_tp3 <- combined_df_g02 %>%
  filter(Time_Point == "t3" & Category %in% c("SW_Mean", "CS_Mean"))


### checking wilcoxon for differences in personal/collective between groups
```

```r
wilcox_rq3_sw_cs_between <- df_tp3 %>%
  group_by(Category) %>%
  rename(Scale = Category) %>%
  summarise(wilcox_p_group = wilcox.test(MeanValue ~ Group)$p.value,
            W_Statistic = wilcox.test(MeanValue ~ Group)$statistic,
    .groups = 'drop') #%>%
  mutate(significance_group = case_when(
      wilcox_p_group < 0.001 ~ "***",
      wilcox_p_group < 0.01 ~ "**",
      wilcox_p_group < 0.05 ~ "*",
      TRUE ~ ""))
# no significant differences between the groups

# Create the table with kable
markdown_table_wilcox_rq3_sw_cs_between <- kable(wilcox_rq3_sw_cs_between, format = "markdown",
                                    col.names = c("Scale", "P-Value", "W Statistic"))

# Save the Markdown table to a text file
writeLines(markdown_table_wilcox_rq3_sw_cs_between, "outputs/markdown_table_wilcox_rq3_sw_cs_be

# Calculate means for each Group and Competence
df_means3 <- df_tp3 %>%
  filter(!is.na(MeanValue)) %>%
  group_by(Group, Category) %>%
  summarise(
    MeanValue2 = mean(MeanValue),
    LowerCI = MeanValue2 - qt(0.975, length(MeanValue) - 1) * sd(MeanValue) / sqrt(length(MeanV
    UpperCI = MeanValue2 + qt(0.975, length(MeanValue) - 1) * sd(MeanValue) / sqrt(length(MeanV
    .groups = 'drop') %>%
  left_join(wilcox_results_sw_cs, by = c("Category"))


# comparing personal/collective within groups ----
```

```r
# Aggregate scores for CS and SW questions within each group
aggregate_scores <- mzp3_cleaner %>%
  pivot_longer(cols = c(starts_with("CS"), starts_with("SW")), names_to = "Question", values_to
  mutate(Type = ifelse(str_detect(Question, "CS"), "CS", "SW")) %>%
  drop_na() %>%
  group_by(Group, Type) %>%
  summarize(Aggregate_Score = mean(Score, na.rm = TRUE), .groups = 'drop')

# Reshape data to wide format to ensure paired observations
wide_data <- aggregate_scores %>%
  pivot_wider(names_from = Type, values_from = Aggregate_Score)

# Function to perform Wilcoxon signed-rank test for aggregated scores
wilcoxon_test_aggregated <- function(data, group) {
  filtered_data <- data %>% filter(Group == group)

  cs_scores <- filtered_data$CS
  sw_scores <- filtered_data$SW

  test <- wilcox.test(cs_scores, sw_scores, paired = TRUE)
  return(data.frame(
    Group = group,
    Statistic = test$statistic,
    P_Value = test$p.value
  ))
}

# Perform Wilcoxon tests for overall comparison within each group
results_group0_overall <- wilcoxon_test_aggregated(wide_data, "group0")
results_group2_overall <- wilcoxon_test_aggregated(wide_data, "group2")

# Combine results into a dataframe
wilcoxon_results_overall <- rbind(results_group0_overall, results_group2_overall)
```

```r
# Add Significance column for stars
wilcoxon_results_overall$Significance <- ifelse(wilcoxon_results_overall$P_Value < 0.05, "*", "

# create table
wilcoxon_rq3_sw_cs_within <- wilcoxon_results_overall %>%
  mutate(Group = as.character(Group)) %>%
  mutate(Group = case_when(
    Group == "group0" ~ "Control group",
    Group == "group2" ~ "Involved group",
    TRUE ~ Group))

# Create the table with kable
markdown_table_wilcoxon_rq3_sw_cs_within <- kable(wilcoxon_rq3_sw_cs_within, format = "markdown
                                                  col.names = c("Group", "W Statistic", "P-Value

# Save the Markdown table to a text file
writeLines(markdown_table_wilcoxon_rq3_sw_cs_within, "outputs/markdown_table_wilcox_rq3_sw_cs_w

# Prepare data for visualization
# Function to calculate confidence intervals
calculate_ci <- function(data, conf_level = 0.95) {
  mean_val <- mean(data, na.rm = TRUE)
  stderr <- sd(data, na.rm = TRUE) / sqrt(length(data))
  error_margin <- qnorm((1 + conf_level) / 2) * stderr
  lower_ci <- mean_val - error_margin
  upper_ci <- mean_val + error_margin
  return(data.frame(Mean = mean_val, Lower_CI = lower_ci, Upper_CI = upper_ci))
}

# Apply CI calculation to the mean scores
mean_scores_with_ci <- mzp3_cleaner %>%
  pivot_longer(cols = c(starts_with("CS"), starts_with("SW")), names_to = "Question", values_to
```

```r
  mutate(Type = ifelse(str_detect(Question, "CS"), "CS", "SW")) %>%
  drop_na() %>%
  group_by(Group, Type) %>%
  summarize(Mean_Score = mean(Score, na.rm = TRUE),
            Lower_CI = mean(Score, na.rm = TRUE) - qnorm(0.975) * (sd(Score, na.rm = TRUE) / sq
            Upper_CI = mean(Score, na.rm = TRUE) + qnorm(0.975) * (sd(Score, na.rm = TRUE) / sq



mean_scores_with_ci1 <- mean_scores_with_ci %>%
  mutate(Group = recode(Group, "group0" = "Control group", "group2" = "Involved group")) %>%
  mutate(Type = recode(Type, "CS" = "Collective efficacy beliefs", "SW" = "Personal efficacy be

# Define custom colors2
custom_colors2 <- c("#D95F02", "#E6AB02")

# Plotting the results with annotation and error bars
(rq2b_boxplot_sw_cs_overall <- ggplot(mean_scores_with_ci1, aes(x = Type, y = Mean_Score, fill
  geom_bar(stat = "identity", position = position_dodge(), width = 0.7) +
  geom_errorbar(aes(ymin = Lower_CI, ymax = Upper_CI), width = 0.2, position = position_dodge(0
  facet_wrap(~ Group, scales = "fixed") +
  labs(y = "\n\nMean Value") +
  theme_clean() +
  scale_fill_manual(values = custom_colors2) +
  theme(axis.title.x = element_blank(),
          axis.text.x = element_blank(),
          axis.ticks.x = element_blank(),
          legend.position = "bottom")+
    guides(fill = guide_legend(title = NULL)))

ggsave(rq2b_boxplot_sw_cs_overall, file = "outputs/rq2b_boxplot_sw_cs_overall.png", width = 7,



### RQ3b: Comparing aim- and action-focused efficacy between and within groups at MP3 ----
```

```r
# comparing between groups aim vs action focussed ----
# add new classification only based on aim vs action
merged_group_scores <- merged_group_scores %>%
  mutate(Theoretical_classification2 = case_when(
    grepl("aim", Theoretical_classification, ignore.case = TRUE) ~ "aim",
    TRUE ~ "action"
  ))

# compare within theoretical classifications:
final_analysis2 <- merged_group_scores %>%
  group_by(Theoretical_classification2, Group) %>%
  summarize(Avg_Mean_Score2 = mean(Mean_Score, na.rm = TRUE), .groups = 'drop')


# only two groups to compare within each theoretical classification
wilcoxon_results_aim_action <- merged_group_scores %>%
  group_by(Theoretical_classification2) %>%
  summarize(
    W_test = list(wilcox.test(Mean_Score ~ Group, data = cur_data())),
    .groups = 'drop'
  )

# Extract p-values and test statistics
wilcoxon_results_aim_action$P_Value <- sapply(wilcoxon_results_aim_action$W_test, function(x) x
wilcoxon_results_aim_action$Statistic <- sapply(wilcoxon_results_aim_action$W_test, function(x)

# Add significance stars
wilcoxon_results_aim_action$Significance <- ifelse(wilcoxon_results_aim_action$P_Value < 0.05,

# create table
wilcox_rq3_aim_action_between <- wilcoxon_results_aim_action %>%
  select(-W_test)
```

```r
markdown_table_wilcox_rq3_aim_action_between <- kable(wilcox_rq3_aim_action_between, format = '
                                                      col.names = c("Theoretical classification", "P

# Save the Markdown table to a text file
writeLines(markdown_table_wilcox_rq3_aim_action_between, "outputs/markdown_table_wilcox_rq3_aim

# Merge Wilcoxon test results back into the main dataframe for plotting
merged_group_scores <- merged_group_scores %>%
  left_join(wilcoxon_results_aim_action %>% select(Theoretical_classification2, Significance),

# Apply CI calculation to the mean scores
mean_scores_with_ci2 <- merged_group_scores %>%
  group_by(Group, Theoretical_classification2) %>%
  drop_na() %>%
  summarize(
    Mean_Score = mean(Score, na.rm = TRUE),
    Lower_CI = mean(Score, na.rm = TRUE) - qnorm(0.975) * (sd(Score, na.rm = TRUE) / sqrt(n()))
    Upper_CI = mean(Score, na.rm = TRUE) + qnorm(0.975) * (sd(Score, na.rm = TRUE) / sqrt(n()))
    .groups = 'drop')

mean_scores_with_ci2 <- mean_scores_with_ci2 %>%
  mutate(Group = recode(Group, "group0" = "Control group", "group2" = "Involved group")) %>%
  mutate(Theoretical_classification2 = recode(Theoretical_classification2, "action" = "Action-f
#rename(`Types of Efficacy` = Type)

# Define custom colors2
custom_colors3 <- c("#7570B3", "#666666")

# Making pretty graph
(rq2b_boxplot_aim_action <- ggplot(mean_scores_with_ci2, aes(x = Theoretical_classification2, y
    geom_bar(stat = "identity", position = position_dodge(), width = 0.7) +
    geom_errorbar(aes(ymin = Lower_CI, ymax = Upper_CI), width = 0.2, position = position_dodge
```

```
    facet_wrap(~ Group, scales = "fixed") +
    labs(y = "\n\nMean Value") +
    theme_clean() +
    scale_fill_manual(values = custom_colors3) +
    theme(axis.title.x = element_blank(),
          axis.text.x = element_blank(),
          axis.ticks.x = element_blank(),
          legend.position = "bottom")+
    guides(fill = guide_legend(title = NULL)))

ggsave(rq2b_boxplot_aim_action, file = "outputs/rq2b_boxplot_aim_action.png", width = 7, height

### compare action vs aim within groups ----
# Aggregate scores for aim and action questions within each group
aggregate_scores <- merged_group_scores %>%
  group_by(Group, Theoretical_classification2) %>%
  summarize(Aggregate_Score = mean(Mean_Score, na.rm = TRUE), .groups = 'drop')

# Reshape data to wide format to ensure paired observations
wide_data2 <- aggregate_scores %>%
  pivot_wider(names_from = Theoretical_classification2, values_from = Aggregate_Score)

# Function to perform Wilcoxon signed-rank test for aggregated scores
wilcoxon_test_aggregated2 <- function(data, group) {
  filtered_data <- data %>% filter(Group == !!group)

  aim_scores <- filtered_data$aim
  action_scores <- filtered_data$action

  test <- wilcox.test(aim_scores, action_scores, paired = TRUE)
  return(data.frame(
    Group = group,
    Statistic = test$statistic,
```

```
    P_Value = test$p.value
  ))
}

# Perform Wilcoxon tests for overall comparison within each group
results_group0_aa <- wilcoxon_test_aggregated2(wide_data2, "group0")
results_group2_aa <- wilcoxon_test_aggregated2(wide_data2, "group2")

# Combine results into a dataframe
wilcoxon_results_aim_action2 <- rbind(results_group0_aa, results_group2_aa)

# Add Significance column for stars
wilcoxon_results_aim_action2$Significance <- ifelse(wilcoxon_results_aim_action2$P_Value < 0.05

# create table
wilcoxon_rq3_aim_action_within <- wilcoxon_results_aim_action2 %>%
  mutate(Group = as.character(Group)) %>%
  mutate(Group = case_when(
    Group == "group0" ~ "Control group",
    Group == "group2" ~ "Involved group",
    TRUE ~ Group))

# Create the table with kable
markdown_table_wilcoxon_rq3_aim_action_within <- kable(wilcoxon_rq3_aim_action_within, format =
                                                       col.names = c("Group", "W Statistic", "P-Valu

# Save the Markdown table to a text file
writeLines(markdown_table_wilcoxon_rq3_aim_action_within, "outputs/markdown_table_wilcoxon_rq3_
```

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