DISSERTATION / DOCTORAL THESIS

Titel der Dissertation /Title of the Doctoral Thesis

„Students' types of interest in physics revisited“

verfasst von / submitted by

Mag. rer. nat. Sarah Maria Zöchling

angestrebter akademischer Grad / in partial fulfilment of the requirements for the degree of

Doktorin der Naturwissenschaften (Dr. rer. nat.)

Wien, 2023 / Vienna 2023

Studienkennzahl lt. Studienblatt / degree programme code as it appears on the student record sheet:	UA 796 605 411
Dissertationsgebiet lt. Studienblatt / field of study as it appears on the student record sheet:	Physik / Physics
Betreut von / Supervisor:	Univ.-Prof. Dr. Martin Hopf

ABSTRACT

Fostering students' interest in physics is a crucial part of physics education. Past empirical studies mainly described students' interest in physics in terms of categories of interesting content and contexts and focused on differences between female and male students. One important past empirical study, the IPN interest study, introduced students' types of interest in physics. Yet, past studies did not include modern physics content areas, such as particle physics. Moreover, the IPN study did not describe how interesting different contexts are relative to each other within the students' different types of interest. In addition, other student characteristics, such as physics-related self-concept, are rarely the focus of research. Therefore, physics education research is faced with three important questions, namely (1) into which different types of interest in physics can students be categorised while additionally considering particle physics as a modern physics content area, (2) how interesting are different contexts within these interest types, and (3) are the interest types described better focussing on physics-related self-concept as student characteristics compared to sex. This doctoral research project set up to provide evidence for answering these three questions by conducting two studies on physics education.

In the first study an instrument to measure students' particle physics interest (IPPI) was developed, since studies on the relationship between interest and other aspects of education, such as achievement and self-concept, require the use of psychometrically sound measurement instruments. Using the findings of past empirical research, interest in particle physics was defined and behaviours that correspond to being interested in particle physics were identified. Based on these definitions, a conceptualisation of students' interest in particle physics as a hierarchy of levels of interest in particle physics was proposed. Then, the IPPI was developed using rating scale items that assessed different degrees of being interested in particle physics. A novel approach was suggested and applied to conducting a Rasch analysis for selecting items from an initial item pool in a clear, stepwise, and reproducible way. The IPPI was tested in student think-aloud interviews and validated in a field test with 99 German-speaking grade 9 students. Evidence supporting the content, construct, statistical, and fit validity of the IPPI was provided by a Rasch analysis. Based on the item hierarchy revealed by the Rasch analysis the hypothesised hierarchy of students' levels of interest in particle physics was revised. Each level could be associated with different contexts of particle physics. For example, it was found that students at a certain level of interest (so-called 'open interest') were interested in particle physics when it was set in an everyday life context. Overall, this first study led to the successful development of the IPPI and the conceptualisation of students' interest in particle physics as a hierarchy of levels of interest in particle physics.

Second, the main study was conducted, a cross-cohort study with German-speaking students aged 14 to 16 years (N = 1219). Students' interest in mechanics and particle physics was assessed using the instrument to measure mechanics interest from the IPN study and the IPPI. In addition, different student characteristics, such as their physics-related self-concept, sex, and previous experience with the content areas in school, were assessed. The main aim of this study was to investigate students' types of interest in physics and their association with different student characteristics. For mechanics, the influence of the introductory text was also analysed by randomly assigning the students to two different versions of the questionnaire. The collected data on students' interest and self-concept, was analysed using mixed Rasch models to unveil qualitative differences in the assessed constructs between different groups of students. Moreover, when interpreting the results of the mixed Rasch analyses, differential item functioning as well as students' response styles were considered. The main study showed that most students can be categorised into one single type of interest in both content areas (86% of the students in mechanics and 79% of the students in particle physics, respectively). For both content areas, students of this first interest type are only interested in physics content set in certain contexts, for example, the context 'one's own body'. For mechanics, the second type of interest comprises students who are relatively more interested in physics relating to the motion of cars (14% of the students); and for particle physics, the second type of interest was referred to as the 'particle physics lovers' reflecting their relatively higher interest in particle physics as a content area and as a scientific endeavour (21% of the students). It was found that whether students belong to one or the other type of interest in both content areas can best be described with a model comprising their degree of physics-related self-concept and their sex. The conceptualisation of interest as a hierarchy of students' levels of interest, originally introduced for particle physics only in the first study, was successfully applied to describe the relative interestingness of different contexts for the first type of interest (i.e. the vast majority of students) in both content areas. Hence, the conceptualisation was suggested as a guideline for physics in general and renamed 'hierarchy of students' levels of interest in physics' (HOLIP). The study also showed that previous experience in school with mechanics and particle physics, respectively, is correlated with students' interest in this particular content area for the first type of interest. Finally, it showed that there was no significant difference in students' degree of interest in mechanics associated with the use of different versions of the introductory text on mechanics.

In sum, the doctoral research project led to the successful development of the IPPI (instrument to measure students' particle physics interest). It showed that most students can be categorised into one single type of interest in physics and that their interest can be described using the HOLIP (hierarchy of levels of interest in physics). Moreover, it showed that the interest types can be described best using models that include physics-related self-concept in addition to sex. The results of this doctoral research project have implications for both physics education and physics education research. Educators, such as teachers, can use the HOLIP as a tool to develop learning activities that are interesting for students with different degrees of interest in physics and physics-related self-concept. For physics education research the first study suggests a novel approach to conducting a Rasch analysis for selecting items from an initial item pool in a clear, stepwise, and reproducible way. The main study supports conducting a mixed Rasch analysis to unveil qualitative differences in an assessed construct, such as interest and self-concept, between different groups of students. Moreover, it provides a strong case for considering differential item functioning as well as the students' response styles when interpreting the results of a mixed Rasch analysis.

ZUSAMMENFASSUNG

Das Interesse von Schüler*innen an Physik zu fördern ist ein wesentlicher Bestandteil des Physikunterrichts. Frühere empirische Studien haben das Interesse der Schüler*innen an Physik hauptsächlich in Form von Kategorien interessanter Inhalte und Kontexte beschrieben und sich auf Unterschiede zwischen weiblichen und männlichen Schüler*innen konzentriert. Eine wichtige Studie, die IPN Interessenstudie Physik, führte die Interessentypen der Schüler*innen an Physik ein. Frühere Studien berücksichtigen jedoch keine Inhaltsbereiche der modernen Physik, wie beispielsweise die Teilchenphysik. Außerdem wurde in der IPN-Studie nicht beschrieben, wie interessant verschiedene Kontexte relativ zueinander für die verschiedenen Interessentypen der Schüler*innen sind. Darüber hinaus stehen andere Schüler*innenmerkmale, wie beispielsweise das physikbezogene Selbstkonzept, selten im Mittelpunkt der Forschung. Für die physikdidaktische Forschung stellen sich daher drei wichtige Fragen: In welche verschiedenen Interessentypen lassen sich Schüler*innen einteilen, wenn zusätzlich Teilchenphysik als Inhaltsbereich der modernen Physik berücksichtigt wird? Wie interessant sind verschiedene Kontexte relativ zueinander innerhalb dieser Interessentypen? Und lassen sich die Interessentypen besser beschreiben, wenn man auf das physikbezogene Selbstkonzept als Schüler*innenmerkmal anstelle des Geschlechts fokussiert? Im Rahmen dieses Promotionsprojekts wurden zwei physikdidaktische Studien durchgeführt, um diese drei Fragen beantworten zu können.

In der ersten Studie wurde ein Messinstrument für Interesse an Teilchenphysik (*Instrument to measure Particle Physics Interest*, IPPI) entwickelt, da Studien über den Zusammenhang zwischen Interesse und anderen Bildungsaspekten, wie etwa Leistung und Selbstkonzept, den Einsatz psychometrisch fundierter Messinstrumente erfordern. Basierend auf den Ergebnissen früherer empirischer Studien wurden das Interesse an Teilchenphysik und Verhaltensweisen definiert, die verschiedenen Graden von Interesse an Teilchenphysik entsprechen. Darauf aufbauend wurde eine Konzeptualisierung des Interesses von Schüler*innen an Teilchenphysik in Form einer Hierarchie von Interessensstufen vorgeschlagen. Anschließend wurde das IPPI entwickelt, bestehend aus 11 Rating Scale-Items, die verschiedene Grade des latenten Merkmals „Interesse an Teilchenphysik“ messen. Es wurde eine neue, schrittweise und reproduzierbare Methode vorgeschlagen und angewandt, um mithilfe einer Rasch-Analyse Items aus einem anfänglichen Item-Pool auszuwählen. Das IPPI wurde zuerst in Interviews mit der Methode „Lautes Denken“ getestet und danach in einem Feldtest mit 99 deutschsprachigen Schüler*innen der 9. Schulstufe validiert. Eine Rasch-Analyse erbrachte Evidenzen für die Inhalts-, Konstrukt-, statistische und Fitvalidität des IPPI. Basierend auf der Item-Hierarchie, die aus der Rasch-Analyse resultierte, wurde die vorgeschlagene Konzeptualisierung des Interesses von Schüler*innen an Teilchenphysik in Form einer Hierarchie von Interessensstufen überarbeitet. Insbesondere konnte jede Interessensstufe mit bestimmten Kontexten der Teilchenphysik assoziiert werden. So wurde beispielsweise gezeigt, dass Schüler*innen auf der Stufe des offenen Interesses an Teilchenphysik interessiert sind, wenn sie in Alltagskontexten präsentiert wird. Insgesamt führte diese erste Studie zur erfolgreichen Entwicklung des IPPI und zur Konzeptualisierung des Interesses von Schüler*innen an Teilchenphysik in Form einer Hierarchie von Interessensstufen.

Danach konnte die Hauptstudie, eine Kohortenquerschnittsstudie mit deutschsprachigen Schüler*innen im Alter von 14 bis 16 Jahren (N = 1219), durchgeführt werden. Das Interesse der Schüler*innen an Mechanik und Teilchenphysik wurde mit dem Messinstrument für Interesse an Mechanik aus der

IPN-Studie und dem IPPI erfasst. Darüber hinaus wurden verschiedene Schüler*innenmerkmale, wie etwa das physikbezogene Selbstkonzept, das Geschlecht und die Erfahrung mit dem Inhaltsbereich in der Schule, gemessen. Das Hauptziel dieser Studie war es, die Interessenstypen der Schüler*innen an Physik und ihre Assoziierung mit verschiedenen Schüler*innenmerkmalen zu untersuchen. Für die Mechanik wurde auch der Einfluss des Einführungstextes analysiert, indem die Schüler*innen zufällig zwei verschiedenen Versionen des Fragebogens zugeordnet wurden. Die erhobenen Daten zum Interesse und Selbstkonzept der Schüler*innen wurden mit Mixed Rasch-Modellen analysiert, um qualitative Unterschiede in den gemessenen Konstrukten zwischen verschiedenen Gruppen von Schüler*innen aufzudecken. Außerdem wurden beim Interpretieren der Ergebnisse der Mixed Rasch-Analyse Differential Item Functioning und das Antwortverhalten der Schüler*innen berücksichtigt. Die Hauptstudie zeigte, dass die meisten Schüler*innen in beiden Inhaltsbereichen einem einzigen Interessentyp zugeordnet werden können (86 % der Schüler*innen in Mechanik bzw. 79 % der Schüler*innen in Teilchenphysik). Schüler*innen dieses Interessenstyps waren nur an Physik in den bestimmten Kontexten interessiert. Schüler*innen, die an der Physik der Bewegung von Fahrzeugen interessiert sind, formten den zweiten Typ des Interesses an Mechanik (14 % der Schüler*innen). Der zweite Typ des Interesses an Teilchenphysik wurde als „Teilchenphysikliebhaber*innen“ bezeichnet, um ihr Interesse an Teilchenphysik als Inhaltsbereich und als Forschungsdisziplin widerzuspiegeln (21 % der Schüler*innen). Für beide Inhaltsbereiche konnte am besten beschrieben werden, zu welchem Interessentyp Schüler*innen gehören, mit einem Modell, das sowohl den Grad des physikbezogenen Selbstkonzepts als auch das Geschlecht der Schüler*innen umfasst. Um zu beschreiben, wie interessant verschiedene Kontexte relativ zueinander für den ersten Interessentyp der Schüler*innen (d.h. für die meisten Schüler*innen in Mechanik und Teilchenphysik) sind, konnte die Konzeptualisierung des Interesses in Form einer Hierarchie von Interessenstufen der Schüler*innen angewandt werden, obwohl sie in der ersten Studie ursprünglich nur für Teilchenphysik entwickelt worden war. Deshalb wurde die Konzeptualisierung des Interesses von Schüler*innen in Form einer Hierarchie von Interessensstufen als Empfehlung für Physik im Allgemeinen vorgeschlagen (*hierarchy of levels of interest in physics*, HOLIP). Die Studie zeigte auch, dass die Erfahrung mit dem Inhaltsbereich in der Schule und das Interesse der Schüler*innen an diesem Inhaltsbereich für den ersten Interessentyp von Schüler*innen korrelieren. Außerdem gab es keinen signifikanten Unterschied im Grad des Interesses der Schüler*innen an Mechanik, der mit dem Einsatz von verschiedenen Versionen des Einführungstextes zur Mechanik zusammenhing.

Insgesamt führte das Promotionsprojekt zur erfolgreichen Entwicklung des Messinstruments für Interesse an Teilchenphysik (*Instrument to measure Particle Physics Interest*, IPPI). Es wurde gezeigt, dass die meisten Schüler*innen einem einzigen Interessentyp an Physik zugeordnet werden können und dass das Interesse dieser Schüler*innen mit der Konzeptualisierung des Interesses an Physik in Form einer Hierarchie von Interessensstufen (*hierarchy of levels of interest in physics*, HOLIP) beschrieben werden kann. Außerdem wurde gezeigt, dass die Interessentypen am besten mit Modellen charakterisiert werden können, die das physikbezogene Selbstkonzept zusätzlich zum Geschlecht beinhalten. Die Ergebnisse dieses Promotionsprojekts haben Implikationen sowohl für den Physikunterricht als auch für die physikdidaktische Forschung. Pädagogen, wie beispielsweise Lehrer*innen, können die Konzeptualisierung des Interesses (HOLIP) als Tool nutzen, um Lernaktivitäten zu entwickeln, die für Schüler*innen mit unterschiedlichen Graden von Interesse und physikbezogenem Selbstkonzept interessant sind. Für die physikdidaktische Forschung schlägt die erste Studie eine neue, schrittweise und reproduzierbare Methode vor, um Items aus einem anfänglichen Item-Pool mithilfe einer Rasch-Analyse auszuwählen. Die Hauptstudie schlägt vor, eine Mixed Rasch-Analyse zu verwenden, um qualitative

Unterschiede in einem gemessenen Konstrukt, wie beispielsweise Interesse und Selbstkonzept, zwischen verschiedenen Gruppen von Schüler*innen aufzudecken, sowie Differential Item Functioning und das Antwortverhalten der Schüler*innen beim Interpretieren der Ergebnisse einer Mixed Rasch-Analyse zu berücksichtigen.

TABLE OF CONTENTS

1. INTRODUCTION

Physics became an interesting subject for me from grade 7 on. I liked questioning, describing, and understanding natural phenomena and daily life objects. I found physics as a subject to be different than the others which added to its interestingness for me. Yet, according to past empirical research my experience is rather the exception than the rule as most studies agree that physics as a subject is less interesting for students than most others (e.g. Galton, 2009; Häußler, Lehrke et al., 1998; Osborne et al., 2003). When it comes to physics as a scientific domain, past empirical studies show that the students' interest is low compared to other sciences such as chemistry and biology (Häußler, Lehrke et al., 1998; OECD, 2016). Moreover, students' interest in physics decreases over time despite various efforts invested by educators in making learning activities more interesting (Galton, 2009; Häußler, Lehrke, et al., 1998; Osborne et al., 2003). In general, students' interest in physics differs across a) *content*, for example, pumps; b) *contexts*, for example, biological; c) *tasks*, for example, hands-on activities; and d) *learning environments*, for example, a science centre (e.g. Blankenburg et al., 2016; Dierks et al., 2016; Häußler, Lehrke, et al., 1998; OECD, 2007a, 2016; Sjøberg & Schreiner, 2012). The context of a learning activity is more important than its content, task, or learning environment when fostering interest in physics (Häußler, Lehrke, et al., 1998; Sjøberg & Schreiner, 2012). In addition to examining the four different aspects of physics listed above, education researchers have examined student characteristics that correlate with their interest in physics, such as age, gender, achievement, and physics-related self-concept (e.g. Cheung, 2017; Häußler, Hoffmann, et al., 1998; Häußler, Lehrke, et al., 1998; Kalender et al., 2019a; Lavonen et al., 2021; Nuutila et al., 2020; OECD, 2007a, 2016; Sjøberg & Schreiner, 2012).

In sum, students' interest in physics has been the focus of many education research projects. Moreover, fostering interest in physics is a key component in national and international physics education standards (National Research Council, 2013; OECD, 2017). There are several reasons for the key role of students' interest for physics education. One reason related to educational practice is that interest enhances persistence and achievement while engaging with an object (de Barba et al., 2016; Kauertz & Fischer, 2006; Nuutila et al., 2020). One reason related to society as a political entity is that today's society needs scientifically literate citizens to be able to make educated and critical decisions in life considering, for example, ecological and social problems related to climate change (Osborne & Collins, 2001). Students' interest in science has been included in definitions and discussions about scientific literacy (Bybee & McCrae, 2011; OECD, 2006). One reason related to society as an economic entity is that an increasing number of jobs related to physics is facing a decreasing number of students aspiring a career in physics (Voss et al., 2019; Bøe et al., 2011; OECD, 2008). Students' interest in physics has been argued and found to be crucial for their course and career choices (e.g. Ainley & Ainley, 2011; Bandura, 1986; Fouad & Smith, 1996; Fouad et al., 2002; Lent et al., 1994, 1999; Maltese & Tai, 2011; Tyson et al., 2007).

Hence, there is a strong educational, political and economic interest in fostering students' interest in physics. Many of the physics education research projects focus on the lack of females who are interested in physics, related courses and careers. Previous research suggests that students' career aspirations are influenced by role models and that women may benefit from female role models (Dasgupta & Asgari, 2004; Young et al., 2013). Across Europe there is an unequal distribution of women and men across different domains (Cross and Bagilhole, 2002; Eurofound and European Commission Joint Research Centre, 2021; European Commission, 2000).[1]

Focusing on the concentration of women and men in (stereo)typically different domains, one might overlook that there are some women and men who are working in an un(stereo)typical domain. Traditionally, sex or gender is considered as a decisive characteristic when conducting analyses in physics education research projects. It is rather easy to assess sex or gender since it requires only one item. Analyses, such as differential item functioning, based on sex or gender are well established. This focus on sex neglects other student characteristics (e.g. physics-related self-concept) that might be better suited at describing students' traits, such as their interest in physics or related course and career choices. Moreover, a focus on sex differences in students' traits might suggest and foster the common misconception of fundamental neurological differences between the sexes (Rippon, 2019). Considering the women and men who are interested in un(stereo)typical domains, related courses and careers, the question arises whether it is not the sex or gender itself that is the most decisive characteristic of a person. Based on a review of 46 meta-analyses, Hyde (2005) proposed the 'Gender similarities hypothesis' stating that women and men mainly are similar from a psychological perspective. The 'Gender similarities hypothesis' questions gender as a relevant variable in social sciences. Hence, Lindqvist et al. (2021) suggest an operationalisation of gender (also comprising sex) and recommend that researchers carefully reflect which aspects of gender are important to their research question. Concerning students' course and career choices, past empirical research indeed points towards other student characteristics, such as self-concept, self-efficacy or perceived recognition, which might be more important than sex or other aspects of gender (e.g. Kalender et al., 2019a, 2019b, 2020; Marshman et al., 2018a, 2018b).

This doctoral thesis presents the findings from my research into interesting contents and contexts in physics education while considering different student characteristics. First, I conducted a literature review to get an overview of this research area. My attention got caught by the 'IPN Interessensstudie Physik' conducted by the Leibniz Institute for Science and Mathematics Education (IPN) at the University of Kiel (Häußler, 1987; Häußler et al., 1996; Häußler, Hoffmann et al., 1998; Häußler, Lehrke et al., 1998; Häußler & Hoffmann, 2000; Langeheine et al., 2001; Rost et al., 1999; Sievers, 1999). The IPN study investigated students' interest in physics while considering different aspects of physics (e.g. context) and different student characteristics (e.g. physics-related self-concept). In this study, students' specific interest profiles and characteristics were used to categorise them into different types of interest (Häußler, Lehrke, et al., 1998). Different types means that the students have qualitatively different interest profiles, and within each type, the students have similar interest profiles but different degrees of interest. It found that students who are generally and highly interested in the broad field of physics differed in their interest profiles and characteristics from those who were highly interested in physics when it was set in contexts related to humans and nature, applications, and society (Sievers, 1999). However, it was not described how interesting different contexts are relative to each other within the students' different types of interest. The interest types were described focusing on the students' sex although the results of the IPN study show that the students' physics-related self-concept describes the distribution of students across the different interest types better. In later publications of the IPN sudy that were based on a re-analysis of the data the description of these three interest types

---

[1] For example, there also is a lack of males interested in the (stereo)typically female domains, such as education (Martino, 2008). The effect of male role models in education has been discussed, for example, by Martino (2008).

was different than the originally published ones. Moreover, although more recent research projects (e.g. Drechsel et al., 2011; Levrini et al., 2017) have also distinguished everyday life contexts from others, such as purely scientific contexts, the findings of the IPN study have not been refined or verified in a follow-up study. Past empirical research on interest in physics has not covered modern physics content areas, such as particle physics. Modern physics content areas are already included in international physics curricula, such as the International Baccalaureate Physics curriculum, and in several national curricula, such as the Austrian, Italian, and Norwegian (Mullis et al, 2016; Austrian federal law consolidated, 2022). Thus, analysing students' interest in these content areas set in different contexts is of crucial educational significance. Hence, the main motivation for my doctoral research project was to investigate types of interest of today's students while comparing a classical and a modern physics content area and considering different student characteristics. My doctoral research project aims to provide evidence for answering three important main questions, namely (1) into which different types of interest in physics students can be categorised while additionally considering particle physics as a modern physics content area, (2) how interesting are different contexts within these interest types, and (3) are the interest types described better focussing on physics-related self-concept as a student characteristic compared to sex.

In this thesis, I first describe the theoretical and empirical background (**Chapters 2** and **3**) and the research interest (**Chapter 4**) of my doctoral research project. Then, the two main parts follow, each representing one of the two studies conducted in this doctoral research project.

In the first study, an instrument to measure students' particle physics interest (IPPI) was successfully developed, tested, and validated using a Rasch analysis. Based on the item hierarchy revealed by the Rasch analysis a conceptualisation of students' interest in particle physics as a hierarchy of levels of interest was introduced. This study, its methods, and findings have already been published in an article, which is in parts presented in this thesis (**Chapter 5**; Zoechling et al., 2022).

Second, the main study, a cross-cohort study with German-speaking students aged 14 to 16 years (*N* = 1219), was conducted. The students' interest in mechanics and particle physics was assessed using the instrument to measure mechanics interest from the IPN study and the IPPI. In addition, different student characteristics, such as their physics-related self-concept, sex, and previous experience with the content areas in school, were assessed. The main aim of this study was to investigate the students' types of interest in physics and their association with different student characteristics. For mechanics, the influence of the introductory text was also analysed by randomly assigning the students to two different versions. This study, its methods, findings, and potential implications for both future research and educational practice are presented in the second main part of this thesis (**Chapter 6**).

Finally, **Chapter 7** summarises the main the findings of this doctoral research project.

2. THEORETICAL BACKGROUND

Since fostering students' interest is a key component of education, interest is a widely used construct in education research. Accordingly, there are different approaches to modelling the construct 'interest' theoretically. In this chapter, different theoretical models of the psychological construct 'interest' are presented and distinguished from similar constructs. Then, different aspects of a physics learning activity, such as content or task, that may be the object of students' interest are introduced. Moreover, selected student characteristics that may be associated with students' interest are presented. Finally, the development of students' interest is summarised.

2.1. Structure of the construct 'interest'

2.1.1. Distinguishing interest from similar constructs

As interest is a widely used construct in education research, the term 'interest' is not used consistently throughout all research projects. For example, some researchers, such as Schreiner and Sjøberg (2004), use the terms 'interest' and 'attitude' synonymously. However, in the Programme for International Student Assessment (PISA) 2006 study interest was put under the headline 'attitudes' together with support for scientific enquiry and responsibility (OECD, 2006). Radišić et al. (2020) also consider interest to be an attitude together with enjoyment. Yet, very little is known about the relation between the two motivational constructs interest and attitude. Gardner and Tamir (1989) refer to interest 'as a highly specific type of attitude'. Moreover, they state that interest and attitude might develop differently (Gardner & Tamir, 1989). For example, even though one's attitude towards an object (e.g. nuclear power plant as a content) is negative, one can have a strong interest in learning more about it (Gardner & Tamir, 1989). Krapp and Prenzel (2011) make a more elaborate distinction, which is also used in this thesis: Attitude means a 'general, non-personal evaluation' of an object, whereas interest refers to the 'subjective value attached to the knowledge about this object' (Krapp & Prenzel, 2011).

Second, there is no clear distinction between the constructs 'interest' and 'curiosity' although some researchers, such as Luce and His (2014) try to distinguish between them. Yet, both constructs are described in terms of long-term dispositions and short-term states. Luce and Hsi (2014) state that '[c]onsidering curiosity as a personality characteristic or trait is less useful for science education'. In addition, interest is part of the 'Integrative interest-deprivation wanting-liking model' of curiosity (Litman, 2005). In this model, curiosity is associated with discrete emotional-motivational states reflecting different combinations of different levels of wanting and liking (Litman, 2005). The state reflecting low level of liking and low level of wanting is referred to as 'ambivalent disinterest or boredom', whereas the state reflecting high level of liking and low level of wanting is referred to as 'curiosity as a feeling of "interest"' (Litman, 2005). Similarly, Silvia (2008) refers to interest as 'the curious emotion' (*interest as emotion:* see below). Kashdan and Silvia (2009) argue that they

> *'use ''curiosity'' and ''interest'' as synonyms: both refer to a positive motivational-emotional state associated with exploration. In everyday speech, people tend to use ''curious'' for upcoming events and ''interested'' for current events, but this doesn't reflect a conceptual difference.'*

Jirout and Klahr (2012) define curiosity as 'the threshold of desired uncertainty in the environment which leads to exploratory behavior'. Luce and Hsi (2014) distinguish between six types of curiosity, namely mechanistic, teleological, inconsistency, cause and effect, engineering and medicine, and general knowledge. Further, they argue that 'understanding more about patterns of curiosity *expression* across a variety of settings, topics, and contexts is potentially more productive for furthering our understanding of interest in science' (Luce & Hsi, 2014). Hence, one could argue that Luce and Hsi (2014) use the term 'curiosity' to refer to interest in different content, contexts, and tasks as objects of interest; that is, they refer to interest in *certain* contents, contexts, and tasks as curiosity. Thus, in this thesis the term 'curious' is not used but only the term 'interested', and curiosity is seen as referring to a subset of interest directed only towards certain contents, contexts, and tasks.

Third, some researchers consider interest to be an emotion due to its strong affective component. For example, some researchers term it together with confusion, surprise and awe the 'knowledge emotions' or 'epistemic emotions' (Sander, 2013; Kashdan & Silvia, 2009; Silvia, 2008). In line with Hidi and Renninger (2006), Sander (2013) argues that interest is always directed towards something. According to his definition this also holds true for emotions (e.g. one can only be happy or angry about something; Sander, 2013). Whereas the term 'interest' refers to the emotion that the individual experiences, the term 'interests' refers to the motives for idiosyncratically engaging with an object of interest (Silvia, 2001). However, in an earlier publication Hidi (2006) suggests that considering interest to be an emotion is only appropriate, when referring to the first phase of interest development, 'triggered situational interest' (see **Section 2.4.2.**). Although, there is an emotional component in all phases of interest development, the further interest develops, the more importance gains the cognitive component of interest (Hidi, 2006). Since interest involves not only affective but also cognitive components, Hidi (2006) calls it a 'unique motivational variable'. Also in this thesis, interest is considered to comprise both affective and cognitive components.

Moreover, some researchers compare the constructs 'interest' and 'enjoyment'. For example, it is argued that interest fosters exploration, investigation, and information seeking, whereas enjoyment is directed towards what is achieved and involves satisfaction and pleasure (Silvia, 2008; Kashdan & Silvia, 2009; Ainley & Hidi, 2014). Interest and enjoyment ideally occur together, even though they do not necessarily do so (Ainley & Hidi, 2014). Hence, Renninger and Hidi (2011) argue that measuring enjoyment is not representative for interest. Also in this thesis, 'interest' and 'enjoyment' are considered to be different constructs.

Fifth, some researchers use the terms 'interest' and 'popularity' interchangeably. For example, Muckenfuß (1995) discusses ambiguities concerning the adjective 'interesting'. He points to its manifold usage in everyday life, which is not in line with the theoretical construct 'interest'. Thus, students may indicate *potential* rather than *actual* interest when asked about their interest (Muckenfuß, 1995). Moreover, although students may refer to an object as interesting, they may still not be willing to engage with it (Muckenfuß, 1995). This holds true for the first phase of interest development (see **Section 2.4.2.**). Moreover, Muckenfuß (1995) criticizes that some researchers, such as Häußler, Lehrke et al. (1998), do not distinguish between popularity ('Beliebtheit') and interest because of using the adjectives 'popular' ('beliebt') and 'interesting' ('interessant') synonymously. He further argues that the term 'popularity' comprises the intrinsic character of interest more than the term 'interestingness' ('Interessantheit'; Muckenfuß, 1995). Thus, he hypothesizes that the term 'popularity' reflects the theoretical construct ‘interest’ better than ‘interestingness’ (Muckenfuß, 1995). This hypothesis has not been investigated yet. Furthermore, in the ‘Expectancy-value theory of achievement motivation’ interest is conceptualised as how much a person *likes* an object of interest (Eccles, 2009; Wigfield & Eccles, 2000). However, Renniger and Hidi (2011) criticise that, when measuring liking, the focus is on positive emotions. This may not be suitable to measure interest since it lacks a measure of the negative emotions that may also be associated with interest, especially in the first phase of interest development (see **Section 2.4.2.**).

Finally, some researchers distinguish interest from the psychological constructs ‘non-interest’ and ‘indifference’. According to the ‘theory of interest and non-interest’ by Vogt (2007), non-interest is the counterpart of interest. Both interest and non-interest evolve when engaging with an object (Vogt, 2007). Non-interest can be classified into neutral disinterest and strong aversion (Blankenburg & Scheersoi, 2018). However, indifference is a neutral predisposition of a person towards an object; that is, one shows indifference towards an object before engaging with it (Vogt, 2007).

2.1.2. The person-object-theory of interest

In this doctoral research project, I use Krapp’s (2002a) ‘Person-object-theory of interest’ to model the structure of the psychological construct ‘interest’, which is a well-established theory in the education research community. In this theory, interest refers to the relationship between a person and an object (Krapp, 2002a). Interest cannot exist in a person unless there is an object of interest; that is, one’s interest is always directed towards something (Krapp & Prenzel, 2011). This characteristic of interest is referred to as content specificity, where the term content means the object of interest. Interest is multifaceted, since it involves cognitive-epistemic, emotional, and value-related components (Krapp & Prenzel, 2011). The cognitive-epistemic component comprises the desire to understand better, learn more, or know more about an object. The emotional component refers to the positive (in early stages of interest development also negative) emotions associated with an object (Krapp & Prenzel, 2011). The value-related component considers the significance of an object for the person who is, therefore, willing to put effort in it (Blankenburg & Scheersoi, 2018). The three components result in an intrinsic character of interest-based activities (Krapp & Prenzel, 2011). Thus, when engaging with an object out of interest, the person may experience flow, that is, complete absorption in the present moment (Krapp & Prenzel, 2011; Nakamura & Csikszentmihalyi, 2009).

2.2. Object of interest

The term ‘object of interest’ in the ‘Person-object-theory of interest’ (see **Section 2.1.2.**) refers to a ‘certain part of the cognitively represented environment’ (e.g. a physical object, an abstract idea, a topic, or an activity; Krapp & Prenzel, 2011). When investigating high-school students’ interest, several objects of interest may be interesting for education researchers, such as general interest in science, or interest in a domain, a subject, a content, a context, a task, or a learning environment. When focussing on physics as a domain or subject, education researchers commonly investigate the four objects (1) ***content***, (2) ***context***, (3) ***task***, and (4) ***learning environment***. I decided to focus on these four objects because they are the defining components of a learning activity. Hence, (1) knowing which contents are more (or less) interesting helps educators to design interesting learning activities; (2) if a content was found to be rather less interesting it can be set in an interesting context for the learning activity to be interesting; (3) similarly, choosing an interesting task may increase the overall interestingness of a learning activity; and (4) the interestingness of a task may vary depending on the learning environment. Below definitions of these four objects are provided. Research findings regarding the four objects are presented in **Chapter 3**.

2.2.1. Content

The term ‘content’ is derived from the Latin verb ‘continere’ which means ‘to hold together, to enclose’. In physics education research, the term ‘content’ often refers to the facts, principles and concepts that students have to learn in school. Specific contents can be grouped into content areas. For example, the contents ‘buoyancy’ and ‘velocity’ can be grouped into the content area ‘mechanics’ and the contents ‘current’ and ‘permanent magnets’ into the content area ‘electromagnetism’.

2.2.2. Context

The term ‘context’ is derived from the Latin verb ‘contexere’ which means ‘to weave together, to connect’. In education research, several authors have attempted to define the term ‘context’, but a general definition is still lacking. For example, according to Parchmann and Kuhn (2018) the term ‘context’ is used in two ways, namely related to the content and related to the learning environment. In this doctoral research project, the term ‘context’ refers to the storyline of a learning activity (Mestre, 2002), that is, the situations and circumstances in which or the motives for which the respective content is meaningful (Häußler, Lehrke et al., 1998). The context of a learning activity is considered the combination of a focal event and its corresponding fields of action (Duranti & Goodwin, 1992; Gilbert, 2006; Habig et al., 2018). This is in line with OECD’s PISA 2006 study, where contexts are defined as ‘life situations that involve science and technology’ (OECD, 2006). For example, the content ‘X-ray’ can be set in a variety of contexts, such as medicine (X-ray images), health (food irradiation for conservation), or technology (airport security). However, presenting a content in no additional storyline other than the content itself may also be considered as a context, namely a purely scientific context.

2.2.3. Task

Exemplar tasks in a classroom are reading, calculation, presentation to the peer-group, or conducting an experiment. Dierks et al. (2014) introduced a new method to categorise tasks in education research, the ‘RIASEC+N-model’. It is based on the ‘RIASEC-model’ which was originally developed for categorising vocational interests by Holland (1997). Following the adaption for education research, tasks are sorted into seven categories as listed in **Table 1**.

Table 1. RIASEC+N tasks.

Task category	Description	Example
Realistic	practical task	perform an experiment
Investigative	intellectual task	plan an experiment
Artistic	intuitional or creative task	draw a draft
Social	informing, help, or training task	support peers
Enterprising	leading or influencing task	lead a work group
Conventional	imitative task	search information
Networking	cooperative task	peer-to-peer knowledge exchange

2.2.4. Learning environment

In education research, there is no clear definition of the term 'learning environment'. In this doctoral research thesis, the term 'learning environment' refers to the local environment, its corresponding equipment, habits, and processes. This definition of learning environment combines aspects of a situation and an idioculture as described by Finkelstein (2005). Whereas situations are characterised by the local environment, the equipment, and participants and are one-time and non-reoccurring, the features of a certain learning environment are relatively stable over time and include habits and processes, which resembles an idioculture. Typical learning environments are schools, science centres, museums, and home. Dierks et al. (2014), who introduced the 'RIASEC+N-model', sorted learning environments into the three categories 'school', 'out-of-school leisure-time', and 'enrichment'. However, in a paper published two years later, Dierks' group focused on the three learning environments 'school', 'enrichment', and '(prospective) vocation' (Dierks et al., 2016).

2.3. Selected characteristics of a person associated with interest in physics

In the 'Person-object-theory of interest' (see **Section 2.1.2.**) a person is characterised by her or his interest (e.g. in the physics content 'buoyancy'; Krapp, 2002). However, a person's interest may be associated with other characteristics of a person. In this section, I present three student characteristics that are commonly investigated in association with students' interest in physics, namely sex or gender, physics-related self-concept, and previous experience with a content area. I decided to focus on these variables because (1) the vast majority of studies focuses on students' sex or gender while analysing students' interest; (2) some studies found that the gender effect on interest is mediated by other variables, such as the students' physics-related self-concept; (3) self-concept involves an ability-related component and an identity-related component, and thus is broader than similar constructs, such as self-efficacy, which is only ability-related and refers to specific physics contents; (4) some studies suggest that self-concept is strongly associated with desirable outcomes of physics instruction, such as students course and career choices; and (5) there are only few studies investigating the association of interest in and previous experience with a content area. Below definitions of the three characteristics (1) ***sex or gender***, (2) ***physics-related self-concept***, and (3) ***previous experience with a content area*** are provided. Research findings regarding these four objects are presented in **Chapter 3**.

2.3.1. Sex and gender

Commonly, the term 'sex' refers to the physiological aspects, whereas the term 'gender' refers to the social and behavioural aspects of a person's identity (Lindqvist et al., 2021; West & Zimmerman, 1987). Lindqvist et al. (2021) propose an operationalisation of gender consisting of the four main aspects '(a) physiological/bodily aspects (sex); (b) gender identity or self-defined gender; (c) legal gender; and (d) social gender in terms of norm-related behaviours and gender expressions'. The terms 'sex' and 'gender' may be confused and used interchangeably in English. Similarly, they may be translated to the same word in other languages, such as German ('Geschlecht'). However, the English term 'gender' may be used in other languages, such as German or Polish, to explicitly refer to the social and behavioural identity. In this thesis, sex is considered to be one component of gender, in line with the operationalisation of gender by Lindqvist et al. (2021).

2.3.2. Physics-related self-concept

Research on self-concept roots in writings of the late 19th century. In 1890, the *self* was described by James using a basic dichotomy of the self as a subject and the self as an object (James, 1890). In this notion, the self as a subject, the ‘I’, is the active component of the self, making immediate experiences and perceiving the self as an object, the ‘me’. This notion of the ‘self as object’ is the basis for the now commonly used definition of self-concept (Lohaus & Vierhaus, 2019). Broadly speaking, self-concept is one’s perception of oneself (Shavelson et al., 1976). When focusing on a certain domain or subject, such as physics, self-concept refers to one’s perception of oneself regarding this domain or subject (Marsh, 1990; OECD, 2007a). James distinguished between the material, spiritual, and social self, that is, between the knowledge of one's own physical body, one's own psychic characteristics, and one’s perception of how others view oneself (Lohaus & Vierhaus, 2019). What one believes others think about oneself was a central aspect of the construction of one’s self-concept according to Cooley (1902), who introduced the term ‘looking-glass-self’ for this. One may think about it as different mirrors; that is, the images that one has about oneself depend on the mirror that one looks at (e.g. relevant others as mirrors).

A similar distinction between different facets of self-concept, such as self-perception and perception by others, can be found in the multifaceted and hierarchical construct of self-concept by Shavelson et al. (1976) and Marsh (1990). On top of the hierarchical structure is the general self-concept. Then, mathematical academic, verbal academic, and non-academic self-concept are distinguished. The non-academic self-concept can be divided into social, emotional, and physical self-concept, which is in turn divided into further sub-categories (Shavelson et al., 1976). Mathematical and verbal academic self-concept both affect the subordinate self-concepts in different domains or subjects, such as English or history (Marsh, 1990). A further subdivision distinguishes, for example, physics class in general and specific situations in physics class.

Similar to the construct ‘interest’, the construct ‘self-concept’ comprises cognitive and affective aspects (Bong & Clark, 1999). Like in James’ and Cooley’s models of self-concept (Lohaus & Vierhaus, 2019; Cooley, 1902), Shavelson et al. (1976) consider self-concept to be evaluative. For example, self-concept includes feelings about one’s own competencies (self-evaluation) and is strongly influenced by social evaluation (Bong & Clark, 1999; Marsh, 1990). In line with this twofold definition of self-concept, Hazari et al. (2010; who introduced the physics identity framework described below) argue that ‘[h]ow others see a student is vitally important to how the student sees her/himself’. Yet, the students’ perception acts as a mediator between the perceptions and expectations of others and the students’ self-perceptions and expectations. For example, Godwin et al. (2016; who applied the physics identity framework in engineering education research) wrote that ‘[s]tudents’ perceptions of how others view them are vitally important to how students see themselves’. In this line students’ physics-related self-concept was investigated in the IPN study; that is, students’ ability self-concept and their perceived recognition in physics by the teacher and by the classmates were considered (Häußler, Hofmann et al. 1998). In the PISA 2006 study in Germany, the perceived recognition by the teacher was also investigated (Frey et al., 2009). Carlone and Johnson (2007) introduced the science identity framework, which considers both (a) self-perceptions related to the own competency and performance and (b) perceptions related to the recognition by others. Hazari et al. (2010) adapted and applied this framework to physics specifically and added students’ interest as an identity component. Building on this, Kalender, et al. (2019a, 2019b, 2020) consider four components of internalised identities in their studies, namely the students' interests, competency beliefs, perceived recognition as a physics person (external identity), and physics identity (internal identity, i.e. the self-perception as a physics person). Here, the internal and external identity comprise aspects of belonging and aptitude.

In general, some researchers (see Oyersman et al., 2012) use the terms 'self-concept' and 'identity' interchangeably or describe them in a similar way (e.g. Erikson's description of identity, 1968). Others make a distinction between self-concept and identity (e.g. Baumeister's description of identity as 'definitions that are created for and superimposed on the self' vs. of self-concept as 'totality of inferences that a person has made about himself or herself', 1998). In this thesis, I use the term 'physics-related self-concept' (unlike the physics identity studies) since the 'looking-glass self' theory suggests that what students think others think of them is still something that the students think themselves. Hence, these beliefs are internalised, or to say it with Baumeister (1998), they are also 'inferences that a person has made about himself or herself' (e.g. 'I am a person that is not seen as a physics person by the teacher').

A construct often defined similar to self-concept is self-efficacy as discussed, for example, by Bong and Clark (1999). Self-efficacy refers to students' 'judgements of their capabilities to organise and execute courses of action required to attain designated types of performances', that is, their own perception of themselves regarding specific tasks (Bandura, 1986). 'Self-efficacy thus appears to provide one ingredient […] to the cognitive dimension of self-concept' (Bong & Clark, 1999). A similar distinction is made in the PISA 2006 study: 'In contrast to self-efficacy in science, which asks students about their level of confidence in tackling specific scientific tasks, self-concept measures the general level of belief that students have in their academic abilities'. (OECD, 2007a). The main question is how specific one defines the construct 'self-efficacy' to be able to distinguish it from the similar construct 'ability self-concept'. In this sense, using a very broad definition of self-efficacy, one might even consider self-efficacy and ability self-concept as equivalent constructs. However, in the PISA studies 'science-related ability self-concept' ('Naturwissenschaftsbezogenenes Fähigkeitsselbstkonzept'; Frey et al., 2009) is investigated. This phrasing points towards one very important characteristic of physics-related self-concept, namely that there is more to self-concept than beliefs about the own ability. In addition to the self-perception of ability, it also covers aspects of social evaluation (comparison with others) and the looking-glass self (belief that others see oneself as apt for or good at physics).

In this thesis, physics-related self-concept refers to a rather broad and general perception of oneself in terms of ability in and aptitude for physics including aspects of self-evaluation, social evaluation (i.e. in comparison to others), and the looking-glass self (i.e. beliefs that others see oneself as good at or apt for physics).

2.3.3. Previous experience with a content area

In this thesis, previous experience means the quantitative experience with a content area in contrast to the quality of these experiences. Yet, the theoretical literature mostly focuses on the association between qualitative experiences that the students may have made throughout their educational paths (e.g. with their parents, with their teachers, in the classroom, or in out-of-school learning environments; i.e. not the quantitative experience with the content area) and interest in physics or related careers. Moreover, the focus often is on the association between interest and knowledge or achievement in a content area instead of quantitative previous experience with it. Here, the theoretical literature suggests a positive association between knowledge and interest. In the ‘Four-phase model of interest development’ (Hidi & Renninger, 2006; **Section 2.4.2.**) the cognitive-epistemic component of interest becomes more relevant the further interest develops. That is, the students want to learn more, even when this is associated with difficulties. Moreover, Höft et al. (2019) explain that both the ‘Expectancy-value theory of achievement motivation’ (Eccles, 2009; Wigfield & Eccles, 2000) and the ‘Social cognitive career theory’ (Sheu et al., 2010; Lent et al., 1994) suggest that

> *‘interest is associated with an increased likelihood to engage in more challenging tasks as well as a higher persistence in performing a given task, resulting in higher achievement which in turn is related to interest development, so that reciprocal relations between interest and achievement can be concluded.’*

One may argue that investigating students’ knowledge regarding a content area also provides some evidence concerning the previous experience in school because experience can be regarded as inevitably preceding knowledge. This hypothesis has not been tested yet.

2.4. Development of interest

In the previous sections, the ‘Person-object-theory of interest’ as well as (a) different objects of interest that are the defining components of a physics learning activity and (b) different characteristics of persons that may be associated with their interest in physics were introduced. Below I will summarise the development of students’ interest.

While describing the development of students’ interest, two different forms of interest must be considered according to Krapp and Prenzel (2011). First, ***individual interest***, often referred to as ‘habitual’ or ‘dispositional interest’, describes a relatively stable personality trait. Second, ***operating interest*** refers to the psychological state of being interested while engaging with an object and is associated with ‘focused attention, increased cognitive functioning, persistence, and affective involvement’ (Krapp & Prenzel, 2011). Operating interest can be caused by an existing individual interest. In this case, operating interest is also referred to as ‘actualized individual interest’ (Krapp, 2002a). Operating interest can also be caused by external factors, that is, the interestingness of an object. In this case, operating interest refers to ‘situational interest’ (Krapp & Prenzel, 2011). One aim of education is not only to trigger students’ situational interest by external factors but also to enable them to develop individual interest.

2.4.1. Linking the development of interest to psychological needs

Some researchers, such as Krapp (2002b, 2005), link the development of interest to the basic psychological needs in the ‘Self-determination theory’ (Ryan & Deci, 2000). For example, Krapp (2002b, 2005) suggests that interest develops when these needs (i.e. autonomy, competence, and social relatedness) are met. However, in this doctoral thesis, this link between the development of interest and the basic psychological needs is not investigated.

2.4.2. The four-phase model of interest development

In this doctoral research project, I use the ‘Four-phase-model of interest development’, which was introduced by Hidi and Renninger (2006) and is well established in the education research community. This model explains how situational interest towards an object develops into individual interest, as illustrated in **Figure 1**. In this model, the four phases of interest development are distinguished by the three components of interest (i.e. cognitive-epistemic, emotional, and value-related component; Blankenburg & Scheersoi, 2018). According to Hidi and Renninger (2006) the relation between and the balance of the three components changes across the four phases. Initially, emotional components prevail but they become less important from phase to phase, wherein the value-related and cognitive-epistemic components become more important (Hidi & Renninger, 2006). However, individual interest cannot develop, unless the person associates cognitive value and positive feelings with the object (Blankenburg & Scheersoi, 2018).

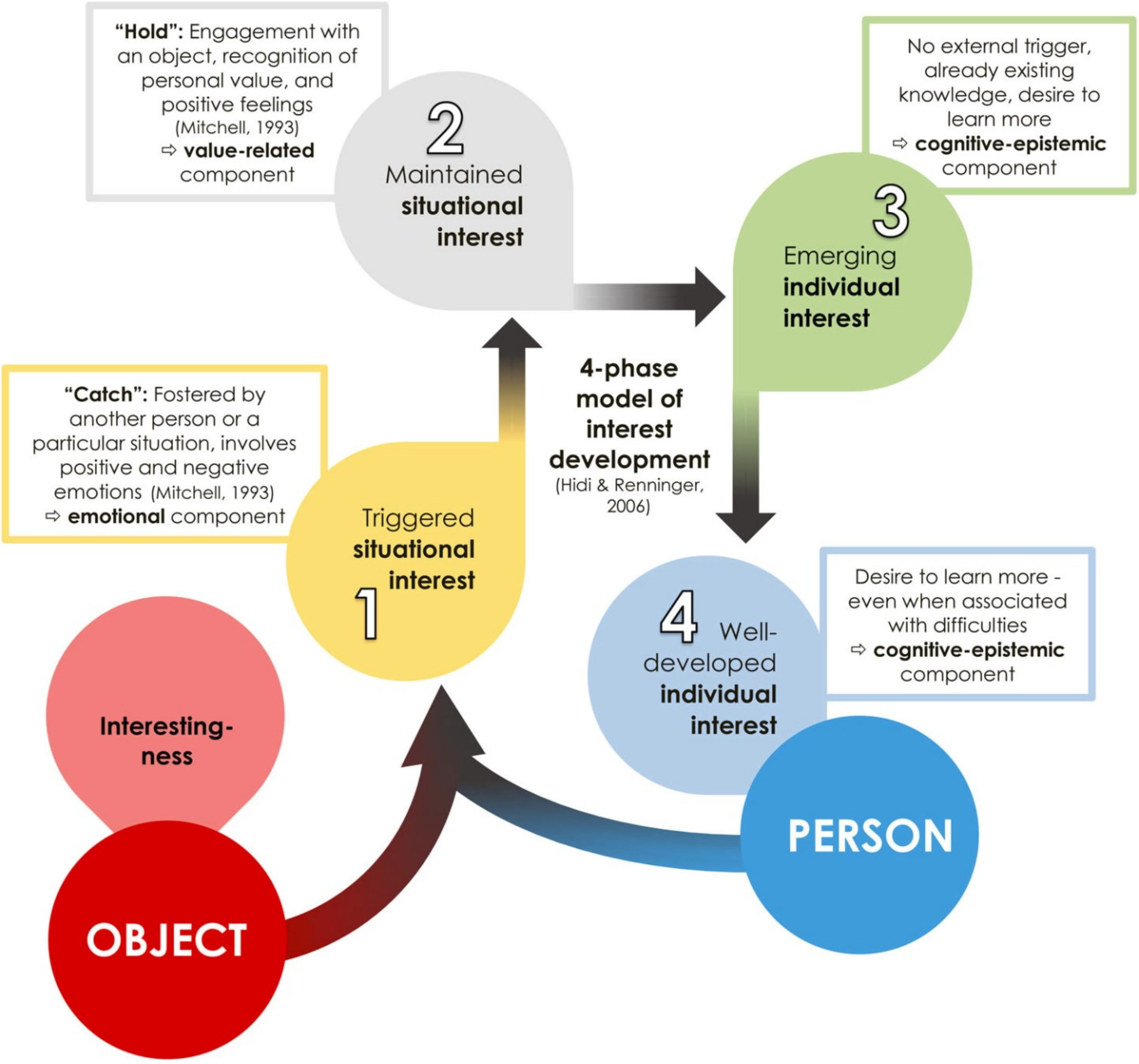


Figure 1. The situational interest is triggered by the interestingness of an object and eventually develops into individual interest in the framework of the 'Person-object theory of interest' (Krapp, 2002a) and the 'Four-phase model of interest development' (Hidi & Renninger, 2006; figure reproduced from Zoechling et al., 2022)

1) The initial phase of interest development is ***triggered situational interest***. Usually, another person (e.g. a teacher or classmate) or a particular situation (e.g. demonstration of an experiment) trigger the arousal of interest towards an object (Blankenburg & Scheersoi, 2018). Mitchell (1993) called them ‘catch factors’. However, this phase can involve positive as well as negative feelings (Blankenburg & Scheersoi, 2018).
2) The second phase of interest development is ***maintained situational interest*** (Hidi & Renninger, 2006). Here, positive feelings are a precondition (Blankenburg & Scheersoi, 2018). The person starts to engage actively with the object and to recognize personal value (Mitchell, 1993). According to Mitchell (1993) situational interest is maintained by ‘hold factors’, such as the perceived personal value.
3) ***Emerging individual interest***, the third phase of interest development, may follow (Hidi & Renninger, 2006). Here, recognition of personal relevance and already existing knowledge about the object are preconditions (Blankenburg & Scheersoi, 2018). This phase involves engagement with the object, even without external trigger. Moreover, the person wants to know more about the object (Blankenburg & Scheersoi, 2018).
4) The final phase of interest development is ***well-developed individual interest*** (Hidi & Renninger, 2006). In this phase, a person engages with the object of interest even when associated with difficulties (Blankenburg & Scheersoi, 2018). Moreover, she or he wants to exchange knowledge about the object with others (Blankenburg & Scheersoi, 2018).

3. EMPIRICAL BACKGROUND

In the previous chapter, the theoretical background of this doctoral research project was presented, especially focussing on the ‘Person-object theory of interest’ (**Section 2.1.2**.) and the ‘Four-phase model of interest development’ (**Section 2.4.2.**). In this chapter, the empirical background is presented. First, commonly applied methods to measure students’ interest are summarised. Then, the topological model of interest is discussed, that is, a multidimensional model of interest comprising the four objects ‘content’, ‘context’, ‘task’, and ‘learning environment’ (introduced in **Section 2.2.**). Moreover, the student characteristics ‘sex or gender’, ‘physics-related self-concept’, and ‘previous experience with a content area’ (introduced in **Section 2.3.**) and their association with students’ interest in physics are discussed. Finally, past empirical studies about students’ interest in physics are presented focusing on the four objects as well as the different student characteristics that have been shown to be associated with interest.

3.1. Measuring students’ interest in physics

To measure certain aspects of the interest construct, usually open questions or rating scale items with students indicating their perceived interestingness of an item are applied (Krapp & Prenzel, 2011). This can happen without any prior stimulus, with a written introduction to the object of interest, or during or after an actual learning sequence (Krapp & Prenzel, 2011). For example, items on interest in science without any prior stimulus are used in the PISA studies in 2006 (8 items) and 2015 (14 items), and in the Relevance of Science Education (ROSE) study (108 items; Frey et al., 2009; Mang & Kristina, 2019; Schreiner & Sjøberg, 2004). In comparison, the IPN study about interest in physics comprises an introductory text on each of the eight assessed physics content areas and 11 corresponding interest items on each content area (Häußler, Lehrke, et al., 1998). Furthermore, the PISA 2006 study also includes embedded items, that is, some test units were accompanied by items to assess the students’ interest in the science content and context of the respective unit (OECD, 2009). A similar approach has also been used by Rösler et al. (2018) in biology education research. Yet, in most studies interest is not assessed while the students are doing an actual task. Moreover, Krapp and Prenzel (2011) also describe some alternative approaches for examining adolescents’ interest in science, such as content analysis of the databanks of an internet search engine. Finally, Renninger and Hidi (2011) as well as Ainley and Hidi (2014) discuss instruments to measure the different phases of interest. For example, they argue that instruments could measure the two phases of situational interest by focussing on the affective component of interest, whereas they could measure the two phases of individual interest by focussing on the cognitive component of interest. In general, Renninger and Hidi (2011) point out that ‘there are no established or agreed-upon methods for measuring interest and its development’.

When measuring traits, such as interest, the measurement instrument as a whole has an influence on the responses to every single item because respondents rate based on the information present in their minds in the very moment (Schwarz, 1999). This information may be accessible all the time resulting in a stability in judgement. It may only temporarily be present caused by contextual effects, such as by an introductory text or the previous and following items (Schwarz, 1999). Respondents assume that the information provided in the questionnaire is relevant, and thus they base their judgement primarily on it (Schwarz, 1999). The order of items can result in assimilation or contrast effects

(Schwarz, 1999). Assimilation is more likely to occur when respondents are asked to judge something that is superordinate to a previous item, whereas contrast is more likely to occur when the previous item is lateral (Schwarz, 1999). For example, when students are asked to express their interest in particle physics, assimilation is more likely to occur when the previous item was about its medical applications. However, if the previous item was about human biology, the contrast effect is more likely to occur. Hence, when using an introductory text prior to a set of items, it is crucial to evaluate its influence on the respondents' expressed degree of the measured trait. For example, different versions of the introductory text could be used and possible differences in responses could be investigated. Moreover, when conducting a replication study using the same measurement instrument, the order of the items could be kept the same to result in the same item order effects. Yet, if the aim is to investigate these item order effects, random order of items could be preferred.

3.2. A topological model of interest

Interest is not equally high in every single aspect of a scientific domain or a school subject. Previous studies focused on students' interest in four different objects that define a learning activity: ***content*** (e.g. pump), ***context*** (e.g. biological), ***tasks*** (e.g. conducting an experiment), and ***learning environments*** (e.g. Science Centre; as introduced in **Section 2.2.**). Krapp and Prenzel (2011) call such differentiations a 'multidimensional topological model' of interest.

In learning activities, the content is usually presented set in a context (i.e. a storyline). Moreover, in empirical research projects, students' interest in a physics content is often measured set in a context. Hence, one may argue that contents are related to contexts. Moreover, contents cannot only be grouped according to a scientific domain or the curriculum of a school subject but also according to a context. This approach is used, for example, in cross-curricular teaching. One content may be set in various contexts and one context may act as a storyline for various contents. In some cases, it is difficult to distinguish between content and context. Häußler, Bünder, et al. (1998) argue that, when comparing the natural science domains, it is more difficult to make such a distinction in biology than in chemistry or physics. For example, when talking about human biology, the context 'human body' is inherent. In chemistry education research, a research group compared the situational interest for the same content set into different context, and for different contents set in the same context (Fechner, 2009; Fechner et al., 2015; Kölbach, 2011). In general, they show that the effects of a context on interest is more distinct, if it can be distinguished clearly from the content (Fechner et al., 2015). Moreover, they found that the positive effect of a context on interest is higher, the less interesting the content itself is (Fechner et al., 2015). However, van Vorst et al. (2014) state that in a meaningful and coherent task content and context are related. In biology education research, Rösler et al. (2018) found that a certain context is more interesting than the corresponding content; for example, the context 'health' is more interesting than the content 'human biology' and the context 'environment' is more interesting than the content 'ecology'. Concerning the interestingness of different contents and contexts, most studies agree that a few contents, such as 'the Universe and its history' (Holstermann & Bögeholz, 2007; OECD, 2016), and contexts, such as 'the possibility of life outside earth' (Sjøberg & Schreiner, 2012) and 'the human body' (Häußler, Lehrke, et al., 1998; Holstermann & Bögeholz, 2007), are extremely and equally interesting for all students. Further research findings concerning the interestingness of different contents and contexts will be presented separately for different studies about students' interest in physics (**Section 3.4.**).

When designing learning activities, the range of possible tasks may depend on the learning environment. Moreover, in empirical research projects, students' interest in a task can be measured set in a specific learning environment. Hence, one may argue that tasks and learning environments are connected. Again, one task may be performed in various learning environments and one learning environment may involve various tasks. Studies considering the students' interest in the RIASEC+N tasks in different learning environments have been conducted, for example, by Dierks et al. (2014, 2016). In general, students' interest in tasks is significantly different in relation to different environments (Dierks et al., 2016). In especial, gender differences were observed: Girls show a higher interest in social tasks in school and in artistic tasks in school and enrichment (Dierks et al., 2016). When focusing on realistic activities, boys indicate a higher interest in vocation, whereas it is the girls in school and enrichment. Dierks et al. (2016) argue that this may be because realistic vocations are perceived as male. Focusing on the learning environment 'school', most studies agree that middle school students show an equal and high interest in hands-on tasks (Häußler, Lehrke et al., 1998; Swarat et al. 2012). Further research findings concerning the interestingness of different tasks and learning environments will be presented separately for different studies about students' interest in physics (**Section 3.4.**).

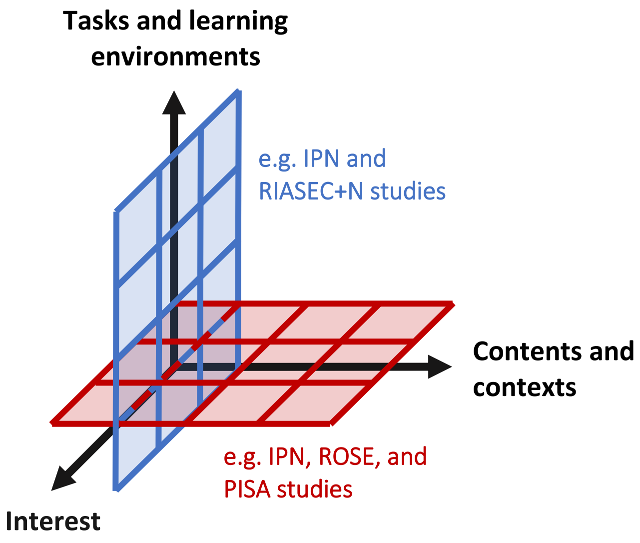


Figure 2. Different contents set in different contexts versus different tasks set in different learning environments

Grouping contents and contexts as well as tasks and learning environments, allows representing interest in form of a coordinate system as shown in **Figure 2**. The horizontal axis represents the contents and contexts, and the vertical axis represents the tasks and learning environments. The most important studies that focus on the horizontal plane are the ROSE and PISA studies (OECD, 2007a, 2016; Schreiner & Sjøberg, 2004). The RIASEC+N studies focus on the vertical axis (Dierks et al., 2014). The IPN study considers both axes (Häußler, Lehrke et al., 1998).

3.3. Selected student characteristics associated with interest in physics

In this section empirical research methodologies and findings regarding the association of students' interest in physics with three student characteristics are discussed: ***sex or gender***, ***physics-related self-concept***, and ***previous experience with the content area*** (as defined in **Section 2.3.**).

3.3.1. Sex and gender

Most previous studies, such as the IPN, PISA, or ROSE studies, analysed the association of students' interest in physics with their sex or gender (e.g. Häußler, Lehrke, et al., 1998; OECD, 2007a, 2016; Sjøberg & Schreiner, 2010). In this thesis, sex is considered to be one component of gender, in line with the operationalisation of gender by Lindqvist et al. (2021; see **Section 2.3.1**.). In many studies, it is not clear whether sex or gender is assessed with a certain item. Moreover, neither sex nor gender are as dichotomous as the items commonly used to assess them (Lindqvist et al., 2021). For example, in the PISA studies students are asked, 'Are you female or male?' (OECD, 2005). When analysing students' interest and its association with sex or gender, three different foci of analysis are commonly applied.

First, male and female students are compared regarding their degrees of interest (e.g. Archer et al., 2013; Häußler et al., 1996; Prenzel et al., 2007). Second, gender differences in interest in physics as a school subject and as a scientific domain are analysed. Most studies agree that the gender difference in favour of boys is less pronounced when assessing domain interest in comparison to subject interest (e.g. Stern et al., 2009; Häußler, Lehrke et al., 1998). Third, different contents and contexts and their interestingness for female in comparison to male students are analysed. Most studies distinguish between contents and contexts interesting for female and those interesting for male students (e.g. OECD, 2016; Sjøberg & Schreiner, 2012). Nevertheless, the studies also show that there are a few contents and contexts that are extremely and equally interesting for all students, for example, 'the Universe and its history' (Holstermann & Bögeholz, 2007; OECD, 2016), 'the possibility of life outside earth' (Sjøberg & Schreiner, 2012), and 'the human body' (Häußler, Lehrke, et al., 1998; Holstermann & Bögeholz, 2007). Further research findings concerning the association of students' interest with their sex or gender will be presented separately for different studies about students' interest in physics (**Section 3.4.**).

3.3.2. Physics-related self-concept

Past empirical studies have provided evidence that how others see a student is important to how the student sees her- or himself. For example, it has been shown that a child's self-perceptions and expectations regarding their science abilities is influenced by their parents' perceptions and expectations of the child's ability (Bleeker & Jacobs, 2001; Jacobs & Eccles, 1992, 2000; Turner et al., 2004). Moreover, the friends' and teachers' perceptions and expectations have been shown to be influential (Jacobs et al., 1998; Speering & Rennie, 1996; Woolnough et al., 1997). Based on such empirical findings and theoretical considerations, physics-related self-concept was defined as a rather broad and general perception of oneself in terms of ability in and aptitude for physics including aspects of self-evaluation, social evaluation (i.e. in comparison to others), and the looking-glass self (i.e. beliefs that others see oneself as good at or apt for physics) in **Section 2.3.2.** of this thesis.

In general, some researchers define and use the constructs 'identity' and 'self-efficacy' similarly to self-concept, as discussed in **Section 2.3.2.**. For example, in the physics identity framework competency beliefs, self-perception (i.e. internal identity; item: 'I see myself as a physics person'), and perceived recognition (i.e. beliefs about how other see oneself, external identity; exemplar item: 'My parents see me as a physics person') are combined (Kalender et al., 2019a, 2019b, 2020). Although not labelled as items to measure perceived recognition, also the item set used to measure physics-related self-concept in the IPN study comprised one item each about the perceived recognition by the teacher and by classmates ('My teacher assesses my performance as …' and 'My classmates think I am … in physics'; students had to complete the statements using the 5-category scale ranging from *very good*, *good*, *medium*, *bad*, to *very bad*; Häußler, Lehrke et al., 1998). One may argue that the physics identity framework as such suggests a holistic description of physics-related self-concept as used in this thesis. Moreover, the competency belief items used in the identity studies resemble well-established ability self-concept items used in the PISA 2006 study (Frey et al., 2009; OECD, 2005; Kalender et al., 2019a, 2019b, 2020). Exemplar competency belief items are ‘I understand concepts I have studied in physics’, ‘If I study, I will do well on a physics test’, and ‘If I encounter a setback in a physics exam, I can overcome it’. However, the researchers argue that they are assessing self-efficacy with items used to assess competency beliefs in the science identity studies in the earlier publications by Marshman et al. (2018a, 2018b) or in the publication about a study on test anxiety by Malespina and Singh (2022). Yet, the wording of the competency belief items is broader and more general than of the PISA self-efficacy items. The similarities or differences become obvious when comparing the items used to assess competency beliefs (physics identity studies; Kalender et al., 2019a, 2019b, 2020), ability self-concept and self-efficacy (PISA 2006 study; OECD, 2005). The self-concept items developed by Marsh (1990)[2] were adapted for and applied in the PISA 2006 study (Frey et al., 2009). In the PISA 2006 study the items used to assess science ability self-concept were, for example, ‘I learn <school science> topics quickly’, ‘When I am being taught <school science>, I can understand the concepts very well’, and ‘I can usually give good answers to <test questions> on <school science> topics’.[3] In contrast, the items used to assess science self-efficacy in the PISA 2006 study refer to very specific tasks. That is, the students are asked ‘How easy do you think it would be for you to perform the following tasks on your own?’ and some exemplar items are ‘Explain why earthquakes occur more frequently in some areas than in others’, ‘Interpret the scientific information provided on the labelling of food items’, and ‘Discuss how new evidence can lead you to change your understanding about the possibility of life on Mars’ (OECD, 2005). One can see that self-concept items are referring to the domain or subject in general unlike self-efficacy items which are focusing on very specific tasks and contents. Although the competency belief items (see examples above) also focus on certain tasks (e.g. overcome a setback in an exam), these tasks are somewhat less specific than the tasks assessed in the PISA self-efficacy items (e.g. scientific interpretation of food labels). In addition, the competency belief items refer to physics as a domain or subject, in contrast to the specific contents referred to in the PISA self-efficacy items. Nevertheless, the differences between the items to assess ability self-concept (PISA 2006 study; OECD, 2005), competency beliefs (physics identity studies; Kalender et al., 2019a, 2019b, 2020), and self-efficacy (PISA 2006 study; OECD, 2005) seem to be small. All items refer to the students’ self-perception of ability, some rather general and broad and others rather specific.

Several previous studies, such as the IPN and the PISA 2006 studies, explored the correlation between physics-related self-concept and interest in physics (Häußler, Lehrke, et al., 1998; OECD, 2007a). When analysing students’ interest and its association with physics-related self-concept as defined in this thesis, two different foci of analysis are commonly applied in empirical studies.

First, the nature of the association of self-concept and interest is investigated. Most studies agree that the students’ ability self-concept is positively associated with interest in physics (Buccheri et al., 2011; DeWitt and Archer, 2015; Dierks et al., 2016; Godwin et al., 2016; Häußler, Hoffmann et al., 1998; Kalender et al., 2019a, 2019b; Stern et al., 2009). Moreover, it has been shown that interest in

---

[2] ‘Compared to others my age I am good at [a specific school subject]’, ‘I get good marks in [a specific school subject]’, ‘Work in [a specific school subject] classes is easy for me’, ‘I'm hopeless when it comes to [a specific school subject]’ (reverse scored), ‘I learn things quickly in [a specific school subject]’, and ‘I have always done well in [a specific school subject]’ (Marsh, 1990).

[3] In the German version of the questionnaire ‘<school science>’ is expressed as ‘Naturwissenschaften’ and ‘<test questions>’ as ‘Prüfungsfragen’ (Frey et al., 2009).

physics is positively associated with self-perception as a physics person and perceived recognition as a physics person (Godwin et al., 2016; Kalender et al., 2019a, 2019b). Second, the structure of self-concept, that is, the relationship among the different aspects of self-concept, are investigated. Some studies point out that the students' ability self-concept and self-perception as a physics person can be impacted by the perceived recognition by others (e.g. Aschbacher et al., 2010; Carlone & Johnson, 2007; Zeldin & Pajares, 2000). Further research findings concerning the association of students' interest with their physics-related self-concept will be presented separately for different studies about students' interest in physics (**Section 3.4.**).

3.3.3. Previous experience with the content area

In this thesis, previous experience refers to the quantitative experience with a content area as defined in **Section 2.3.3.**. There are only few studies that explore the association of interest in a physics content area and the students' previous experience with the content area. The empirical literature mostly focuses on knowledge in a content area instead of quantitative previous experience with it. Hypothesising that previous experience with a content area precedes knowledge, some conclusions about the association between interest and knowledge in a content area may be applied to the association between interest and quantitative previous experience with a content area. For example, Höft et al. (2019) conducted a longitudinal empirical study to investigate the association between interest in different tasks and achievement. In line with other past studies, they found that, overall, the interest decreases whereas the achievement increases over time. However, they also showed that interest in cognitively activating tasks or in tasks that involve the communication of knowledge (e.g. solving theoretical problems, debating with classmates, or explaining something to classmates) is increasingly associated with achievement, which aligns with theoretical expectations based on the 'Four-phase model of interest development' (Höft et al., 2019).

Moreover, many researchers focus on the association between interest in the broad field of physics or related careers and qualitative experiences that the students may have made throughout their educational paths (e.g. with their parents, with their teachers, in the classroom, or in out-of-school learning environments; i.e. not the quantitative experience with the content area). Here, Maltese et al. (2014) provide a detailed overview of the literature on what, when, and how such experiences contribute to the students' interest in STEM. They summarise that 'by high school, the majority of students know what career they will pursue' (*When?*), that 'formal and informal learning experiences play a role in generating and maintaining interest in STEM' (*What?*), and that 'individuals can become interested and persist in STEM through positive interpersonal relationships with teachers or family members, or they can develop an interest on their own' (*How?*; Maltese et al., 2014). Regarding their own study, Maltese et al. (2014) conclude that

> *'individuals who complete STEM degrees have quite varied histories and that the triggering of their interest can happen across a wide age spectrum. Similarly, in terms of who was involved or what their triggering experience was, the general results support the notion that there are many pathways to STEM and no clear preferential pathway.'*

However, findings concerning the association between interest in a content area and qualitative physics-related experiences (e.g. with their parents, with their teachers, in the classroom, or in out-of-school learning environments) cannot be applied to the association between interest and quantitative previous experience with a content area.

Only few empirical studies indeed investigated the relationship between previous experience with the content area and interest in it. For example, studies about the construct 'novelty' point towards a positive association between previous experience and interest. Orion and Hofstein (1994) in a study about field trips found that preparing the students to reduce their novelty space (e.g. previous knowledge) improves their attitudes (e.g. interest) and achievement after the trip. Based on this, Woithe (2020) in a study about a hands-on learning laboratory on particle physics also investigated three novelty influence factors, namely previous experiences with the content area particle physics, with hands-on experiments, and with out-of-school learning opportunities. She found the students' previous experiences were not significantly associated with their situational interest (Woithe, 2020). Further research findings concerning the association of students' interest in and their previous experience with a content area will be presented separately for different studies about students' interest in physics (**Section 3.4.**)

3.4. Previous studies about students' interest in physics

One of the first studies about students' interest in physics came from a German researcher (Todt, 1978). Not long after, the importance of students' interest for learning was discussed at the first and the second conference on interest at the Leibniz Institute for Science and Mathematics Education (IPN) at the German University of Kiel in 1984 and 1996 (Hoffmann et al., 1998). Throughout the years similar calls for fostering students' interest in physics were repeated. Now, several national and international physics education recommendations already include fostering interest in physics as a key component (e.g. National Research Council, 2013; OECD, 2017). Furthermore, a relatively large number of studies investigating students' interest in domains, such as physics, and its association with learning have been published. In this section, an overview of some of the most important publications in the field of interest research related to physics is given. I also present some chemistry and biology education research projects that I consider to be relevant also for physics education. **Table 2** provides an overview of the presented studies, author teams, and the variables investigated in these studies. To simplify the list of other assessed variables, I am not consistently using the study authors' original naming in **Table 2**. Instead, I used the PISA naming as a reference if possible. I investigated whether the items used to assess a variable resembles items used in the PISA studies. If yes, I used the naming of the PISA studies. For example, the items to assess 'emotional valence' used in a study by van Vorst et al. (2018) resemble items used to assess 'enjoyment of science' in the PISA studies (Frey et al., 2009; Mang et al. 2019). In a study by Habig, van Vorst et al. (2018), the items to assess 'value-related valence' and 'topic-related interest' resemble items used to assess 'personal value of science' in the PISA studies (Frey et al., 2009; Mang et al. 2019), and the items to assess 'activity-related intrinsic motivation' are similar to items used to assess 'enjoyment of science' in the PISA studies (Habig, van Vorst et al., 2018; Frey et al., 2009; Mang et al. 2019).

Table 2. Past empirical studies (using abbreviated names), author teams, and selected investigated variables associated with students' interest. In the right column I list other investigated variables. The coding (bold, inclined, or coloured) of some author teams and investigated variables indicates that the coded author teams only discuss the equally coded variables. | * Several publications by the author team and other researchers

		Interest in						Associated variables				
Short name	Authors	School subject	Scientific domain	Contents	Contexts	Tasks	Learning environments	Sex/Gender	Self-concept	Performance	Self-efficacy	Others
IPN	Häußler, Lehrke et al., 1998*	x	x	x	x	x		x	x	x		Interest induced by physics instruction, leisure interest, vocational interest, attitude towards women in science, attitude towards physics instruction, fascination with nature and technology, instrumental motivation to learn physics, characteristics of family, characteristics of physics instruction
SIQ	Schiefele et al. (1993)		x									
PISA 2006	OECD (2007b; questionnaire and data, EN)	x	x	x	x			x	x	x	x	Leisure activities, sources of science information, preparation for a career in science, information about a career in science, time spent studying science, characteristics of physics instruction, interest of the teacher, perceived recognition by the teacher, *general value of science, personal value of science*, *vocational aspirations*, ***instrumental motivation to learn science, future-oriented motivation to learn science, enjoyment of science***
	Frey et al., 2009 (questionnaire and data, DE)	x	x	x	x			x	x	x	x	
	OECD (2007a)		x							x	x	
	Stern et al. (2009)	x	x	x	x			x	x	x	x	
	Prenzel et al. (2007)		x					x	x	x	x	
	Buccheri et al. (2011)		x	x	x			x	x	x		
	Drechsel, et al. (2011)		x	x	x			x		x		
	Olsen and Lie (2011)		x	x	x							

cont.

Short name	Authors	Interest in						Associated variables				
		School subject	Scientific domain	Contents	Contexts	Tasks	Learning environments	Sex/Gender	Self-concept	Performance	Self-efficacy	Others
PISA 2015	Mang et al., (2019; questionnaire and data, DE)	x	x	x	x			x		x	x	*Leisure activities*, time spent studying science, **characteristics of physics instruction**, support by the teacher, *vocational aspirations,* ***enjoyment of science, instrumental motivation to learn science***
	OECD (2016)		x	x	x			x		x		
	Salchegger et al. (2016)		x					x		x	x	
	Radišić et al. (2021)		x					x		x	x	
Embedded items	Rösler et al. (2018)		x	x	x	x				x		expectations about how successfully a task will be completed, value of the expected result
ROSE	Schreiner and Sjøberg (2004)*		x	x	x			x				Leisure activities, *characteristics of physics instruction*, vocational aspirations, future-oriented motivation to learn science, general value of science, myself as a scientist
	Jenkins and Nelson (2005)		x	x	x			x				
	Holstermann & Bögeholz (2007)		x	x	x			x				
Studies about different context characteristics	van Vorst et al. (2018)		x	x	x							personal value of science, enjoyment of science
	Habig, van Vorst, et al. (2018)	x	x	x	x							
RIASEC+N studies	Blankenburg et al. (2016)		x	x	x	x		x	x			
	Dierks et al. (2016)	x				x	x	x	x	x		
	Höft et al. (2019)					x		x	x	x		
	Höft and Bernholt (2019)					x		x	x	x		

cont.

		Interest in						Associated variables				
Short name	Authors	School subject	Scientific domain	Contents	Contexts	Tasks	Learning environments	Sex/Gender	Self-concept	Performance	Self-efficacy	Others
ASPIRES	DeWitt, et al. (2013)*	x							x			Vocational aspiration, leisure activities, characteristics of parents, characteristics of peers
HOPE	Levrini et al. (2017)	x	x	x	x	x	x	x	x	x		Vocational aspiration, leisure activities, characteristics of physics instruction, information about a career in science
Physics identity	Kalender et al. (2019a)*		x					x	x			Physics identity, perceived recognition

3.4.1. IPN Study

The 'IPN Interessensstudie Physik' is one of the first systematic studies about **school students**' interest in physics (Häußler, 1987; Häußler et al., 1996; Häußler, Hoffmann et al., 1998; Häußler, Lehrke et al., 1998; Häußler & Hoffmann, 2000; Langeheine et al., 2001; Rost et al., 1999; Sievers, 1999). It was conducted by the IPN (Leibniz Institute for Science and Mathematics Education at the University of Kiel) in the 1980s and 90s. The IPN study was cross-sectional (grade 9) and longitudinal (grades 5 to 10). In total, 10 954 interest records were collected whereby students participating in the longitudinal study are counted several times. In the IPN study interest in eight curricular content areas of physics (e.g. mechanics and electronics) was assessed utilising an introductory text and eleven rating scale items per content area. The introductory text on each curricular content area preceded the corresponding part of the questionnaire to cover students' possible deficit in previous experience with the content area by giving them an idea of what it is about. The items tackled three objects of interest, namely different contents, contexts, and tasks. The eleven interest items for each content area were based on eleven item categories as listed in **Table 3**. Each item category represents a certain combination of seven different contexts and four different tasks (Häußler, Lehrke, et al., 1998). The students express their degree of interest in each item on a 5-category rating scale ranging from 'My interest in it is …' *very high*, *high*, *medium*, *low*, or *very low* (Häußler, Lehrke, et al., 1998).

Table 3. The eleven item categories of the IPN study (Häußler et al., 1996; Häußler, Hoffmann et al., 1998; Rost et al., 1999) and one exemplar item for each category for the content area 'heat'.

#	Item category	Context	Task	Items for the content area 'heat'
01	Learning more about the functional principle of technical devices	Understanding of technical devices in everyday life	Receiving information (observing, reading, listening)	Learning more about how thermos jugs function
02	Learning more about natural phenomena	Enrichment of emotional experiences		Learning more about what causes the weather
03	Learning more about the relevance of physics for society	Relevance for society		Learning more about how improved insulation of houses may save a lot of energy
04	Learning more about qualitative physics	Science I (qualitative)		Learning more about what heat essentially is
05	Learning more about quantitative physics	Science II (quantitative)		Learning more about why heat may not completely be transformed into movement
06	Getting insight into technical jobs	Vocation I (technical, scientific)		Gaining an insight into how people work in a thermal power station
07	Getting insight into jobs related to humans	Vocation II (medical, artistic, counselling)		Gaining an insight into how people work in a weather station
08	Constructing technical devices	Enrichment of emotional experiences	Hands-on (constructing objects, conducting experiments)	Building and testing a device from simple materials (e.g. chips of wood or straw) that keeps things warm
09	Planning experiments	Science I (qualitative)	Minds-on (devising, calculating)	Planning experiments to find out what influences the speed at which an object cools down
10	Calculating physical quantities	Science II (quantitative)		Calculating how much of a certain amount of heat energy may be transformed into kinetic energy
11	Discussing about the societal relevance of physics	Relevance for society	Evaluation and discussion	Learning about the disadvantages of thermal power stations and discussing alternatives

Additionally, in the IPN study interest in physics as a scientific domain and as a school subject were distinguished. The former was measured by summing up the responses to all 88 items on the curricular content areas. The latter was measured with two additional specific items, namely one single item ('I find the subject physics …') and one item as part of an item set consisting of 14 items on interest in 14 different school subjects including physics. All of these items use a rating scale ranging from *very interesting*, *interesting*, *medium*, *boring*, to *very boring* (Häußler, Lehrke, et al., 1998).

Häußler, Lehrke, et al. (1998) found that the domain interest decreases continuously the older the students get. Male students indicate a significantly higher interest than females. Yet, the gender difference in domain interest is way less pronounced than the gender difference in subject interest (Häußler, Lehrke, et al., 1998). One can see that, whereas male students report equal or higher interest in the subject than in the domain, female students report lower interest in the subject than in the domain. That is, gender differences are more pronounced when students are asked directly about their subject interest than when calculating the mean domain interest across all items (Häußler, Lehrke, et al., 1998).

Moreover, Häußler and Hoffmann (2000) compared students' interest in physics as a school subject and as a scientific domain. They found that, although interest in physics as a subject and as a domain have a relatively high correlation (*correlation coefficient* $r \approx 0.7$), they differ in predictor variables in multivariate regression analyses. Interest in the school subject is associated with positive physics-related self-concept (*standardised regression coefficient* $\beta = 0.43$) which involves current achievement, corresponding expectations, and how competent students feel to be perceived by the peer group (Häußler & Hoffmann, 2000). In addition, interest stimulating instruction by the teacher may contribute to the students' interest in the subject ($\beta \approx 0.2$). Häußler and Hoffmann (2000) point out that interest in the domain fairly contributes to interest in the school subject ($\beta \approx 0.1$). Interest in the scientific domain is associated with fascination by technical objects ($\beta \approx 0.4$) and fascination by natural phenomena ($\beta \approx 0.2$) (Häußler & Hoffmann, 2000).

Häußler, Bünder et al. (1998) highlighted that gender is not associated with interest differences across students. They concluded that gender differences are mediated by other variables (e.g. fascination by technical objects), which can be stereotypical, gender-related characteristics (Häußler, Bünder, et al., 1998).

As mentioned above, the IPN study distinguished between eight content areas aligned with the German physics curriculum and between seven context areas that were identified in a preceding Delphi-study. The eight content area factors resulting from the data analysis are slightly different than the initial eight content areas. In particular, the content areas 'electricity' and 'electronics' got merged and the content area 'acoustics' was split into 'sound production' and 'noise protection' (Häußler, Lehrke, et al., 1998). Furthermore, only two out of the seven context areas remain as factors, namely 'emotional experiences', in especial 'natural phenomena', as well as 'vocation', in especial 'medical devices' (Häußler, Lehrke, et al., 1998). All the other context items were assigned to the activity factors instead as described below (Häußler, Lehrke, et al., 1998).

Häußler, Lehrke, et al. (1998) found that students in general show a stable and high interest in both context areas 'natural phenomena' and 'medical devices' with female students being more interested than males. Moreover, they report a general high interest in the content area 'radioactivity'. Here, the initial gender difference in favour of boys evens out when students get older (Häußler, Lehrke, et al., 1998). The second most interesting content area is 'nuclear physics'[4], for which no notable gender difference was shown. The other content areas are all similarly interesting for the students (Häußler, Lehrke, et al., 1998). However, boys indicate a continuous and pronounced higher interest in 'motion of vehicles' and 'electricity and electronics'. In comparison, the gender difference is smaller concerning 'optical instruments' and 'thermodynamics'. No gender differences were shown concerning 'sound production', nor for 'noise protection' (Häußler, Lehrke, et al., 1998).

---

[4] Häußler, Lehrke, et al. (1998) call this content area 'How the world is built on a small scale'. The introductory text is about protons, neutrons, and electrons, as well as lasers.

Four kind of tasks were distinguished in the IPN study, namely receiving information (e.g. reading, listening), hands-on task (e.g. conducting experiments), minds-on task (e.g. calculation), and evaluation and discussion (Häußler, Lehrke et al., 1998). As a result of the data analysis only three of the four tasks remained as factors since all the items on the task 'receiving information' were assigned to the content area factors instead. Häußler, Lehrke et al. (1998) found that students show an equal and high interest in hands-on tasks, although it decreases when they get older. Moreover, they report a general high interest in the task 'evaluation and discussion'. Here, the initial gender difference in favour of boys evens out when students get older. However, boys indicate a higher interest in the task 'calculation', even though it is the least interesting task for all students (Häußler, Lehrke et al., 1998).

The learning environment is not specified because, as Häußler (1987) argues, the interest in the tasks is regardless of the learning environment. However, the instruction text preceding the questionnaire starts with 'Dear students, in physics class …' (‚Liebe Schülerinnen, liebe Schüler, Im Physikunterricht …'; Häußler, Lehrke et al., 1998). Thus, their focus was on school as learning environment.

In addition, the students' interest profiles were used to categorise the students into different types of interest in physics (Häußler, Lehrke et al., 1998). The interest profiles resulted from a mixed Rasch analysis. In general, three types of interest in physics were distinguished. In different publications of the IPN study, the interest types are presented in different versions (Häußler, 1987; Häußler et al., 1996; Häußler, Hoffmann et al., 1998; Häußler, Lehrke et al., 1998; Langeheine et al., 2001; Rost et al., 1999; Sievers, 1999). The earlier publications (Häußler, 1987; Häußler et al., 1996; Häußler, Hoffmann et al., 1998; Häußler, Lehrke et al., 1998) present the interest types based on the initial analysis of the data. Here, the hypothesis was that there are two different types. As a result of the analysis the three interest types A, B, and C were distinguished and presented. Sievers (1999) conducted the latest re-analysis of the data in his doctoral research project. In this thesis, I refer to the definitions of the interest types presented in the more recent publications (Langeheine et al., 2001, Rost et al.,1999) that are based on Sievers' thesis (1999). Here, the hypothesis was that there are three different types A, B, and C. As a result of the analysis the three types A, C, and NG were distinguished and presented.

The first type (A) describes students that are generally and highly interested in physics, from mathematical calculations to discussions about the relevance for society. This type consists mainly of boys and/or students with high physics-related self-concept (Sievers, 1999). The second type (C) describes students that are only interests in physics when set in a context that relates to humans and nature, applications, and relevance for society. This type consists mainly of girls and/or students with low physics-related self-concept (Sievers, 1999). Type C was found to be a combination of the initially found types B and C. The initial type B was described as being interested in 'the more practical side of physics', and the initial type C as being interested in 'the social implications of physics' (Häußler, Hoffmann et al., 1998). The third type (NG) consists of students whose interests are either similar to type A or type C depending on the content area (Sievers, 1999). The third type (NG) was discussed to result from the students' different use of the rating scale. That is, they mostly used the extreme categories of the rating scale which results in non-ordered ('nicht geordnet' in German) thresholds between the rating scale categories in the interest profile of these students (Sievers, 1999). Since the third type (NG) is similar to type A or type C depending on the content area, one may argue that only two types of interest in physics were found in the IPN study.

Furthermore, Häußler, Lehrke, et al. (1998) found that 60% of the explainable inter-item-variance can be traced back to the context, and 20% to each, the content and the task. They recommend the following context characteristics for raising students' interest: '(1) providing opportunities to be amazed; (2) linking content to prior experiences for both boys and girls; (3) providing first-hand experiences; (4) encouraging discussions and reflections on the social importance of science; (5) letting physics appear in application-oriented contexts; (6) showing physics in relation to the human body; (7) letting students experience the benefit and use of treating physics quantitatively.' (Häußler, Hoffmann et al., 1998).

3.4.2. SIQ

The Study Interest Questionnaire (SIQ) was developed by Schiefele et al. (1993) to measure **university students**' interest in a scientific domain. In total it consists of 27 items; 11 items assess the emotional component of interest (e.g. 'After a long weekend or holiday, I look forward to my studies'), 9 items assess the value-related component (e.g. 'It is more important to me to engage with my studies' contents than to have leisure time and entertainment'), and 7 items assess the intrinsic character of interest (e.g. 'I chose my studies because of its interesting contents'; Schiefele et al., 1993). The students express their degree of agreement to each item on a 4-category rating scale ranging from *strongly disagree*, *disagree*, *agree*, to *strongly agree*. The SIQ does not involve any introduction to the object of interest (i.e. to the scientific domain). The SIQ was validated in a study with university students aged 22 to 25 years (*N* = 298) from different programmes (e.g. economics, engineering, and education). Schiefele et al. (1993) also assessed other variables, such as intrinsic motivation, using scales developed by other researchers. For example, they showed that interest in one's field of study is significantly and positively associated with intrinsic motivation ($r = 0.46$), achievement ($r = 0.33$), and the learning strategies elaboration ($r = 0.48$) and information seeking ($r = 0.45$; Schiefele et al., 1993).

3.4.3. PISA studies

In the PISA (Programme of International Student Assessment) studies **15-year-old students**' ability to use their reading, mathematics and science knowledge and skills to meet real-life challenges are measured. Since 2000, a PISA study usually takes place every three years with different focus. So far, PISA has involved more than 90 countries and economies and around 3 million students worldwide.

PISA 2006

The PISA 2006 study focused on science. For example, it included eight rating scale items on general interest in science (OECD, 2007a). The students had to rate their interest in learning about the eight different areas of natural science 'physics contents', 'chemistry contents', 'plant biology', 'human biology', 'astronomy contents', 'geology contents', 'how scientists design experiments', and 'what is required for scientific explanations' (Frey et al., 2009). The students were asked: 'How much interest do you have in learning about …?' (OECD, 2005). The rating scale comprised the four categories 'high interest', 'medium interest', 'low interest', and 'no interest' (OECD, 2005). These items focus on the cognitive-epistemic component of interest. In addition, other student characteristics were assessed, such as self-efficacy, ability self-concept, and perceived recognition by the teacher in science. For example, the ability self-concept in science was assessed with six items as listed in **Table 4**. The students had to indicate their agreement to each item on a 4-category rating scale ranging from *strongly agree*, *agree*, *disagree*, to *strongly disagree*.

Table 4. The six items used to assess ability self-concept in the PISA 2006 study (OECD, 2005; Buccheri et al., 2011)

#	Item wording
1	Learning advanced <school science> topics would be easy for me
2	I can usually give good answers to <test questions> on <school science> topics
3	I can learn <school science> topics quickly
4	<School science> topics are easy for me
5	I can easily understand new ideas in <school science>
6	When I am being taught <school science>, I can understand the concepts very well

Note: In the German version of the questionnaire '<school science>' is expressed as 'Naturwissenschaften' and '<test questions>' as 'Prüfungsfragen' (Frey et al., 2009).

In accordance with the SIQ results, the PISA 2006 study suggests that higher general interest in science is associated with higher performance values (OECD, 2007a). For the pooled OECD sample the correlation between interest in science and science performance is significant, yet relatively low ($r = 0.13$; OECD, 2007a). Stern et al. (2009) analysed the Austrian[5] results in detail. They point out that the significant correlation between interest in science and performance ($r = 0.24$) is lower than the significant correlation of science performance with science self-efficacy ($r = 0.44$; Stern et al., 2009). Indeed, also for the pooled OECD sample the significant correlation between science performance and self-efficacy is relatively higher ($r = 0.33$; OECD, 2007a). In addition, Stern et al. (2009) also analysed the significant correlations between interest and other student characteristics, namely science self-efficacy ($r = 0.45$), science ability self-concept ($r = 0.47$), instrumental motivation to learn science ($r = 0.43$), future-oriented motivation to learn science ($r = 0.48$), enjoyment of science ($r = 0.62$), general value of science ($r = 0.43$), and personal value of science ($r = 0.53$).

Concerning the assessed areas of natural science, two thirds of participants indicated interest in learning about 'human biology' on average across OECD countries, whereas about half of students indicated interest in learning about the other areas (OECD, 2007a). A little less interest was shown in learning about 'geology' and 'what is required for scientific explanations' (OECD, 2007a). Across most participating countries no interest differences related to gender were observed for the eight general interest items (OECD, 2007a). For the Austrian participants, Stern et al. (2009) highlight the relatively higher interest in learning about 'human biology' and 'how scientists design experiments' compared to the OECD average. Moreover, they sum up the participants who indicate high or medium interest and point out gender differences. Whereas girls indicate higher interest in learning about 'human biology' (19%), 'plant biology' (18 %), and 'astronomy' (6%), boys indicate higher interest in learning about 'physics' (26%), 'chemistry' (16%), 'how scientists design experiments' (8%), 'geology' (7%), and 'what is required for scientific explanations' (7%; Stern et al., 2009). Stern et al. (2009) argue that from a student's point of view the eight general interest items rather concern the school subjects than the scientific domains. In the German student questionnaire students had to choose their favourite subject out of biology, physics, and chemistry (Frey et al., 2009). Similar to the results of the general interest items, 42% of German participants preferred biology, contrary to 24% who preferred chemistry and 21% who preferred physics (Frey et al., 2009).

---

[5] Stern et al. (2009) discuss that their analysis of the Austrian participants can also be applied to participants from Switzerland, Germany, and Luxembourg.

In addition to the general interest items, the PISA 2006 study consisted of 37 test units. Each unit began with an introductory text and a graph or a picture about the science content set in a context followed by the cognitive test items. 18 of the test units were accompanied by in total 52 embedded items on interest in the science content and context of the respective units (OECD, 2007a). The focus of the embedded interest items in PISA 2006 is on 'learning more', 'knowing more', and 'understanding better', that is, on the cognitive-epistemic component of interest. Stern et al. (2009) analysed the results of the embedded interest items and compared them to the results of the non-embedded interest items for the Austrian participants. They found that when using embedded interest items, 73% of girls and 75% of boys indicate high or medium interest in physics. When using non-embedded interest items, 36% of girls and 62% of boys indicate high or medium interest in physics (Stern et al., 2009). Stern et al. (2009) conclude that the expressed interest in physics is higher and the gender difference is much smaller when using embedded items on interest in physics. They argue that the participants indicate their interest in the school subject when confronted with the general interest items and their interest in specific aspects of a scientific domain when confronted with embedded items (Stern et al., 2009). Stern et al.'s findings (2009) about the difference between domain interest and subject interest align with findings of the IPN study (Häußler, Lehrke et al., 1998).

Prenzel et al. (2007) also analysed the interest expressed in the embedded interest items by the students focusing on the fourth performance quartiles, that is, the highest-performing 25% of students in each country. They compared the data of Australia, Austria, Canada, Germany, Finland, the Netherlands, and Switzerland. They found that in all analysed countries there is a balanced gender distribution in the fourth performance quartile (percentage of girls ranging from 43% to 49%). Prenzel et al. (2007) also divided the participants of each country into four interest quartiles. Then, they analysed the percentage of high-performing students per interest quartile. They found that in the third and fourth interest quartile the high-performers are slightly overrepresented which is true for all analysed countries (e.g. 29% of German high-performers in the fourth and 28% in the third interest quartile). Prenzel et al. (2007) conclude that performance and interest are positively correlated, when using embedded items for measuring interest. Moreover, they analysed if gender is relevant for the expressed interest. They found that in all analysed countries and corresponding interest quartiles high-performing girls are neither over- nor underrepresented (percentage of high-performing girls ranging from 41% to 51%; Prenzel et al., 2007). Finally, Prenzel et al. (2007) investigated which aspects are associated with the interest expressed by German students utilising a regression analysis. They considered gender, social background, type of school, and different instruction characteristics ('Unterrichtsmerkmale') in their regression model. They found that only one aspect significantly contributes to modelling the students' interest, namely the instruction characteristic 'frequency of scientific modelling and applying' ('Häufigkeit des naturwisschenschaftlichen Modellierens und Anwendens'). They point out that gender does not significantly contribute to modelling the students' interest (Prenzel et al., 2007).

Drechsel et al. (2011) also focused on the embedded interest items and analysed the students' expressed interest in the contents and contexts of the cognitive test units. They used a random sample of about 25 500 students from the 30 participating OECD countries representing 10% of the cases of the complete dataset. They found that a 4-dimensional content model that distinguishes between 'earth and space', 'living', 'physical', and 'technology' best describes the sample, followed by a 4-dimensional context model that distinguishes between 'health and environment', 'frontiers', 'hazards', and 'resources'. Yet, for further analyses they used a 2-dimensional content model distinguishing between 'living systems' (22 items) and 'physical and technology systems' (30 items) because this distinction 'resembles the content structure of science as a domain very well, as seen from a student perspective' (Drechsel et al., 2011). They found that boys and girls are equally interested in living systems but there are gender differences concerning physical and technology systems. Boys are more interested in physical and technology systems than in living systems, whereas it is the opposite for girls. The gender difference concerning physical and technology systems is statistically significant (Drechsel et al., 2011). When focusing on students of the highest performance quartile, the gender differences and gender specific preferences are similar (Drechsel et al., 2011). Moreover, Drechsel et al. (2011) pointed out that high-performing boys are equally interested in both systems. Overall, high-performing students are significantly more interested in both systems than other students (Drechsel et al., 2011).

Olsen and Lie (2011) did cluster analysis with regard to countries and items. They distinguished between European and non-European countries, as well as between items on living systems and on physical/technological systems (Olsen & Lie, 2011). The latter distinction is in line with Drechsel et al. (2011). Utilising a cluster analysis, Olsen and Lie (2011) found that within the cluster 'living systems', most items relate to plants and agriculture, and within the cluster 'physical/technological systems', most items relate to fires and pure or frontier science (e.g. evolution or genetics). Moreover, they found that non-European countries are more interested in items on living systems which may be explained by their relation to basic needs for survival. European countries, on the contrary, indicate higher interest in items on physical/technological systems (Olsen and Lie, 2011).

Buccheri et al. (2011) analysed the data of high-performing Finish, Swiss, Korean, and Australian students. In this study 'high-performing' refers to the students that have achieved the two highest levels of performance in science or in mathematics. They found that in Korea, Finland, and Switzerland girls show significantly more interest in human biology than in chemistry, in which they indicate again significantly more interest than in physics. Finish and Swiss girls are overrepresented in medicine as a vocational aspiration (Buccheri et al., 2011). Males in Australia, Finland, and Switzerland 'are equally interested in chemistry and physics and more interested in these than in biology' (Buccheri et al., 2011). In all four countries male students are overrepresented in vocational aspirations such as engineering and physics (Buccheri et al., 2011). Buccheri et al. (2011) also found that in all four countries 'female high-performers have a significantly lower [ability] self-concept in sciences than their male colleagues'. They also found that the gender differences in vocational choices are only partly due to interest and argue based on Eccles et al. (1998) that fostering the students' ability self-concept in science and their perceived utility of a science career might also foster corresponding career aspirations (Buccheri et al., 2011).

In sum, the PISA 2006 study assessed students' interest in physics using one non-embedded interest item ('physics contents'), and additionally it assessed students' interest in physics using several items embedded in the different test units, which may resemble a typical classroom situation better. It confirmed findings from the IPN study that gender differences in interest are less pronounced when using embedded items and that interest is associated with other student characteristics, such as science-related ability self-concept.

PISA 2015

The PISA 2015 study focused again on science. The German student questionnaire assessed students' interest in different school subjects (Mang et al., 2019). The students were asked 'How much interest do you have in the following school subject?' and had to answer on a four-category rating scale ranging from *high interest*, *medium interest*, *low interest* to *no interest*. 29% of the German students reported high or medium interest in biology, compared to 22% in chemistry and 20% in physics (Mang et al., 2019). These results are in line with the findings of the PISA 2006 study (Mang et al., 2019).

Moreover, the PISA 2015 study included 14 general interest items on five broad science topics or subjects (biosphere, motion and forces, energy and its transformation, the universe and its history, how science can help us prevent disease; OECD, 2016). The students were asked 'How much interest do you have in learning about the following science topics?' and had to answer on a four-category rating scale ranging from *high interest*, *medium interest*, *low interest* to *no interest*. In contrast to the PISA 2006 study, there were no interest items embedded in the test units. However, other student characteristics, such as self-efficacy in science, were also assessed.

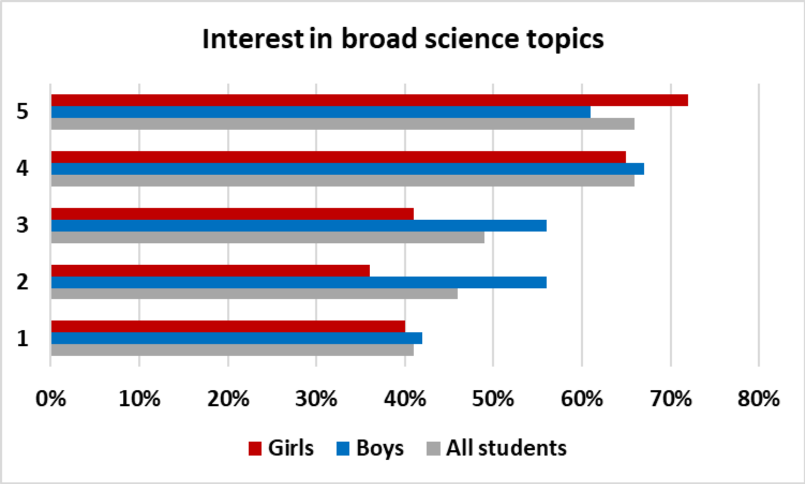


Figure 3. Students' interest in five broad science topics.

1) Biosphere (e.g. ecosystem services, sustainability),
2) Motion and forces (e.g. velocity, fiction, magnetic and gravitational forces),
3) Energy and its transformation (e.g. conservation, chemical reactions),
4) The Universe and its history, and
5) How science can help us prevent disease (OECD, 2016)

On average, 66% of OECD participants indicate high or medium interest in 'how science can help us prevent disease' and 'the universe and its history', compared to 49% being interested in 'energy and its transformation', 46% in 'motion and forces', and 41% in 'biosphere' as shown in **Figure 3** (OECD, 2016). Moreover, there are some striking gender differences. According to OECD (2016) 'boys are more interested than girls in physics and chemistry ("motion and forces", "energy and its transformation"), while girls tend to be more interested in health-related topics ("how science can help us prevent disease")'. A weakness with this argument is that it involves categorising into boys' and girls' content areas. Yet, it is apparent in **Figure 3** that the supposedly girls' content areas are even more interesting for boys than the supposedly boys' content areas.

Salchegger et al. (2016) further analysed the Austrian results in comparison to some OECD countries. They highlight that the average interest of Austrian participants is higher than the OECD average and that there is a pronounced gender difference in Austria with boys being more interested than girls. In addition, in Austria and on OECD average there is a significant positive correlation between performance and interest in science ($r_{AT} = 0.29$ and $r_{OECD} = 0.26$), although this does not apply to all participating countries (Salchegger et al., 2016). In comparison to some OECD countries, Austrian participants report the least self-efficacy and there is a pronounced gender difference with boys having a more positive self-efficacy (Salchegger et al., 2016). Finally, in Austria and on OECD average there is a significant positive correlation between interest and self-efficacy ($r_{AT} = 0.34$ and $r_{OECD} = 0.32$; Salchegger et al., 2016).

Radišić et al. (2021) further analysed the data obtained from students in Italy. Utilising a latent profile analysis, five groups of students were distinguished based on their interest, enjoyment, instrumental motivation, self-efficacy, epistemological beliefs, and activities in science (Radišić et al., 2021). The biggest group (59%) are the 'informal inquirers' who practice their interest in science through leisure activities. The second largest group (24%) are the 'uninterested' students who, on average, enjoy science less and engage less in school or leisure science activities. In contrast, the 'scientists' (12%), on average, enjoy and are interested in school science and in leisure science activities the most and have the highest self-efficacy and instrumental motivation. The 'uncommitted' (3%) have the lowest levels on all considered variables. Finally, the 'practical inquirers' (2%) have a high self-efficacy and instrumental motivation, engage in leisure science activities but have the lowest interest in science topics (Radišić et al., 2021). The 'scientists' have the highest performance mean, followed by the 'informal inquirers', and the 'uninterested' (Radišić et al., 2021).

In sum, the PISA 2015 study only assessed students' interest in broad physics topics using non-embedded items. It confirmed findings from the PISA 2006 study that gender differences in interest are more pronounced regarding topics relating to technological systems (e.g. motion and forces) compared to topics relating to living systems (e.g. biosphere). Moreover, it confirmed findings from the IPN and PISA 2006 studies that interest is associated with other student characteristics, such as science-related self-efficacy.

3.4.4. Embedded items in biology education research

Rösler et al. (2014) examined how interesting and motivating contexts are associated with performance in biology tests, focusing on **grade 10 students** (*N* = 1543). They distinguished between the competencies 'content knowledge' and 'decision-making' (Rösler et al., 2018). The items were set in the four different context areas 'health' (contexts: *nutrition* and *drugs*), 'environment' (contexts: *water quality* and *air quality*), 'technology' (contexts: *environmental technology* and *food technology*), and 'natural resources' (contexts: *soil resources* and *agriculture*). For each context they developed a common stimulus text and four corresponding items for both competencies, as well as competency-specific stimulus texts and four corresponding items, resulting in 16 items per context (Rösler et al., 2014, 2018). Summing up across the eight contexts, their competency test consisted of 128 items. Furthermore, they developed a questionnaire about interest (*individual* and *situational*) and motivation (*expectations about how successfully a task will be completed* and *value of the expected result*; 116 items adapted from past empirical studies; Rösler et al., 2018). Using a multi-matrix-design, all items were distributed across 32 booklets with 21 items each (Rösler et al., 2018). Like in the PISA 2006 study, the questionnaire items were embedded; that is, they were presented after each set of four competency items. Rösler et al. (2018) observed that the context does not affect the average item difficulty. However, they found that students' situational interest differs across the contexts. Here, they confirmed previous findings that health is the most interesting context area for students, followed by environment, whereas students express rather low interest in technology and natural resources. In addition, they found that competence correlates significantly with individual interest ($r = 0.18$), situational interest ($r = 0.28$), expectancy (i.e. expectations about how successfully a task will be completed, $r = 0.40$) and value (i.e. value of the expected result, $r = 0.27$). They describe the influence of interest and motivation on performance utilising a structural equation model. In this model, there is a significant positive association of situational interest with expectancy ($r = 0.38$) and value ($r = 0.63$), and of expectancy with value ($r = 0.57$). Individual interest has no direct influence on competence but a significant direct influence on situational interest ($r = 0.39$), expectancy ($r = 0.27$) and value ($r = 0.33$). The indirect influence of individual interest on competence is mediated by situational interest and the expectancy component of motivation. Situational interest and expectancy have a significant influence on competence ($r = 0.14$ and $r = 0.37$). Individual and situational interest, and expectancy together account for 18% of the variance in competence (Rösler et al., 2018).

3.4.5. ROSE study

The ROSE (Relevance of Science Education) study focused on **15-year-old students**' interest in science and was carried out from 2002 to 2006 (Schreiner & Sjøberg, 2004; Sjøberg & Schreiner, 2010, 2012). The ROSE questionnaire comprised 108 rating scale items asking the participants what they want to learn about. Thus, the focus of the ROSE survey was on the cognitive-epistemic component of interest which is similar to the PISA 2006 study. The survey did neither include a stimulus text, picture, or a graph, nor were the students directly involved with a task. Schreiner and Sjøberg (2004) distinguished between 10 content areas ('subject') and 10 context areas. However, 'some matters can be regarded as a subject *and* as a context' (Schreiner & Sjøberg, 2004). Moreover, Holstermann and Bögeholz (2007) argue that some matters were not considered although they are common part of science education (e.g. structure and function of cells), whereas other topics that are not relevant for science education (e.g. ghosts, witches, and whether they exist) were investigated. The ROSE study found that the general interest in science is lower, the more developed a country is (Sjøberg & Schreiner, 2012). In line with the IPN study, it showed that the context is key for understanding the students' interest. When analysing the data, the focus of the ROSE study was on gender differences. Sjøberg and Schreiner (2012) concluded that boys 'and **NOT** the girls' are interested in 'technical, mechanical, electrical, spectacular, violent, [and] explosive' matters. On the other hand, girls 'and NOT the boys' are interested in '[h]ealth and medicine, beauty and the human body, ethics, aesthetics, wonder, speculation', and 'the paranormal'. As outlined above for findings of the PISA 2015 study, such argumentation may not describe the analysis results well because it suggests that, for example, girls' topics are not interesting for boys.

In contrast, Jenkins and Nelson (2005) analysed the data obtained from a sample of 1277 students from the ROSE study in England. When reporting their results, they point not only to gender differences but also highlight similarities:

> *'Although the distribution of interests among boys and girls is very different, it is important to acknowledge that a high level of interest in a given topic by one gender does not necessarily mean that the same topic is of no interest to the other. For example, although "Black holes, supernovae and other spectacular objects in outer space" is a topic which most boys indicate strongly they would wish to learn about, the topic is also of interest to girls […]. Likewise, boys would also like to learn about "Why we dream when we are sleeping and what the dreams may mean" […]. In contrast, boys are not interested in learning about "Eating disorders" (mean score 2.03) and girls show little enthusiasm for understanding "How the atom bomb functions" (mean score 2.27).'* (Jenkins & Nelson, 2005)

Holstermann and Bögeholz (2007) analysed a sample of 265 German students in detail who had participated in the pilot study of the ROSE study in Germany. Their factor analysis resulted in 13 factors. To enhance the reliability, they excluded nearly half of the items and assigned 57 items to the 13 factors. Their results indicate that girls and boys express equal and high interest in 'bodily harms', 'the universe', and 'research', equal and relatively lower interest in 'natural phenomena', 'animals', and 'humans and the environment', as well as equal but relatively even lower interest in 'agriculture and plants' (Holstermann & Bögeholz, 2007). However, girls show high interest in 'diseases and epidemics', 'bodily functions', and 'body awareness'. Boys indicate interest in these topics as well, although it is less pronounced. The interest difference is the largest for 'body awareness'. Moreover, girls express high interest in 'the paranormal' unlike boys. Boys show high interest in 'physics and technology' and in 'dangerous applications'. Girls indicate interest in these topics as well, although it is less pronounced (Holstermann & Bögeholz, 2007).

3.4.6. Studies about different context characteristics in chemistry education research

van Vorst et al. (2014) conducted a study about context characteristics and their influence on grade 9 students' interest in chemistry education. They set tasks in one of three context areas (nature, leisure time, and traffic). Fechner et al. (2015) state that it is only possible 'to design contexts with a high potential to generate content-related situational interest' if 'detailed characteristics of a context are worked out and substantiated'. A model for operationalising context characteristics was proposed by van Vorst et al. (2014) based on a literature review as well as their own study. They describe three main characteristics, namely authenticity, familiarity, and topicality. Several past empirical studies have investigated the association of students' interest with the familiarity and the topicality of a context as presented in the following.

a) Familiarity: 'Familiarity' refers to the subjective perception of frequency and regularity with which a person encounters a context (van Vorst et al., 2014). van Vorst et al. (2014) distinguish between contexts with relation to everyday life and unique contexts. They propose that a context is strongly related to the everyday life of a person when the frequency of encounters is high, whereas the characteristic 'uniqueness' represents the opposite (van Vorst et al., 2014). The term 'uniqueness' has already been used by Häußler, Hoffmann et al. (1998) for explaining unexpected high interest arousals. For example, some items are particularly interesting in combination with certain content areas (e.g. nuclear power plant shortly after the accident in Chernobyl; Häußler, Hoffmann et al., 1998). However, in their study van Vorst et al. (2018) show that the emotional valence of unique contexts is higher than the emotional valence of daily life contexts. The same holds true for learning with unique contexts in comparison to daily life contexts (van Vorst et al., 2018). Overall, the interest in contexts was rated higher than the interest in learning with these contexts (van Vorst et al., 2018). In addition, Habig et al. (2018) examined the interest in contexts over a longer time-span, in particular when using the same context in three consecutive tasks. They show that the emotional valence decreases in general but it decreases more when using unique contexts in comparison to daily life contexts. The value-related valence also decreases and again the decrease is less pronounced when using daily life contexts (Habig et al., 2018). Habig et al. (2018) also analysed the effect of individual subject and domain interest. Here, they show that students with high individual interest indicate higher interest in unique contexts than in daily life contexts, whereas it is the opposite for students with low individual interest (Habig et al., 2018).

b) Topicality: Furthermore, van Vorst et al. (2014) distinguish between contexts that are up-to-date or not. They argue that, although topicality is subjective, societal trends exist (e.g. due to the seasons of the year or media). Here, van Vorst et al. (2018) show that topicality is not significantly associated with the emotional valence of contexts nor of learning with these contexts. However, they point to limitations of their study with respect to the topicality of their contexts (van Vorst et al., 2018).

3.4.7. RIASEC+N studies

The **RIASEC+N studies** were based on the RIASEC+N-model which was introduced by Dierks et al. (2014; see **Section 2.2.3.**).

Blankenburg et al. (2016) focused on German **grade 6 students** (*N* = 474) and compared tasks regarding three different contents set in a daily-life context. In particular, the physics content 'buoyancy' was set in the context 'floating and sinking of objects like ships', the chemistry content 'combustion' in the context 'candles', and the biology content 'flora and ecosystem' in the context 'plants'. The learning environment is school. Blankenburg et al. (2016) found that in all three content-context-combinations the realistic and investigative tasks are the most interesting, whereas the social and enterprising tasks are the least interesting. Moreover, students who indicate higher interest in a certain context are more interested in all activities in that context. This is also true for tasks (Blankenburg et al., 2016). However, girls indicate a higher interest than boys in artistic tasks in the physics and chemistry content-context-combinations as expected by Blankenburg et al. (2016). Boys, on the contrary, show higher interest than girls in social tasks in the physics content-context-combination. This result was unexpected and Blankenburg et al. (2016) explain it by the higher science-related self-concept of boys.

Dierks et al. (2016) focused on German **grade 8 to 12 students** (*N* = 247) and compared students' interest in different tasks in three different learning environments, namely school, enrichment, and vocation. They found that, in general, social and networking tasks are the most interesting, whereas artistic tasks are the least interesting. Students with high self-concept, and students with high general interest are more interested in all tasks and environments (Dierks et al., 2016). Moreover, high-achievers indicate higher interest in all tasks (Dierks et al., 2016).

Höft et al. (2019) examined the interest in tasks, while focusing on school as learning environment and the three chemistry contents 'chemical reaction', 'energy', and 'matter'. Their focus was on the long-term relationship between conceptual understanding and interest (Höft et al., 2019). They conducted a cross-sectional study with **grade 5 to 11 students** (*N* = 2510). Their questionnaire consisted of a part on conceptual understanding involving 30 classical or ordered multiple choice items (i.e. four grade-specific and six anchor items on each of the three contents). The part of the questionnaire on interest in tasks consisted of the 28 RIASEC+N items (i.e. four items per each RIASEC+N task; Höft et al., 2019). In line with previous findings, Höft et al. (2019) found that students' interest in all tasks in school declines more or less monotonically from grade to grade, whereas conceptual understanding generally increases. However, the interplay between interest in tasks and conceptual understanding differs across tasks (Höft et al., 2019): For investigative, social, networking, and enterprising tasks, it is a small to middle-sized positive association, which increases from grade 7 to 11. Höft et al. (2019) argue that this is due to the specific characteristics of these tasks, which come along with higher attention, persistence, and advanced learning strategies. For realistic tasks (i.e. solely the hands-on part of experiments), they found a small positive, and almost stable association between interest and conceptual understanding. Thus, they argue that without the minds-on part, experiments do not contribute to a higher conceptual understanding. For artistic and conventional tasks, the association is negligible. Höft et al. (2019) argue that this is due to the specific characteristics of these tasks, which are usually not cognitively demanding nor contribute to a feeling of relatedness.

Furthermore, Höft and Bernholt (2019) examined how students' interest in school science tasks develops over **grades 9 to 11** and whether this relation is directed or reciprocal. They used the same questionnaire as Höft et al. (2019). In line with Höft et al. (2019), Höft and Bernholt (2019) found that, whereas the interest in all tasks declines from grade to grade, the conceptual understanding increases. In addition, they observed that conceptual understanding is reciprocally associated with interest in investigative, social and networking tasks. Furthermore, Höft and Bernholt (2019) found that interest in enterprising tasks is positively associated with conceptual understanding. They explain this with the characteristics of these tasks. Investigative and networking tasks usually result in cognitive activation and knowledge gain, and enterprising and social tasks involve top-down-like exchange of knowledge between students (Höft & Bernholt, 2019). They suggest that further analysis is necessary. For example, to understand the effect of enterprising tasks better, focusing on the students' self-concept is recommended (Höft & Bernholt, 2019).

3.4.8. ASPIRES

The ASPIRES project explored the science aspirations and career choices among **10- to 14-year-old students** in England in a five-year longitudinal study (Archer et al., 2013; DeWitt et al., 2013, 2015). A quantitative online survey was administered to 10/11-year-old students in the first phase. The students were tracked and surveyed again at ages 12 and 14 (second and third phase). More than 9000 Year 6 students participated in the first phase.

> *'There was no notable gender difference within the 648 children who were classified as "uninterested in science" (i.e. those who recorded the lowest scores on all the five science aspirations items), but notably fewer girls (N = 92, 37%) than boys (N = 159, 63%) were classified as being "science keen" (N = 251) (i.e. those scoring very highly on all five science aspirations items). That is, of the overall sample, 3.4% of the boys were classified as "science keen" and 2.0% of girls.'* (Archer et al., 2013)

However, most of the students (more than 70%) indicated that they enjoy science, have positive views of scientists, take part in science-related activities in their free time, and that their parents value science. Yet, only few (less than 17%) indicated that they aspire a science career. More than 4500 grade 9 students (aged 14 years) participated in the third phase (out of which only about 1000 had participated in the first phase too). For both grade 6 and grade 9 students, aspirations in science were positively associated with the parents' positive attitudes towards science, positive science-related ability self-concept, and positive experience of school science (DeWitt et al., 2013, 2015).

In line with findings from the ROSE study carried out in England (Jenkins & Nelson, 2005), the ASPIRES study points to the difference between students' positive attitudes towards science or doing science and their rather rare vocational aspirations to be a scientist (Archer et al., 2010; DeWitt et al., 2013). Archer et al. (2010) termed this disparity the 'being-doing divide' (Archer et al., 2010). DeWitt and Archer (2015) suggest that students' identity is crucial for translating attitudes into aspirations.

3.4.9. HOPE

The Horizons in Physics Education (HOPE) project investigated the views of **first-year physics students** (Levrini et al., 2017). In the HOPE questionnaire, students were asked to indicate the importance of different factors for their choice to study physics. The HOPE questionnaire consisted of 20 rating-scale items covering the six aspects ‘personal interest’, ‘job perspectives’, ‘school experience’, ‘out-of-school experience’, ‘targeted recruitment efforts’, and ‘science/physics capital’. The students’ interest was measured using five items (e.g. ‘A wish to acquire a deep understanding of the Universe’, ‘Wanting to understand how things work’, ‘A wish to learn advanced physics’). The questionnaire was inspired by the ASPIRES (described above) and the Colorado Learning Attitudes about Science Survey (CLASS) instruments. The CLASS assesses university students’ beliefs about physics and learning physics (e.g. personal interest in physics; Adams et al., 2005).

In total, 2485 first-year university physics students from 31 partners in 18 different European countries participated in the study. Levrini et al. (2017) found that factors related to individual interests are more important than those related to the school experience (e.g. teacher and school marks) or recruitment efforts (e.g. visits to and from universities). There was a moderate statistically significant gender difference as female students tend to be more motivated by recruitment efforts (Levrini et al., 2017). Second, 94 semi-structured individual interviews (50 male and 44 female students) were carried out in 16 universities on a selection of students who had previously answered the questionnaire. In particular, the main motivational factor ‘personal interest and curiosity’ was unfolded into different categories to gain ‘a more fine-grained picture of the different kinds of interest and curiosity the students have’ (Levrini et al., 2017). Levrini et al. (2017) distinguished between the ‘two macro categories: (A) curiosity to understand the world, natural phenomena and the Universe; and (B) interest in physics knowledge as a *special* way of knowing, investigating, questioning and thinking’. Following Luce and Hsi (2014), category A was further described in terms of the three sub-categories ‘Mechanisms underlying phenomena’ (e.g. asking how something works), ‘“teleological” cluster’ (e.g. asking why processes occur), and ‘inconsistency/surprise/wonder’ (Levrini et al., 2017). Category B was further described in terms of the six sub-categories ‘mindset of physicists, rational thinking and problem solving’, ‘“Think different-and-critical”’, ‘math cluster’, ‘experiment/real world connection’, ‘theoretical modelling’, and ‘never-ending questioning’ (Levrini et al, 2017). Students who did not specify their interest or curiosity in such a clear way were categorised into category AB ‘Generic fascination’. Other categories were the ‘application cluster (C)’, the ‘philosophy cluster (D)’, the ‘societal engagement cluster (E)’, the ‘job cluster (F)’, the ‘personal challenge cluster (J)’, and ‘other (K)’. Most students (80 out of 94) gave answers that were categorised in category A or B or both (Levrini et al., 2017). Nine of the remaining students gave answers categorizable only in category AB and five students gave answers categorizable only into C to K. Gender differences in the two macro categories A and B were not significant which might be due to the small sample size as Levrini et al. (2017) argue.

3.4.10. Physics identity studies

The physics identity studies investigated the views of **university students** enrolled in an introductory calculus-based physics course, which is usually taken by engineering or physical sciences students in their first year (Kalender et al., 2019a, 2019b, 2020). In the physics identity questionnaire, students were asked to indicate their agreement to 14 items on their physics identity, perceived recognition, competency beliefs, and interest (see **Table 5**). The students had to rate their agreement to the identity and perceived recognition items on a 4-category rating scale ranging from *strongly agree*,

agree, *disagree* to *strongly disagree*. The items on physics identity and perceived recognition were inspired by a study by Hazari et al. (2010) who introduced the physics identity framework (as discussed in **Sections 2.3.2.** and **3.3.2.**). The physics identity item refers to the students' internal identity and the physics perceived recognition items to their external identity (Kalender et al., 2019a). The items on physics competency beliefs were inspired by the CLASS and other instruments. The items on physics interest were inspired by the Fascination in Science instrument by Chung et al. (2016).

Table 5. The items used to assess physics identity, physics perceived recognition, physics competency beliefs, and physics interest in the physics identity studies (Kalender et al., 2019a, 2019b).

Assessed variable	Item(s)
Physics identity	I see myself as physics person
Physics perceived recognition	My parents see me as physics person
	My friends see me as physics person
	My TA or Instructor see me as physics person
Physics competency beliefs	I am able to help my classmates with physics in the laboratory or in recitation
	I understand concepts I have studied in physics
	If I wanted to, I could be good at physics research
	If I study, I will do well on a physics test
	If I encounter a setback in a physics exam, I can overcome it
Physics interest	I wonder about how physics works
	In general, I find physics
	I want to know everything I can about physics
	I am curious about recent physics discoveries
	I want to know about the current research that physicists are doing

Utilising a structural equation model, the physics identity studies suggest that the relation between gender and physics identity is mediated by the students' perceived recognition, and additionally by their competency beliefs and interest (Kalender et al., 2019b). The model showed that gender has a significant direct effect on perceived recognition ($\beta = 0.27$) with female students indicating less perceived recognition than males and that students' perceived recognition has a significant direct effect on their physics identity ($\beta = 0.59$), competency beliefs ($\beta = 0.54$), and interest ($\beta = 0.64$). The significant direct effect of gender on interest is comparatively weak ($\beta = 0.16$; Kalender et al., 2019b). Utilising a principal component analysis by gender, they showed that physics identity and interest are strongly aligned and do not factor out separately for either gender (Kalender et al., 2019a). There was one exception, namely for female students the perceived recognition (i.e. external identity) item 'My TA or instructor see me as a physics person' loads strongly with the factor 'self-efficacy or belonging' instead of 'interest or identity' (Kalender et al., 2019a). That is, female students' perceived recognition by their TA or instructor is closely linked to their self-efficacy and sense of belonging in a physics classroom (Kalender et al., 2019a). Here, it is important to remember that sample comprised university students enrolled in an introductory calculus-based physics course, which is usually taken by engineering or physical sciences students in their first year.

Godwin et al. (2016) applied the physics identity framework to engineering education. Utilising a structural equation model, they found that perceived recognition in physics has the largest significant direct effect on physics identity ($\beta = 0.72$). They emphasise that '[s]tudents who feel recogni[s]ed by their peers, family, and teachers are more likely to identify as a […] "physics person"' (Godwin et al., 2016). The significant direct effect of interest on physics identity is comparatively smaller ($\beta = 0.37$). The model also showed that competency beliefs have very large significant direct effect on interest and on perceived recognition ($\beta = 0.91$ and $\beta = 0.81$; Godwin et al., 2016).

4. RESEARCH INTEREST

In the previous chapters, I introduced the theoretical and empirical background (**Chapters 2** and **3**) of my doctoral research project. Both the theoretical and empirical literature show that students' interest in physics is intertwined with a broad range of desired educational outcomes, such as physics-related self-concept, achievement, and course and career choices. However, several open questions about students' interest in physics remain that my doctoral research project aims to address, for example, 'Into which different types of interest in physics students can be categorised while additionally considering particle physics as a modern physics content area?' and 'How interesting are different contexts relative to each other within these interest types?'. Even though we already know a lot about some of the possible instructional approaches based on contexts that foster students' interest, studying them in terms of the students' types of interest and extending them to modern physics content areas, such as particle physics, may be useful for educational practice.

4.1. Limitations of past empirical studies

Although several studies examined students' interest in physics, the generalisability of their results for educational practice is limited. Most past studies assess students' general interest in physics as a domain or school subject instead of specific content areas. In addition, the focus usually is on sex- or gender-related issues of physics education, which has several limitations and may increase the perception of differences between different sexes or genders. Since teachers commonly encounter mixed-sex or -gender classrooms, they require recommendations on how to foster interest of most students (not only of 'the girls' or 'the boys'). Here, describing students' interest types, as done in the IPN study, might be a more generalisable approach and helpful in providing educators with recommendations for the design of interesting learning activities.

Yet, in different publications of the IPN study, the interest types were presented in different versions (see Häußler, 1987; Häußler et al., 1996; Häußler, Hoffmann et al., 1998; Häußler, Lehrke et al., 1998; Langeheine et al., 2001; Rost et al., 1999; Sievers, 1999; as discussed in **Section 3.4.1.**). Moreover, when describing the students' categorisation into the different types of interest, the focus was on their sex or gender and to a lesser extent also on other characteristic, such as their physics-related self-concept. In addition, the IPN study on interest types was conducted more than 30 years ago, which raises questions on whether its results still apply to today's learners. The questionnaire used in the IPN study comprised an introductory text on a physics content area as well as corresponding interest items. Here, it is not clear how the introductory text affected students' expressed interest. Moreover, there were some constructional flaws, that is, some items did not fit the item configuration. For example, concerning the content area 'radioactivity' the item on radioactivity in our surrounding does not match the item category 'technical devices'. Hence, 23 out of 88 items were excluded and only 65 items remained for analysis (Häußler, Hoffmann et al., 1996, 1998). Finally, modern physics content areas, such as particle physics, have not been included in most previous studies about students' interest in physics. Yet, some previous studies, such as the IPN study (Häußler, Lehrke et al. 1998), indicate that students are highly interested in the content areas 'radioactivity' and 'nuclear physics', which are strongly linked to particle physics. Moreover, there were no gender differences in interest regarding nuclear physics and the gender differences in interest regarding radioactivity decreased with increasing age. In sum, the content area ‘particle physics’ may be highly and equally interesting for all students (Häußler, Lehrke et al. 1998). Assessing students’ interest in physics in empirical studies relies on suitable test instruments. A measurement instrument to assess students’ interest in particle physics does not exist yet. The unfamiliar terminology and lack of (daily life) previous experiences make it challenging to assess students’ interest in this content area.

Finally, when investigating the association of the students’ interest with other variables than gender, different studies focus on various constructs. However, when looking at the item wording, the items used to assess one construct may resemble other items used in a different study to assess a different construct (as discussed in **Section 3.4.**). In addition, the same constructs may be defined differently in different studies. For example, Kalender et al. (2019a) used a rather broad definition of self-efficacy in physics. Hence, they referred to physics as a domain (or subject; in contrast to specific physics content areas, such as mechanics, or contents, such as buoyancy) in their items. Consequently, these items to assess self-efficacy in physics resemble items used in other studies to assess physics-related ability self-concept (e.g. in the PISA 2006 study). Moreover, some studies show that different constructs do not factor out separately. For example, Kalender et al. (2019a) found that sense of belonging in physics and self-efficacy in physics were so strongly aligned with one another that they did not factor separately for either gender. The same was found for physics identity and interest. Kalender et al. (2019a) argue that ‘[r]esearchers rarely examine the factor structure of multiple motivational factors together’ and that their ‘findings show that it is important to do so’.

4.2. Motivation for this project

Most past studies agree that the context is key for fostering the students’ interest in physics. The IPN study also showed that the context is more important than the content or the task for the interestingness of an item. Yet, the research approach (i.e. longitudinal and cross-cohort study), analysis method (i.e. mixed Rasch model), and results (i.e. students’ types of interest in physics) of the IPN study were unique for the physics education research field. Because of the inconsistency in results (see **Section 3.4.1.**), I initially planned to re-analyse the original IPN data but, unfortunately, I could not get this data. Hence, my motivation was to take data with one part of the original IPN instrument. I decided to focus on grade 9 students in my study because students in grade 9 typically already had to decide in favour of or against a school path that focuses on natural sciences, such as physics. Investigating students’ interest in physics after such a crucial and conscious decision may be even more informative for the design of interesting learning activities. In addition, other large-scale studies, such as PISA and ROSE, and the cross-cohort part of the IPN study also focused on grade 9 students. This approach also allows for a later comparison of my results to those of these studies.

When deciding on which part of the original IPN instrument to use in my study, I considered the constructional flaws reported by Häußler, Hoffmann et al. (1996, 1998). I could not find any reported constructional flaws for the content area ‘mechanics’. I also analysed the item wordings and compared it with the predefined item categories. My analysis showed that the items are overall in line with the category descriptions. Only the second item ‘pumping oil from great depths’ does not represent the category ‘Learning more about natural phenomena’ well. When deciding on investigating the content area ‘mechanics’ one motivation was also that students in grade 9 are already familiar with this content area since mechanics is usually covered in grade 9 and/or earlier grades.

Since the IPN questionnaire comprises an introductory text on the content area in addition to corresponding interest items, it is not clear whether the introductory text affects students' expressed interest. Hence, I decided to use two different versions of the mechanics introductory text and to investigate whether there are differences in students' interest in mechanics associated with the use of different versions. One version is the original IPN mechanics introductory text. I decided to develop a second new version and to randomly assign the students to one of the two versions.

Moreover, my aim was to compare a classical physics content area 'mechanics', with which the students are familiar, with a rather rarely covered modern physics content area. In particular, I decided on investigating students' interest in particle physics as this content area is strongly linked to the content areas 'nuclear physics' and 'radioactivity', which were found to be the most interesting for students in the IPN study (Häußler, Lehrke et al. 1998). The IPN instrument design comprising an introductory text on each content area in combination with the respective interest items is well suited to compensate for the possible lack of previous experience that students may have with particle physics. Most students will not have had prior instruction on particle physics because it is not yet fully established in school curricula, especially below grade 11. Many students will not have engaged with particle physics outside of school either. Therefore, the IPN study method seems best suited to measure students' interest in particle physics, that is, to present the students with an introductory text on the content area in addition to the corresponding interest items. The introductory text aims to cover students' possible deficit in previous experience with the content area by giving them an idea of what it is about.

Since particle physics was not included in the IPN study, the first step was to develop and validate an instrument to assess the students' interest in particle physics based on the item category descriptions of the IPN study. Combining the already existing measurement instrument about interest in mechanics and a newly developed measurement instrument about interest in particle physics, enables me to investigate whether the results of the IPN study can be replicated with new data from nowadays' students and whether they can even be applied to particle physics.

Finally, my motivation was to shift the focus from sex or gender differences in interest on other student characteristics. In especial, I chose to investigate students' physics-related self-concept and its association with their interest in physics as previous studies indicate that the effect of sex or gender may be mediated via physics-related self-concept. Moreover, my motivation was to investigate the structure of the multiple aspects together that the construct 'physics-related self-concept' comprises, that is, to investigate whether the different aspects vary in their contribution to the physics-related self-concept of different students.

In sum, I decided (1) to base my research on students' interest in physics on the IPN study because of its unique analysis method (i.e. mixed Rasch model) and results (i.e. the students' types of interest in physics), (2) to investigate students' interest in particle physics since this content area may be equally and highly interesting for all students, (3) to investigate the structure of students' physics-related self-concept and its association with their interest because it may act as a mediator for the commonly reported differences in interest between the sexes, and (4) to investigate the association of students' interest with their previous experience with the content areas in school because the previous experience may be very different regarding mechanics and particle physics.

4.3. Research questions and hypotheses

The main research interest of my doctoral research project is to investigate students' types of interest and whether related previous findings about the classical physics content area mechanics can be replicated and extended to the modern physics content area particle physics. In the first study, I developed the Instrument to measure Particle Physics Interest (IPPI) guided by the following research questions:

RQ1: What psychometric evidence can be found to support the use of the IPPI using a Rasch analysis?

RQ2: To what extent do the results of the Rasch analysis of the data collected with the IPPI match the theoretical hierarchy of the students' levels of interest in particle physics?

RQ3: To what extent can the students' levels of interest in particle physics be described qualitatively by associating each level with different contexts?

Second, the main study, a cross-cohort study to investigate students' types of interest, was guided by the following research question:

RQ4: To what extent does the introductory text on the physics content area mechanics affect students' expressed interest, when using the same items?

H4: The students' expressed interest in mechanics differs, when using a different version of the introductory text in combination with the same items.
Background: This hypothesis is based on Schwarz' study (1999) which showed that the information provided in a questionnaire influences the participants' responses.

RQ5: Into which different types of interest in physics can German-speaking students aged 14 to 16 years be categorised, while comparing a classical and a modern physics content area (namely mechanics and particle physics)?

H5.1: The interest types (A, C, NG; Sievers, 1999) are still valid for today's students and for classical as well as modern physics content areas.
Background: Different types means that the students have qualitatively different interest profiles, and within each type, the students have similar interest profiles but different degrees of interest. This hypothesis is based on the interest types found in the IPN study (Langeheine et al., 2001; Rost et al.,1999; Sievers, 1999). Although they have not been replicated yet, more recent findings, such as the interest categories revealed in the HOPE study (Levrini et al. 2016) and the PISA 2006 study (Drechsel et al. 2011), provide further evidence for the existence of interest types.

H5.2: Within each type of interest, different contexts are more (or less) interesting relative to each other.
Background: This hypothesis is based on past empirical studies which show that different contexts are more (or less) interesting relative to each other for different students (e.g. OECD, 2016; Sjøberg & Schreiner, 2012; Häußler, Lehrke, et al., 1998).

RQ6: To what extent can students' types of interest be described in terms of a hierarchy of students' levels of interest? ⇨ related to **RQ2** and **RQ3**

RQ7: In which physics content area are the students overall more interested, mechanics or particle physics?

H7: High-school students are more interested in particle physics than in mechanics.
Background: This hypothesis is based on studies which show that modern physics contents, such as radioactivity or the history of the universe, are more interesting than classical ones (e.g. Häußler, Lehrke, et al., 1998; OECD, 2016; Sjøberg & Schreiner, 2012).

RQ8: To what extent can the students be categorised into different types of physics-related self-concept?

H8: There are no types of students with qualitatively different physics-related self-concept; that is, the students do not differ in their item hierarchies depending on the aspect of physics-related self-concept addressed in an item.
Background: This hypothesis is based on the definition of physics-related self-concept that includes different aspects, namely students' self-perceptions of ability (competency beliefs) and beliefs about the perception as a physics person by others (perceived recognition). Here, the self-perceptions of ability and the beliefs about the perception as a physics person by others may align or differ from one another. However, the physics identity studies suggest that perceived recognition and competency beliefs are strongly associated with one another (Godwin et al., 2016; Kalender et al., 2019b).

RQ9: To what extent is physics-related self-concept a better independent variable than sex[6] for distinguishing between the different types of interest in mechanics and in particle physics?

H9: When using physics-related self-concept instead of sex as an independent variable, the interest types are described better.
Background: This hypothesis is based on the IPN study which showed that the distribution of students with different degrees of physics-related self-concept is more descriptive for the interest types than the distribution of students with different sex (Häußler, Lehrke, et al., 1998). Moreover, the science identity studies showed that physics identity and interest are strongly aligned and do not factor out separately for either gender (Kalender et al., 2019a); that the effect of gender on students' physics identity is mediated by other characteristics, such as self-concept (Kalender et al., 2019a); and that competency beliefs have a very large direct effect on interest and on perceived recognition (Godwin et al., 2016).

---

[6] I decided on using the term 'sex' for describing my study (as defined in **Section 2.3.1**) because the study was conducted in German and there is only one German term ('Geschlecht') for both 'sex' and 'gender'.

RQ10: To what extent does the students' previous experience with the content areas in school affect their expressed interest in these content areas?

H10: Students' interest in a content area differs depending on how much previous experience they had with the specific content area in school.
Background: Previous experience in school refers to the quantitative experience with a content area in contrast to the quality of these experiences. Since previous experience with a content area precedes knowledge, some conclusions about the association between interest and knowledge in a content area may be applied to the association between interest in and quantitative previous experience with a content area. A longitudinal study by Höft et al. (2018) in chemistry education research found that while the students' conceptual understanding generally increases with time, their interest decreases. In addition, a meta-analysis by Schiefele et al. (1992) showed that interest and achievement are positively correlated.

4.4. Structure of this research project

To answer all my research questions, I conducted two studies in the framework of my doctoral research project. In the first study (addressing the RQs 1 to 3), I developed the Instrument to measure students' Particle Physics Interest (IPPI). This study, its methods, and findings have already been published in an article, parts of which are presented in the next chapter (i.e. **Chapter 5**; Zoechling et al., 2022).

Second, the main study, a cross-cohort study with German-speaking students aged 14 to 16 years (N = 1219), was conducted (addressing the RQs 4 to 10). The students' interest in mechanics and particle physics, and the association of interest with different student characteristics, such as physics-related self-concept, sex, and previous experience with the content areas in school, were investigated. This study, its methods, and findings are presented in **Chapter 6**.

5. DEVELOPING THE IPPI

My first study aimed to investigate students' interest in particle physics (i.e. addressing the RQs 1 to 3; see **Section 4.3.**). Modern physics content areas are already included in international physics curricula, such as the International Baccalaureate Physics curriculum, and in several national curricula, such as the Austrian, Italian, and Norwegian (Mullis et al, 2016; Austrian federal law consolidated, 2022). Thus, analysing students' interest in these content areas set in different contexts is of crucial educational significance. Particle physics can be set in various contexts, from everyday life (e.g. digital cameras as particle detectors) to medicine (e.g. particle accelerators in cancer treatment) and existential questions of humankind (e.g. 'Where do we come from?'). I hypothesise that there are different levels of interest in particle physics, each associated with different contexts of particle physics. To examine these levels of interest in particle physics, a conceptualisation of such interest is necessary. Subsequently, the conceptualisation of students' interest in particle physics was the core objective of this study. First, using the findings of past empirical research, the construct 'interest in particle physics' is defined and behaviours that correspond to different degrees of being interested in particle physics are identified. Based on these definitions, a hierarchy of students' levels of interest in particle physics is proposed. Then, the IPPI is developed, using rating scale items that assess different degrees of interest in particle physics. This study, its methods, and findings, have already been published in an article. In this chapter I present the parts of the original manuscript of this article that focus on the methods and findings of this study. The Version of Record of this manuscript has been published and is freely available in the *International Journal of Science Education* (2022) http://www.tandfonline.com/doi/10.1080/09500693.2022.2122897. Please note that the section numbers and cross-links as well as the captions of the figures and tables have been adapted when embedding the manuscript in this thesis to ease the flow of reading. Moreover, unlike the rest of this thesis, the sections below are written in the first-person-plural perspective.

5.1. Method

We developed the IPPI using the Rasch approach described by Liu (2010), Boone et al. (2014), and Planinic et al. (2019). The Rasch procedures are described below. These steps included conceptualising interest in particle physics, creating an initial item pool, piloting potential instrument items, selecting items for the final version of the IPPI, conducting a psychometric analysis of the IPPI, and comparing the theoretical conceptualisation of interest in particle physics to the item hierarchy revealed by the Rasch analysis.

5.1.1. Conceptualising the construct 'interest in particle physics'

When defining the construct to be measured, one basic assumption is that the construct is a unidimensional latent trait that ranges from a lower to a higher level (Liu, 2010). The construct to be measured in our study is interest in particle physics. The linear trait underlying this construct is the degree of interest. Based on previous research (Drechsel et al., 2011; Häußler, Lehrke, et al., 1998; Levrini et al., 2017; Sievers, 1999; Sjøberg & Schreiner, 2012), we identified behaviours that represent different degrees of interest in particle physics. We hypothesised that there are several levels of interest in particle physics, wherein each is associated with different contexts. Students progress through these levels as their interest increases. That is, they become interested in additional contexts. Our focus on the context is based on previous empirical studies that have found or discussed the importance of the context for students' learning progression and achievement in cognitive assessments (Bennett et al., 2007; Härtig et al., 2020; Häußler, Lehrke, et al., 1998; Mesic & Muratovic, 2011; Neumann et al., 2013; Rösler et al., 2018; Sjøberg & Schreiner, 2012; Yao et al., 2017).

We hypothesised a level of focused interest in particle physics by being interested in particle physics solely when it is set in an everyday context, such as the human body or nature. The hypothesised level of focused interest is based on the IPN interest type C, which describes students who are highly interested in physics as it relates to humans, nature, applications, and society (Sievers, 1999); on category A found in the Horizons in Physics Education (HOPE) study, which describes students' 'curiosity to understand the world, natural phenomena and the universe' (Levrini et al., 2017, p. 8); and on the interest category 'living systems' found by Drechsel et al. (2011) for PISA 2006 data. Moreover, we hypothesised that students at a level of broad interest are also interested in particle physics when it is set in a purely scientific context, such as qualitative or quantitative science. This hypothesis is based on IPN interest type A, which describes students who are generally and highly interested in the broad field of physics, that is, even when set in a purely scientific context (Sievers, 1999); on category B found in the HOPE study, which describes students' 'interest in physics knowledge as a special way of knowing, investigating, questioning and thinking' (Levrini et al., 2017, p. 8); and on the interest category 'physical/technology systems' found by Drechsel et al. (2011).

5.1.2. Characterisation of item categories

Based on the conceptualisation of 'interest in particle physics', we characterised item categories for the IPPI. We decided to model the IPPI on the IPN instrument in German (Häußler, Lehrke, et al., 1998) for the following reason. The IPN instrument examines interest in eight different physics content areas, such as mechanics and optics. We found that the structure of the IPN instrument is also well suited to assess interest in the content area 'particle physics': For each content area, the IPN instrument comprises (a) an introductory text and (b) 11 rating scale items regarding students' interest.

(a) The introductory text can cover students' possible deficit in previous experience (see **Section 3.1.**). It provides the students with a short overview of the respective content area set in different contexts aligned with the items. (b) For a certain content area (e.g. mechanics) different content (e.g. lever or pump) are presented in the items. Each item represents a specific item category, that is, a combination of context and task, as listed in **Table 6**[7]. Häußler, Lehrke, et al. (1998) based the distinction of content, contexts, and tasks as well as their definition of item categories on the results of their corresponding preceding Delphi study. Recent empirical studies about interest in physics also consider different learning environments (e.g. Dierks et al., 2016; Blankenburg et al., 2016) while the IPN study focuses on school as a learning environment. Nevertheless, the IPN study was innovative because it presented students with item categories based on unique combinations of different contexts and tasks. Häußler, Lehrke, et al. (1998) explained that they defined 11 item categories to limit the length of the instrument. Although length is certainly an important factor when creating an instrument, we argue

---

[7] In different publications of the IPN study, these 11 item categories are presented in slightly different versions (see Häußler, 1987; Häußler et al., 1996; Häußler, Hoffmann et al., 1998; Häußler, Lehrke et al., 1998; Rost et al., 1999; Sievers, 1999). In Table 6, we list the item category descriptions presented in Häußler et al. (1996) translated into English. Moreover, our ordering of item categories corresponds to the ordering of items as presented to the students in the IPN study.

that it is difficult to formulate every possible combination of context and task because the boundaries between the different contexts and tasks, respectively, are somewhat blurry.

While the distribution of items across the different tasks is uneven, it is even across the different contexts (see **Table 6**). Hence, we found that the IPN item categories are well suited to analysing the interestingness of different contexts but not of different tasks. The contexts used in each IPN item category varied from humans and nature to pure science, and this variety aligns well with our theoretical hierarchy of students' levels of interest in particle physics. We found that the IPN item categories also cover the interest categories revealed by other past empirical studies, such as HOPE and PISA 2006 (Levrini et al., 2017; Drechsel et al., 2011; see **Section 5.1.1.**). Consequently, we decided to model our items on interest in particle physics on the item categories used in the IPN study, including the variation of tasks, because we found that its structure (i.e. introductory text plus items based on item categories) is well suited to assess 'interest in particle physics', as described above. This approach also allows for a later comparison of our results to those of the IPN study.

5.1.3. Creation of the initial version of the IPPI

The IPPI, which is in German, comprises an introductory text on particle physics and items regarding interest in particle physics. The students were asked to read the introductory text and express their degree of interest in each item on a 5-category rating scale ('My interest in it is …' *very high* (=5), *high* (=4), *medium* (=3), *low* (=2), or *very low* (=1)). In March 2020, we developed an initial version of the IPPI. We created a draft introductory text on particle physics and an item pool based on the above-detailed item categories. The introductory text provides the students with a short overview of different particle physics content set in different contexts aligned with the contexts used in the items. The item pool comprised at least three items per category. Exemplar items translated into English are listed in **Table 6** for each item category.

5.1.4. Review and trial of the initial version of the IPPI

Following the creation of the initial version of the IPPI, the draft introductory text and item pool were reviewed by the team of authors. Following this review, the comprehensibility of the draft introductory text and the items was assessed in one-on-one interviews with 16 German-speaking students (9 female, 7 male; grades 8 to 11) in April and May 2020 using a think-aloud protocol according to Sandmann (2014). The students were asked to read aloud and explain their understanding of both the text and items. They were asked to respond to each item and were given the opportunity to provide reasons for the degree of interest they each expressed. Each interview lasted between half an hour and an hour based on the student. Each interview was audio-recorded and transcribed. We conducted a content analysis of the transcripts (Ericsson & Simon, 1993). As a result, we rephrased and shortened parts of the introductory text. For example, while describing the structure of a hair, we initially started at the molecular level and shortened the description so that it began at the atomic level. We also selected three items per category from the pool based on whether students easily understood them. For example, one student commented on one of the items: 'Wie ist das jetzt gemeint? Was ist das? [What is that supposed to mean? What is that?]'. Hence, we did not select this item.

Table 6. Item categories, underlying contexts, and tasks as used in the IPN study (Häußler, Lehrke, et al., 1998) and exemplar items translated into English for each item category developed for the IPPI

#	Item category	Context	Task	Exemplar item (translated into English)
01	Learning more about the functional principle of technical devices	Understanding technical devices in everyday life	Receiving information (observing, reading, and listening)	Learning more about the functional principle of devices that detect particles (e.g. digital camera)
02	Learning more about natural phenomena	Enrichment of emotional experiences		Learning more about how particle physics helps understand the northern lights
03	Learning more about the relevance of physics for society	Relevance for society		Learning more about how a particle accelerator contributes to the peaceful collaboration of diverse nations
04	Learning more about qualitative physics	Science I (qualitative)		Learning more about which interaction binds the elementary particles in the nucleus space together
05	Learning more about quantitative physics	Science II (quantitative)		Learning more about how many elementary particles constitute an object, such as a pen
06	Getting insight into technical jobs	Vocation I (technical, scientific)		Getting insight into how particle accelerators are used in the electronics industry
07	Getting insight into jobs related to humans	Vocation II (medical, artistic, and counselling)		Getting insight into the work flow in a medical diagnostic centre
08	Constructing technical devices	Enrichment of emotional experiences	Hands-on (constructing objects, conducting experiments)	Building a particle detector out of everyday objects
09	Planning experiments	Science I (qualitative)	Minds-on (devising and calculating)	Planning an experiment to explore the structure of an atom
10	Calculating physical quantities	Science II (quantitative)		Calculating the energy when two particles moving with nearly the speed of light collide
11	Discussing the societal relevance of physics	Relevance for society	Evaluation and discussion	Discussing why research in particle physics is important for society

5.1.5. Field testing

Following the think-aloud interviews, we conducted a field test. We utilised the introductory text on particle physics and the 33 corresponding interest items that we developed to create an online questionnaire. The original introductory text and items in German and an English paraphrase of each item are provided in the online appendix for this paper [and in **Appendix 10.1.** for this thesis]. Our field test sought to provide information that would optimise the measurement functioning of the IPPI. In addition, our goal was to lessen the number of developed rating scale items to 11 instead of 33, that is, to one item per category. While conducting a field test to collect data for a Rasch analysis, it is important that the sample population represents the target population and is spread along the construct to be measured (Liu, 2010). The minimum sample size suggested by Linacre (2002) is ten times the number of answer categories, which was 50 for the 5-category rating scale used in our items on interest in particle physics.

To identify a respondent pool, we invited several randomly selected Gymnasium (secondary school) teachers in Austria and Germany via email to participate in our field test. One class each from Vienna (AT), Graz (AT), and a city close to Munich (DE), as well as individual students from three schools in Tyrol (a federal state in AT) completed our online questionnaire. In all, 99 German-speaking grade 9 (aged 15 years) students (57 female, 41 male, 1 not specified) participated in the field test in June 2020.

5.1.6. Rasch analysis

We evaluated the psychometric functioning of the IPPI using a Rasch analysis, which is commonly used while developing new instruments in science education research (e.g. Kirschner et al., 2016; Luo et al., 2019; Neumann et al., 2011; Vorholzer et al., 2016). There are many reasons for using a Rasch analysis: 1) it facilitates the computation of linear measures for persons and items, 2) numerous indices are provided to evaluate the measurement functioning of the instrument, and 3) Wright Maps can be created to evaluate the construct (Wright & Stone, 1979). In our study, the person measure reflects a person's degree of interest in particle physics. The higher the person measure, the higher the person's interest. The item measure reflects the interestingness of an item. The lower the item measure, the higher its interestingness. Person and item measures are expressed on the same linear scale and in the unit of logits.

We used the Rasch partial credit model to analyse our field test data because it allows for the quantitative difference in the degree of interest for each pair of adjacent rating scale categories, for example, from categories 1 to 2, to vary for different items of the instrument (Masters, 1982). Thus, the partial credit model provides insights into the functioning of the rating scale for each individual item. We see this as a potential benefit for achieving the aims of our field test, that is, optimising the measurement functioning of the IPPI and selecting one item per item category from the initial item pool. Additionally, we conducted a comparative analysis of the person and item measures and measurement functioning indices of the IPPI while utilising the partial credit and rating scale models. The data collected were analysed using the Winsteps Rasch programme (version 4.8.1.0), of which the manual provides detailed documentation for users (Linacre, 2021).

5.1.7. Selection of items for the IPPI

Based on the Rasch analysis utilising the partial credit model, we selected one item per item category from the item pool for the final version of the IPPI. We also analysed the data using the Rasch rating scale model. We compared person and item measures and measurement functioning indices of the initial item pool while utilising the partial credit and rating scale models and found them to be very similar. The selection procedure based on the partial credit model analysis is described below.

Category probability curves

First, we examined the category probability curves. According to our theoretical hierarchy of students' levels of interest, students should progress from one rating scale category of an item to another, that is, from categories 1 to 2 and from 2 to 3 and so on, as they progress in terms of interest. Thus, we checked whether the average person measure advances with the advancing rating scale category. This selection criterion is based on Linacre's (2002) Guideline 3 for optimising rating scales. We found that item I033 does not fulfil this criterion. Thus, we removed it from the item pool of the IPPI. In the Rasch partial credit model, the Andrich threshold marks the point where one rating scale category becomes more (or less) probable than another (Linacre, 2021). The items selected for the IPPI must have ordered threshold measures, and every category must be the most probable for some combination of person interest and item interestingness. These selection criteria are based on Linacre's (2002) Guideline 5. For example, for item I081, categories 3 and 4 were never the most probable, and thus the thresholds were not ordered, as seen in **Figure 4a**.

In comparison, for item I042, every category has an individual probability peak, as seen in **Figure 4b**. Linacre (2002) suggested considering non-ordered thresholds as problematic if there are at least ten observations in each rating scale category; this is related to Guideline 1 from Linacre (2002). We found that items I081, I092, and I102 have non-ordered thresholds and ten observations in each category. Thus, we removed these items from the item pool of the IPPI.

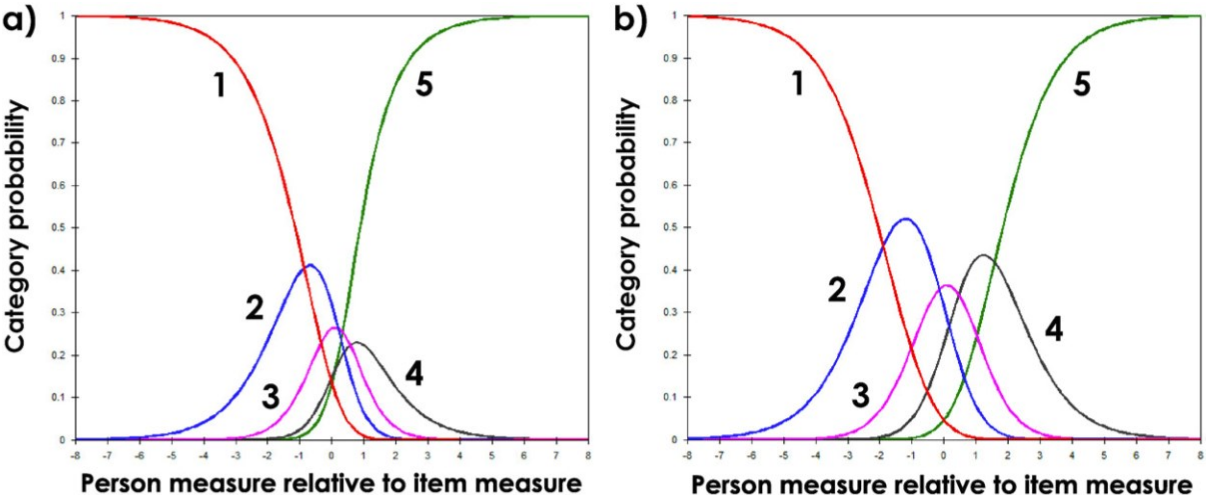


Figure 4. Probability of response for all five rating scale categories as a function of person minus item measure, a) item I081 with non-ordered thresholds (categories 3 and 4 are never the most probable), b) item I042 with ordered thresholds.

Item fit statistics

Second, we examined the item fit statistics, which are based on the difference between what is observed and what is expected by the Rasch model. This difference is considered residual. In general, items with Infit and Outfit mean square residuals (MNSQs) ranging between 0.75 and 1.3 are accepted as having a good model-data-fit (Bond et al., 2020). Although MNSQ values from 0.6 to 1.4 are satisfactory for rating scale data (Boone et al., 2014), we chose the smaller range (0.75 to 1.3) while selecting items for the IPPI. This item fit range is commonly applied while developing new instruments in science education research, as seen in Vorholzer et al. (2016). While analysing the fit statistics for each item (see **Table 7**), we started with the Outfit MNSQs, and if they were not within the acceptable range, we investigated the Infit values (Boone et al., 2014). We excluded five more items from the item pool (I022, I031, I041, I043, and I093) because they presented a possible Outfit and Infit MNSQ misfit. We decided to retain item I032, although its Infit and Outfit MNSQs were slightly above the acceptable range because it had the best fit in this item category. Based on our examination of the category probability curves and item fit statistics, the item selection for the IPPI was already completed for categories 03, 04, and 09 (I032, I042, and I091).

Wright Map

Third, we examined the Wright Map to select one item each for the remaining eight item categories. In the Wright Map, all 33 items of the initial item pool are distributed along the vertical axis according to their item measures (see **Figure 5**). Ideally, the items of an instrument are evenly spread so that they do not measure a similar portion of the trait. In line with our conceptualisation of the interest construct, the map illustrates that many items within the same item category are of a comparable item measure, that is, comparable interestingness, for example, all Category 04 items. Some items within the same item category were not of comparable interestingness (Categories 02, 03, 05, 07, and 08). Thus, we examined the item wording of these categories and found that the contextualisation of items within these categories was not consistent. For example, in Category 02, item I023 was rather set in the context of qualitative physics than in that of the enrichment of emotional experiences. In all, we excluded six items because of their non-comparable interestingness based on the re-evaluation of the respective item wording (I023, I051, I052, I072, I073, and I082). The items for the remaining four item categories (01, 06, 10, and 11) were selected based on the desire to have an even distribution of items on the Wright Map (I012, I063, I101, and I111).

5.1.8. Functioning of the IPPI

After selecting 11 items, that is, one per category, we analysed the respective data subset using the Rasch partial credit model. We conducted a new analysis of the subset of the data collected in the field test that included just the 11 items selected for the IPPI.

To investigate the functioning of the IPPI, we first explored its unidimensionality in several ways. We analysed the item fit statistics because we expected the relative fit indices to change after reducing the number of items by two thirds. Here, we applied the item fit range that is generally accepted for rating scale data, that is, MNSQ values from 0.6 to 1.4 (Boone et al., 2014). We also examined unidimensionality using point measure correlations. Values greater than 0.3 indicate that items measure the same latent trait (Li et al., 2018). Additionally, we evaluated the unidimensionality of the IPPI with a principal component analysis of residuals (PCAR) as described by Boone and Staver (2020). The variance unexplained by the Rasch model caused by the first contrast may be evidence for multidimensionality. A minimum of two items must be considered for a dimension. We examined the wording of the items in the clusters identified in the PCAR and the disattenuated correlation of person measures computed using these clusters of items. A high correlation is evidence that the items of each cluster are measuring the same trait.

Table 7. Rasch item measures and statistics for the initial item pool of the IPPI.

Item ID	Total score	Total count	Item measure [logits]	Model SE [logits]	Outfit MNSQ	Infit MNSQ
I011	310	99	0.11	0.12	0.72	0.76
I012	329	99	-0.23	0.12	0.93	0.97
I013	308	99	-0.04	0.12	0.82	0.81
I021	371	99	-0.99	0.13	1.23	1.16
I022	330	99	-0.2	0.12	1.33	1.38
I023	303	99	0.1	0.11	1.31	1.16
I031	285	99	0.2	0.11	1.96	1.38
I032	374	99	-0.71	0.13	1.31	1.31
I033	342	99	-0.42	0.12	1.72	1.48
I041	253	99	0.59	0.11	0.65	0.69
I042	269	99	0.51	0.12	0.79	0.81
I043	266	99	0.68	0.12	0.62	0.64
I051	324	99	-0.18	0.13	0.91	0.95
I052	303	99	0.2	0.12	0.8	0.81
I053	263	99	0.79	0.13	0.89	0.83
I061	304	99	0.11	0.11	1.11	1.16
I062	300	99	0.12	0.11	1.71	1.3
I063	291	99	0.25	0.12	0.86	0.83
I071	361	99	-0.49	0.12	1.29	1.32
I072	403	99	-1.3	0.15	1.08	1.13
I073	360	99	-0.51	0.13	1.09	1.11
I081	364	99	-0.47	0.10	0.9	0.97
I082	307	99	0.05	0.11	0.89	0.86
I083	370	99	-0.47	0.11	0.94	1.06
I091	319	99	0.02	0.11	0.78	0.83
I092	274	99	0.41	0.11	0.67	0.69
I093	290	99	0.28	0.11	0.69	0.71
I101	279	99	0.41	0.10	0.9	0.91
I102	293	99	0.3	0.11	1.22	1.31
I103	241	99	0.82	0.11	0.78	0.81
I111	303	99	0.14	0.12	0.82	0.87
I112	306	99	0.09	0.12	0.8	0.83
I113	328	99	-0.17	0.11	1.24	1.2

Note: The first two digits of the item ID indicate the item category. Total score refers to the total raw score of all respondents who answered the item. Total count refers to the total number of respondents who answered the item. Measure refers to the Rasch item measure in logit units. Lower and higher item measures represent more and less interesting items, respectively. Model SE refers to the standard error of the item measure in logit units. Outfit MNSQ refers to a fit statistic that is sensitive to extreme responses. Infit MNSQ refers to a fit statistic utilising weighted means.

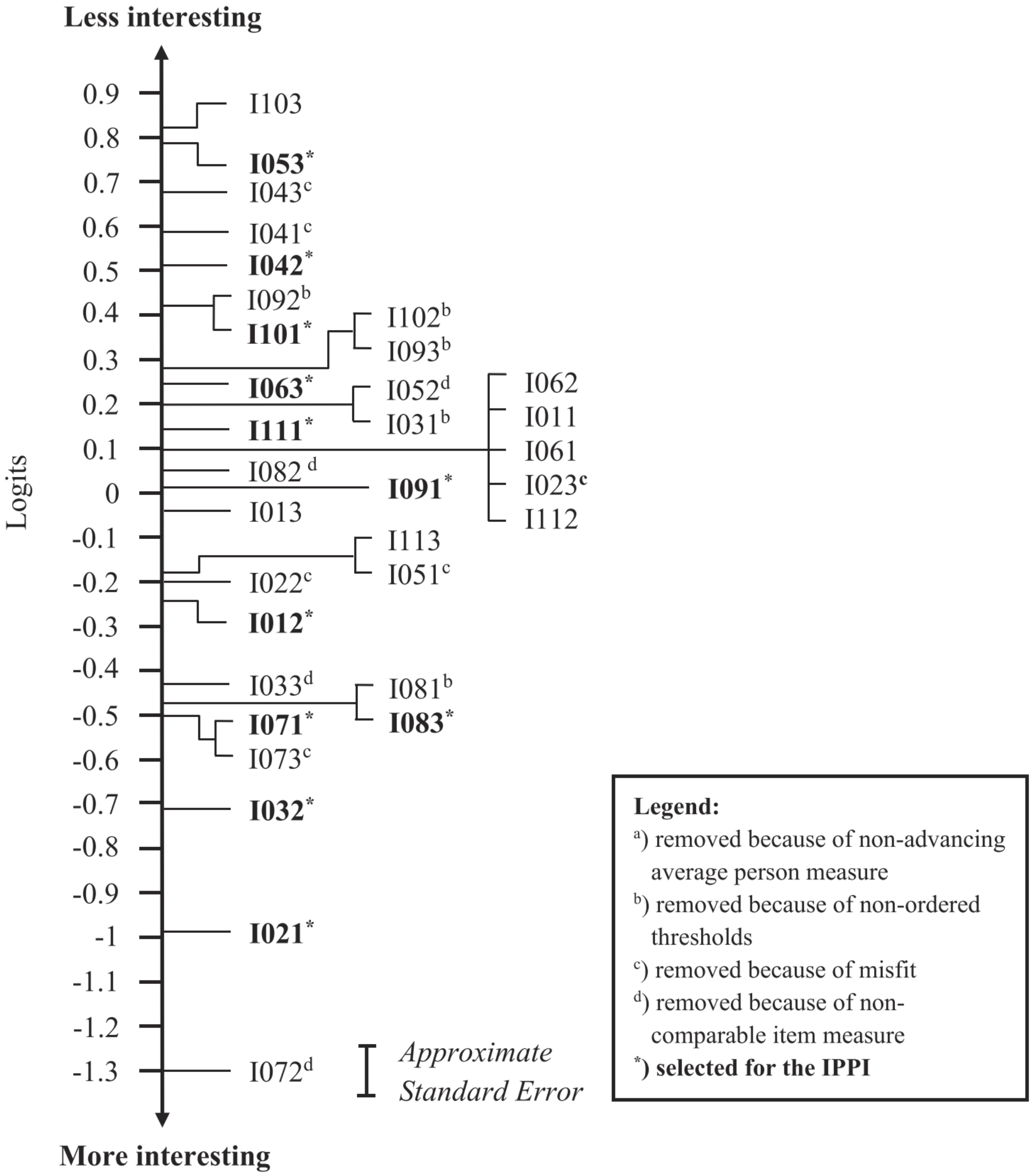


Figure 5. Wright Map of the initial item pool of the IPPI.
Note: Items are represented with their item ID. Items are sorted according to their item measures. Lower and higher item measures (base and top of the map, respectively) represent more and less interesting items, respectively.

To check for item independence, we examined the residual correlations between items by pairs. Correlation values smaller than 0.7 imply that two items are independent, that is, the response to one is independent of that to another (Linacre, 2021). We also explored the distribution of items across the latent trait by analysing the Wright Map.

Finally, we examined the summary statistics, which provide several indices that are used for monitoring instrument functioning, such as item and person separation and their respective reliabilities. Item separation and reliability values greater than 4 and 0.9, respectively, imply that the sample size is large enough to verify the item hierarchy (Linacre, 2021). Person separation values greater than 2 indicate a good level of separation, and person reliability values greater than 0.8 imply that the measurement instrument can distinguish between two or three levels of interest (Linacre, 2021). We also explored the mean item and person measures listed in the summary statistics to draw conclusions on their relationship.

We also analysed this data subset using the Rasch rating scale model. We compared person and item measures and measurement functioning indices of the IPPI while utilising the partial credit and rating scale models.

5.1.9. Validation of the conceptualisation of interest in particle physics

To investigate how the latent trait defined by the IPPI aligns with the theoretical hierarchy of levels of interest in particle physics, we examined the item measures and wordings. In line with previous findings, we hypothesised that there are two levels of interest, each associated with different contexts of particle physics. If the IPPI displays our hypothesised hierarchy of interest levels, item measures should depend on the context that the item in question represents. Items set in an everyday context, such as the human body or nature, should be among the most interesting, whereas those set in a purely scientific context, such as qualitative or quantitative science, should be among the least interesting.

5.2. Results

The data obtained with the IPPI were analysed in two steps. In the first step, we investigated whether the IPPI functioned in a psychometrically sound manner, and in the second step, we investigated whether the item hierarchy documented a hierarchy of students' levels of interest in particle physics as hypothesised in our conceptualisation.

5.2.1. Instrument functioning

Table 8 lists the Rasch item measures and statistics for the IPPI. The Outfit MNSQ values ranged from 0.79 to 1.36, which is within the acceptable range. All point measure correlation values were above the suggested 0.3 cut-off value (ranging from 0.59 to 0.71). As a result of the PCAR, the unexplained variance of the first contrast was found to be 2.1 (item) units. This may be evidence for a secondary dimension among the items with a strength of about two items. Thus, we analysed the item wordings in each cluster identified in the PCAR (cluster 1: I071, I032; cluster 2: I021, I083, I111, I012, I053; and cluster 3: I042, I101, I063, I091). We found that the items did not share any substantive latent trait other than the single Rasch dimension we hypothesised based on our theory, that is, interest in particle physics. We evaluated the disattenuated correlation of person measures computed through the clusters of items ((a) clusters 1 and 2: r=0.91, (b) clusters 2 and 3: r=0.93, and (c) clusters 1 and 3: r=0.35). The high correlations obtained between person measures (a) only with cluster-1 and cluster-2 items and (b) only with cluster-2 and cluster-3 items suggest that the items defined a single trait. However, the correlation obtained (c) only with cluster-1 and cluster-3 items is lower.

Table 8. Rasch item measures and statistics for the IPPI.

Item ID	Total score	Total count	Item measure [logits]	Model SE [logits]	Outfit MNSQ	Infit MNSQ	PT. corr.
I053	263	99	0.94	0.13	0.92	0.85	0.70
I042	269	99	0.63	0.12	0.81	0.83	0.70
I101	279	99	0.51	0.11	0.91	0.86	0.67
I063	291	99	0.35	0.13	0.83	0.80	0.71
I111	303	99	0.23	0.12	0.98	1.02	0.66
I091	319	99	0.09	0.11	0.79	0.83	0.69
I012	329	99	-0.17	0.12	0.96	0.96	0.66
I083	370	99	-0.44	0.11	1.00	1.09	0.63
I071	361	99	-0.45	0.12	1.36	1.27	0.59
I032	374	99	-0.67	0.13	1.32	1.25	0.59
I021	371	99	-1.02	0.13	1.14	1.26	0.62

Note: Total score refers to the total raw score of all respondents who answered the item. Total count refers to the total number of respondents who answered the item. Measure refers to the Rasch item measure in logit units. Lower and higher item measures represent more and less interesting items, respectively. Model SE refers to the standard error of the item measure in logit units. Outfit MNSQ refers to a fit statistic sensitive to extreme responses. Infit MNSQ refers to a fit statistic utilising weighted means. PT. corr. refers to the point measure correlations.

Nevertheless, based on the item fit statistics, the point measure correlations, and the analysis of item clusters revealed by the PCAR, we felt that the data supported one trait. Thus, we consider the IPPI unidimensional yet broad because, according to our definition, interest in particle physics includes several aspects such as contexts or tasks (cf. Linacre (2021) about mathematics as a broad domain). Local independence for each item was supported, as the correlation values between item residuals were smaller than 0.7 for all item pairs. The Wright Map of the IPPI revealed that the items were spread along the latent trait (**Figure 6**). Examining the summary statistics, the item separation of the IPPI was determined to be 4.43 with an item reliability of 0.95. The person separation for our data sample was found to be 2.53 with a person reliability of 0.86. The item measures ranged from -1.02 to 0.94 logits. The mean item measure was set to zero logits in Rasch analysis and the model standard error of items was 0.12 logits. The mean person measure was found to be 0.26 logits with a model standard error of 0.11 logits and ranged from -5.34 to 5.55 logits.

Finally, we compared person and item measures and measurement functioning indices of the IPPI while utilising the partial credit and rating scale models. **Table 9** presents a selection of results from this comparative analysis. All key indices that are commonly reviewed for Rasch analyses are very similar while comparing the results of the rating scale and partial credit model analyses.

Table 9. Selected Rasch statistics for the IPPI utilising the Rasch partial credit model and Rasch rating scale model (The IPPI comprises a total of 11 items) | *Item I083: 1.42

Rasch analysis method	Item		Person		Outfit MNSQ	Infit MNSQ
	Separation	Reliability	Separation	Reliability	*Number of items falling within the range of 0.6-1.4*	
Partial credit model	4.43	0.95	2.53	0.86	11	11
Rating scale model	4.48	0.95	2.50	0.86	11	10*

5.2.2. Validation of the conceptualisation of interest in particle physics

To investigate how the latent trait defined by the IPPI aligns with the theoretical hierarchy of the levels of interest in particle physics, item measures and wordings were analysed as illustrated in the Wright Map (**Figure 6**) and **Table 8**. When interpreting the Wright Map, it is crucial to consider that it illustrates a hierarchy of items. This means that the items with a low item measure, that is, the most interesting items (bottom of the map), are interesting for most of the students, and the least interesting items (top of the map) are interesting only for some students, the most interested ones. This also means that the persons with a high person measure, that is, the most interested persons, are interested in most of the items, even in the least interesting ones, and the least interested persons are only interested in some items, the most interesting ones.

The three least interesting items (I053, I042, and I101) present particle physics set in the context of qualitative or quantitative science. This is in line with our hypothesis that only students at a level of broad interest are interested in particle physics when it is set in a purely scientific context. The slightly more interesting item I063 presents particle physics set in the context of technical vocation. The even more interesting items (I111, I091, and I012) present particle physics set in the context of everyday life. The most interesting items present particle physics in different contexts. Only item I071 (specific context 'medical diagnostics') is in line with our hypothesis that students at a level of focused interest are solely interested in particle physics when it is set in an everyday context, such as the human body or nature.

In general, the analysis of the item wordings demonstrated that the specific context mentioned in each item is crucial for the degree of interest expressed. We argue that when the specific context mentioned in the item was more precise, the students expressed a higher interest in an item. For example, in item I012, a very precise everyday example ('digital camera') is provided as the specific context, whereas the specific context mentioned in item I111 is very broad ('everyday life'). We see this pattern as item I012 is perceived as more interesting than item I111. We also observed that students expressed higher interest in items that mentioned a hands-on task. For example, although item I091 is set in a purely scientific context, students expressed higher interest in I091 than in the other items set in a purely scientific context because the word 'experiment' is mentioned in I091.

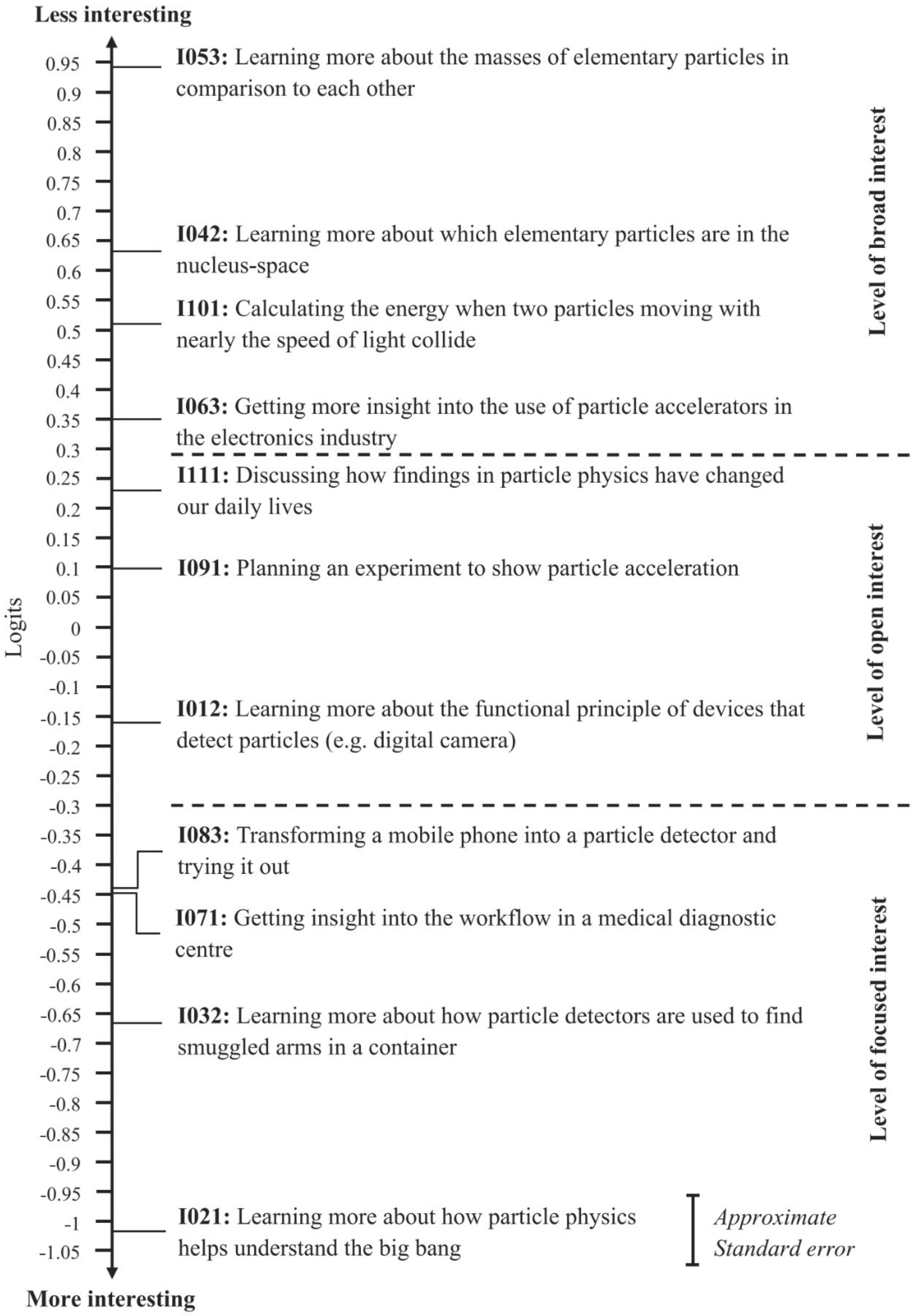


Figure 6. Wright Map of the IPPI.

Note: Items are represented with their item ID and their item wording translated into English. Items are sorted according to their item measures. Lower and higher item measures (base and top of the map, respectively) represent more and less interesting items, respectively. The map also shows the refined hierarchy of levels of interest in particle physics (dotted lines mark the transition from one level to another).

5.3. Discussion

This study sought to conceptualise students' interest in particle physics and develop and evaluate the IPPI. We discuss the use of the Rasch partial credit model in developing and evaluating the IPPI, its functioning, and whether the findings on students' interest in particle physics matched the theoretical conceptualisation of the construct.

5.3.1. Using the Rasch partial credit model

We used the Rasch partial credit model, where each item is considered to have its own rating scale. That is, the partial credit model allows for the quantitative difference in the degree of interest for each pair of adjacent rating scale categories, such as from category 1 to 2, to vary for different items of the measurement instrument. We see this as a potential benefit for developing and evaluating instruments. We also analysed the data using the Rasch rating scale model, where the quantitative difference in the degree of interest for each pair of adjacent rating scale categories is the same for all items of the instrument. All key indices that are commonly reviewed for a Rasch analysis were very similar while comparing the results of the rating scale and partial credit model analyses. There are benefits to using both models. Thus, researchers should use both models to analyse rating scale data to gain additional insight into the functioning of the rating scale of each item.

5.3.2. Instrument functioning

The results concerning our first research question are discussed in this section. For our measurement instrument, four aspects of validity evidence are relevant: content, construct, statistical, and fit validity. We present the details below.

To ensure content validity, the construct to be measured must be conceptualised in accordance with a theoretical grounding and previous findings and represented through items forming the measurement instrument (Boone & Staver, 2020). The IPPI was designed to have content validity evidence as we first defined the construct, that is, interest in particle physics, and identified related behaviours based on previous findings and theoretical grounding. Then, we developed representative items forming the IPPI.

Second, construct validity evidence means the degree to which the item hierarchy matches the predictions based on the theoretical construct. In a Rasch analysis, construct validity is evaluated by analysing the item ordering and spacing on the Wright Map (Boone & Staver, 2020). One could argue that the item hierarchy could be a matter of item construction rather than of context. Analysis of the Wright Map of the initial item pool consisting of three items per IPN item category demonstrated that items constructed for the same item category are not consistently of comparable item measure (see **Figure 5**). Thus, we did not refer to the IPN item categories to describe the item hierarchy of the IPPI. Instead, we introduced the three levels of interest in particle physics based on context. These levels describe the item hierarchy of the IPPI and, ultimately, students' interest in particle physics. Using this approach, we found that the item hierarchy can be interpreted in keeping with previous findings on interest in physics. To verify the item hierarchy, we examined item separation and reliability and found that our data fell in the rule of thumb guideline ranges as needed for both item separation (>4.0) and reliability (>0.9). To investigate statistical validity, person reliability is evaluated with a Rasch analysis (Boone & Staver, 2020). Person reliability refers to the reproducibility of the person ordering, which can be interpreted on the lines of Cronbach's Alpha in classical test theory. Our data fell within the rule of thumb guideline ranges as needed for both person separation (>2.0) and reliability (>0.8). This means that the IPPI is sensitive enough to distinguish between persons with two or three different levels of interest. Previous findings and the current results suggest two to three different levels of interest. Thus, the IPPI provides useful and informative measures for the intended purpose.

Finally, fit validity evidence refers to the degree to which the data fit the Rasch model (Boone & Staver, 2020). To ensure fit validity, we analysed the dimensionality of the IPPI by examining the item fit statistics and conducting a PCAR. Our data fit the Rasch model, which supports the fit validity evidence of the IPPI.

In summary, our Rasch analysis provides evidence supporting the content, construct, statistical, and fit validity of the IPPI.

5.3.3. Validation of the conceptualisation of interest in particle physics

The results concerning our second and third research questions are discussed in this section. We hypothesised that students at a level of focused interest in particle physics evince such interest solely when particle physics is set in an everyday context, such as the human body or nature, and that students at a level of broad interest are also interested in particle physics when it is set in a purely scientific context, such as qualitative or quantitative science.

Based on our analysis of the item hierarchy, we refined this conceptualisation of interest in particle physics (see **Figure 6**). First, we characterised the level of focused interest in particle physics as being interested in particle physics when it is set in a context that is related to (1) one's own body, for example, 'medical diagnostics' (I071); (2) socio-scientific issues, for example, 'smuggled arms' (I032); or (3) existential questions of humankind, for example, 'big bang' (I021). We found that aspect (3) caused a high interest in particle physics, although this context is theoretical, like purely scientific contexts. We believe that these three aspects can be assigned to the same level of interest because they all sparked interest by arousing positive or negative emotions. This aligns with the first phase in interest development, 'triggered situational interest', in which the emotional component prevails (Hidi & Renninger, 2006).

Second, we suggest introducing an additional level of interest, the level of open interest, to our hierarchy of levels of interest. Students at the level of open interest were additionally interested in particle physics when it was set in the broad context of 'everyday life'. Our definition of the level of open interest aligns with the second phase of interest development, namely 'maintained situational interest' as proposed by Hidi and Renninger (2006). In this phase, a person begins to recognise personal value based on already existing positive feelings, that is, the value-related component of interest prevails.

Third, to align with the least interesting items, we refined our definition of the level of broad interest in particle physics as being interested in particle physics, even when it is set in a purely scientific or technical context. Our definition of the broad level aligns with the third and fourth phases in interest development, namely 'emerging' and 'well-developed individual interest', where the cognitive-epistemic component of interest prevails (Hidi & Renninger, 2006).

In summary, we found that the data collected with the IPPI results in an item hierarchy that aligns with earlier findings on students' interest in physics (Bennett et al., 2007; Drechsel et al., 2011; Häußler, Lehrke, et al., 1998; Levrini et al., 2017; Sjøberg & Schreiner, 2012) and the hierarchy of interest components in the 'Four-phase model of interest development' (Hidi & Renninger, 2006). In this section, we also refined and described the proposed levels of interest in particle physics qualitatively by associating each level with different contexts.

5.4. Strengths, limitations, and directions for future research

This study conceptualised interest in particle physics and developed and evaluated a corresponding measurement instrument, the IPPI, utilising a Rasch analysis.

One strength of our study is that by conducting a Rasch analysis for developing and evaluating the IPPI, we could draw conclusions on the interestingness of each item with respect to other items and the students' sample. Moreover, utilising the partial credit model helped us gain additional insights into the functioning of each item's rating scale and introduce a novel approach in selecting items from an initial item pool. This novel approach may be useful for other researchers developing instruments.

Our sample size (N = 99) was large enough to conduct a Rasch analysis. Similar sample sizes have been used in developing new instruments, such as N = 103 in Luo et al. (2019). However, larger sample sizes are needed to evaluate differential item functioning, such as for gender, and to collect additional evidence for the different levels of interest in particle physics. Another limitation is the lack of generalisability of the results as the measurement instrument was developed in German and the sample was German-speaking. To enhance generalisability, the instrument must be translated, and data from other countries must be considered as well.

While defining different levels of interest, one limitation is that the latent trait, that is, the degree of interest, is continuous and not discrete. Nevertheless, we could define three levels of interest based on the qualitative descriptions of the contexts mentioned in the items. However, the definitions of contexts used to characterise the students' levels of interest cover broad and overlapping ranges of specific contexts. Another limitation of our study is that, although different tasks were mentioned in the items, we did not examine their effects on the degree of interest expressed in detail.

In our discussion, we interpreted the Wright Map so that the levels of interest in particle physics are cumulative. For example, students at the level of broad interest are interested in additional contexts compared to the level of open interest. This is an assumption of the Rasch model used for analysing the data, and we have shown that the data fits the Rasch model well. We also discussed that the hierarchy of levels of interest aligns with the four-phase model of interest development (Hidi and Renninger, 2006). It remains an open question for future longitudinal studies whether students indeed progress through the levels of interest as their interest increases. Moreover, future studies could investigate whether this hierarchy of students' levels of interest in particle physics can also be applied to other modern physics content areas.

This study is part of a larger research project on different types of interest in physics among students. The project aims to examine and compare students' levels of interest in particle physics and mechanics. A large dataset (N > 1000) is now being collected with the IPPI developed in this study and the original IPN instrument to measure interest in mechanics (Häußler, Lehrke, et al., 1998). In this project, we will also examine whether the different levels of interest correlate with different student characteristics, such as sex and physics-related self-concept.

5.5. Conclusions and implications for practice

Based on previous findings, we conceptualised interest in particle physics. Context was a crucial aspect of fostering interest among students in past empirical research. Accordingly, we suggested a theoretical hierarchy of students' levels of interest in particle physics based on context. That is, different levels of interest among students were mainly determined by the context in which the physics content was set. We created the IPPI based on the 11 item categories introduced in the IPN study (Häußler, Lehrke, et al., 1998). Initially, we created at least three items for each category. Applying the Rasch partial credit model, we ultimately selected one item per category following clear, stepwise, and reproducible criteria based on the category probability curves, item fit indices, and the sign of an even distribution of items on the Wright Map. The Rasch analysis provided evidence supporting content, construct, statistical, and fit validity of the IPPI. The results demonstrate that we have successfully developed a valid and reliable instrument to measure interest in particle physics, and we conclude that the IPPI can be used in future studies. We also interpreted the hierarchy of students' levels of interest based on the results of past empirical research and the four-phase model of interest development. We found that: (1) students at a focused level of interest are interested in particle physics when set in a context, which arouses emotions; (2) students at an open level of interest were additionally interested in particle physics when it was set in an everyday life context; and (3) students at a level of broad interest were even interested in particle physics when it was set in scientific and technical contexts. When interpreting the hierarchy of students' levels of interest, it is crucial to consider that these levels are cumulative. This means that the level of broad interest includes the level of open interest, which further includes the level of focused interest. We conclude that there are groups of students that are characterised by different levels of interest in particle physics and that these levels describe cumulative, not mutually exclusive interests.

Comparing our results to those of the IPN study, we draw the following conclusions. For teaching physics Häußler, Hoffmann, et al. (1998) recommended 'providing opportunities to be amazed', 'encouraging discussions and reflections on the social importance of science', and 'showing physics in relation to the human body' (p. 236-237). This aligns well with our description of the level of focused interest in particle physics and our finding that most students are interested in contexts that arouse emotions, that is, contexts related to one's own body, socio-scientific issues, or existential questions of humankind. Moreover, they recommended 'linking content to prior experiences for both boys and girls' and 'letting physics appear in application-oriented contexts' (p. 236). This aligns well with our description of the level of open interest, which describes students that are additionally interested in particle physics when set in the broad context of 'everyday life'. We also conclude that Häußler, Hoffmann, et al.'s (1998) recommendation that 'the teaching of physics should de-emphasise physics for physics' sake' (p. 236) aligns well with our finding that only some students are at the level of broad interest in particle physics, that is, are even interested in purely scientific or technical contexts. Additionally, we conclude that the IPN interest types can be described better by applying our hierarchy of levels of interest. In particular, our levels of focused and open interest provide a more detailed description of the IPN interest type C, which comprises students who are highly interested in physics related to humans, nature, applications, and society (Sievers, 1999). Similarly, our level of broad interest also describes the

IPN interest type A, which comprises students who are generally and highly interested in the broad field of physics, that is, even when set in a purely scientific context (Sievers, 1999).

For educational practice, we imply that knowing and understanding this hierarchy of students' levels of interest in particle physics can help educators who seek to foster their students' interest. They can match the design of their learning activities with the different levels of interest in particle physics among their students. Here, we outline the following recommendations for educators. For most students, it is crucial to trigger emotions by highlighting the relationship between particle physics and one's own body, socio-scientific issues, or existential questions of humankind to catch their interest. For fewer students, it is important to highlight the personal value of particle physics by setting it in an everyday context to hold their interest. Only when educators aim to tackle the interest of the even fewer students at the level of broad interest, we recommend cognitive-epistemic learning activities set in a purely scientific or technical context.

Finally, we suggest that educators implement the recommendations given based on this hierarchy of levels of interest in particle physics to other modern physics content areas, especially if they can be set in similar contexts.

6. MAIN STUDY

In the first study, I developed an Instrument to measure Particle Physics Interest (IPPI) and proposed a hierarchy of levels of interest in particle physics (see **Chapter 5**). Second, the main study of my doctoral research project focused on the students' types of interest in mechanics and particle physics, as well as on the association of students' interest with their physics-related self-concept and with their previous experience with the content areas in school (i.e. addressing the RQs 4 to 10; see **Section 4.3.**). In this chapter I present the design and findings of this study.

6.1. Research design of the main study

The main study of my doctoral research project was a cross-cohort study with German-speaking students aged 14 to 16 years ($N = 1219$). The measurement instrument was administered online to the students from May to September 2021. The students' interest in mechanics and particle physics was assessed using the instrument to measure mechanics interest from the IPN study and the IPPI. In addition, different student characteristics, such as their physics-related self-concept, sex, and previous experience with the content areas in school, were assessed. The main aims of this study were to investigate students' types of interest in physics utilising a mixed Rasch analysis and their association with different student characteristics utilising different regression and correlation analyses. I also analysed whether the hierarchy of students' levels of interest that I had developed in the first study for particle physics can be applied to students' different types of interest in mechanics and in particle physics. In addition, I analysed to what extent the students' previous experience with a content area in school affects their interest in it. For mechanics, the influence of the introductory text was also analysed by randomly assigning the students to two different versions of the measurement instrument. Below I describe the measurement instrument used in this study, the participants, and the preparation of the collected data.

6.1.1. Measurement instrument

I used an online measurement instrument in German language to assess the students' interest in mechanics and particle physics, physics-related self-concept, sex, previous experience with the content areas in school, and background as listed below. I implemented all the items in the 'SoSci Survey' platform. I created two different versions of the online questionnaire, one with the original IPN mechanics introductory text and one with a newly developed text. The 'SoSci Survey' platform was chosen because it allows for the creation of one link that can be sent to the students and randomly assigns them to one of the two different versions of the questionnaire. Both versions of the **complete online measurement instrument in German** used in my main study are provided in **Appendix 10.2.5.** and **10.2.6.** for this thesis. Below I will present the original English versions or English translations[8] of the questions, exemplar items, and corresponding scales to ease the flow of reading.

(1) The **students' interest in mechanics** was assessed using the corresponding introductory text and interest items used in the IPN questionnaire for half of the students. In addition, the other half of the students was presented with a new introductory text on mechanics

---

[8] The English translations provided have not been validated yet.

instead of the original IPN text, while keeping the original interest items. This enabled me to evaluate the influence of the introductory text on the expressed interest (RQ4, see **Section 4.3.**). The original IPN introductory text and the newly developed introductory text (both in German and an English translation) and the original IPN items (in German and an English paraphrase[9] of each item) are provided in **Appendix 10.2.1.**, **10.2.2.**, and **10.2.3.** for this thesis. The 11 items on students' interest in mechanics were based on predefined item categories (Häußler, Lehrke et al., 1998; see **Section 3.4.1.**). Each item category describes the context and the task, in which a content is set, when formulating an item. **Table 10** lists three exemplar item categories and corresponding items in English translation. Each item has a corresponding item ID. For example, the item M01 measures interest in mechanics (M) and the two digits indicate the item category (01). For all items the question was: 'How interested are you in doing the following related to this topic?'. The students expressed their degree of interest in each item on a 5-category rating scale ('My interest in this is …' *very high* (=5), *high* (=4), *medium* (=3), *low* (=2), or *very low* (=1); Häußler, Lehrke et al., 1998)[10]. I used the original IPN instrument for half of the students to enable a comparison of my results about the students' types of interest in mechanics to the results of the IPN study that was conducted in the 1980s and 90s. This approach allows for drawing conclusions about the applicability of the IPN results to today's students. The other half of the students was presented with a new introductory text on mechanics, while keeping the original items to measure students' interest in mechanics. The original IPN introductory text presents mechanics mostly in contexts related to cars and accidents. When developing the new text, the idea was to present mechanics in a broader variety of contexts that is closer to the life of today's students. This new text on mechanics was developed in March 2021. First, I created two different draft versions of a new introductory text on mechanics. The draft introductory texts were reviewed by my supervisors. As a result, one text was chosen because its content is more similar to the original text and better aligns with the items. Next, the comprehensibility of the chosen draft introductory text was assessed in one-on-one interviews with 8 German-speaking students (5 female, 3 male; grades 8 to 9) in April and May 2021 using a think-aloud protocol according to Sandmann (2014). In the interviews the students were also presented with the IPN items to measure mechanics interest and the IPPI. The students were asked to read aloud and explain their understanding of the texts and items. Each interview lasted between half an hour and an hour based on the student. Each interview was audio-recorded and transcribed. I conducted a content analysis of the transcripts (Ericsson & Simon, 1993) focusing on the new mechanics introductory text. The analysis confirmed that the introductory text is comprehensible for the students without any changes. I argue that students express operating interest in each item while filling in such a measurement instrument consisting of an introductory text and corresponding items. Operating interest in the form of situational interest may be caused by the interestingness of an item or of the introductory text that acts as a prior stimulus. Operating interest may also be caused by the students' already existing individual interest (see the **Section 2.4.**).

---

[9] I am only providing an English paraphrase since an English translation of these items has not been validated yet.

[10] **German original:** ‚Wie gerne würdest du im Zusammenhang mit diesem Thema das Folgende tun?' ⇨ ‚Mein Interesse daran ist …' *sehr groß* (=5), *groß* (=4), *mittel* (=3), *gering* (=2), oder *sehr gering* (=1).

(2) The **students' interest in particle physics** was assessed using the IPPI (instrument to measure students' interest in particle physics) consisting of an introductory text and 11 items that I have developed in my first study (Zoechling et al., 2022; see **Chapter 5**). The introductory text (in German and an English translation) and the items (in German and an English paraphrase of each item) are provided in the **Appendix 10.1.1.** and **10.1.2.** for this thesis. **Table 10** also shows three exemplar items. Each item has a corresponding item ID. For example, the item PP01 measures interest in particle physics (PP) and the two digits indicate the item category (01). The IPPI uses the same 5-category rating scale as described above for the instrument to measure mechanics interest. The IPN instrument acted as a model for the IPPI to enable a comparison of our results about the students' types of interest in particle physics to our results about mechanics. Hence, I can draw conclusions about the universality of the students' types of interest. As above for the items used to measure mechanics interest, I argue that students express operating interest in each item while filling in such a measurement instrument consisting of an introductory text and corresponding items.

Table 10. Three exemplar IPN item categories (Häußler, Lehrke et al., 1998), the corresponding IPN items on interest in mechanics (Häußler, Lehrke et al., 1998), and the items on interest in particle physics developed for the IPPI. Categories and items are translated into English. The items in German and an English paraphrase[11] of each item are provided in the **Appendix 10.1.2.** and **10.2.3.** for this thesis.

#	IPN item category	IPN item on interest in mechanics (ID)	IPPI item (ID)
04	Learning more about qualitative physics	Learning more about how the kinetic energy of a car is transformed into other forms of energy (e.g. in the brakes or in the crumple zone) (M04)	Learning more about which elementary particles are in the nucleus space of an atom (PP04)
07	Getting insight into jobs related to humans	Learning more about the artificial organs (e.g. heart as blood pump) and joints used in medicine (M07)	Gaining insight into the workflow in a medical diagnostic centre (PP07)
08	Constructing technical devices	Constructing different pulleys and trying them out (M08)	Transforming a mobile phone into a particle detector and trying it out (PP08)

(3) The **students' interest in the content areas** was assessed using two items, one on mechanics aligned with the titles of the introductory texts ('About movements' and 'Mechanics', respectively) and one on particle physics ('Particle physics'). The students had to rate their interest on a 5-category rating scale ('My interest in this is …' *very high* (=5), *high* (=4), *medium* (=3), *low* (=2), or *very low* (=1).

---

[11] I am only providing an English paraphrase since an English translation of the IPPI has not been validated yet.

(4) The **students' physics-related self-concept** was assessed using a combination of two measurement instruments to align with the definition of physics-related self-concept given in **Section 2.3.2.**. In particular, the two aspects 'ability self-concept' (self-perception of physics ability) and 'perceived recognition' (beliefs about the perception as a physics person by relevant others) were assessed. The item set used to assess self-concept in the IPN study also addressed both aspects (Häußler, Lehrke et al., 1998). The items in German and their English version are provided in **Appendix 10.2.4.** for this thesis. For all items the question was: 'How much do you agree with the following statements?'. The students expressed their degree of agreement in each item on a 4-category rating scale ranging from *fully agree* (=4), *rather agree* (=3), *rather disagree* (=2), to *fully disagree* (=1) (Frey et al., 2009)[12]. Each item has a corresponding item ID. For example, the item SC01 measures physics-related self-concept (SC) and the two digits indicate the item number (01).

  a) The **physics ability self-concept** (self-perception of ability) was assessed using the corresponding instrument in German to assess science ability self-concept from the PISA 2006 study (see **Section 3.4.3.**; Frey et al., 2009). It consists of 6 items. I adapted the item wording to measure the physics ability self-concept instead of the science ability self-concept (e.g. 'Physics topics are easy for me' instead of 'Science topics …'). This is based on Marsh (1990) who argues that ability self-concept is domain specific.

  b) The **physics perceived recognition** (beliefs about the perception as a physics person by others) was assessed using the corresponding instrument consisting of three items from Kalender et al. (2019a, 2019b, 2020; see **Section 3.4.10.**; e.g. 'My parents see me as a physics person'). I translated the original items from English into German and added a fourth item about students' perceived recognition by classmates. Evidence for the validity of these four items in German language are provided in my study.

(5) The **students' previous experience with the content areas in school** was assessed using seven items each for mechanics and particle physics. For all items the question was: 'How exhaustively was this content covered in class?'. The students expressed their degree of experience with the content presented in each item on a 5-category rating scale ranging from *very exhaustively* (=5), *exhaustively* (=4), *medium* (=3), *just briefly* (=2), to *not at all* (=1)).[13] The seven different contents presented in the items to assess previous experience with mechanics were: 'force', 'velocity', 'energy of motion', 'pump', 'pulley', 'lifting platform', and 'braking distance'. The seven different contents presented in the items to assess previous experience with particle physics were: 'elementary particles', 'particle detector', 'particle accelerator', 'nucleus space', 'electron', 'the big bang', and 'speed of light'.

---

[12] **German original:** ‚Wie sehr stimmst du mit den folgenden Aussagen überein?' ⇨ *stimme ganz zu* (=4), *stimme eher zu* (=3), *stimme eher nicht zu* (=2), oder *stimme gar nicht zu* (=1).

[13] **German original:** ‚Wie ausführlich wurde dieses Thema im Unterricht behandelt?' ⇨ *sehr ausführlich* (=5), *ausführlich* (=4), *mittel* (=3), *nur kurz* (=2), oder *gar nicht* (=1).

(6) The **students' sex** was assessed using one item ('Please tick your sex': *female* (=1), *male* (=2), *not specified* (=3)).

(7) The **students' age** was assessed using one item ('How old are you?': *14 years* (=1), *15 years* (=2), *16 years* (=3), *other* (=4)).

(8) The **students' language most spoken at home** was assessed using one item ('Which language do you speak most at home?': *German* (=1), *another language* (=2), *both equally* (=3)).

(9) The **students' school grade** was assessed using one item ('What grade are you in?': *grade 8* (=3), *grade 9* (=1), *other* (=2)).

(10) The **students' school type** was assessed using one item ('What school type are you in?': *Gymnasium* (=1), *other* (=2)).

(11) **Class code:** Each participating class was assigned to a class code consisting of three characters. The first is a letter, indicating the country (A=Austria, C=Switzerland, D=Germany). The second and third are digits, representing the consecutive numbers of the classes. For example, 'A01' is the class code of class number one from Austria. Using a class code system enables a class-wise comparison of the data later. The students' class code was assessed using one item ('What is your class code?') and a free text field limited to three characters. Students were instructed which class code to use by their teachers who received the respective class codes via e-mail.

(12) **Feedback:** The students could express further comments in a free text field using one item ('Your opinion matters!').

6.1.2. Participants

I sent out an e-mail invitation to participate in the study to more than 100 teachers in Austria, Germany, and German-speaking Switzerland. As a result, the questionnaire was filled in by 62 school classes from different urban and rural areas of the three countries. The teachers who agreed to participate in the study with their classes were asked for some background information about their classes, namely number of students, grade, and whether they have physics classes in the current school year. The teachers were provided with an email invitation specific for each class and asked to send it to their students. In this email invitation, the students were provided with the class code and told that they will need to enter this class code at the end of the questionnaire. Moreover, the students received the link to access the questionnaire in the email. When clicking on the link the students were randomly assigned to one of the two versions of the questionnaire. The participating school classes were grade 9 (*N* = 48, 77%), grade 8 (*N* = 12, 19%), grade 10 (*N* = 1, 2%), and grade 8 and 9 mixed (*N* = 1, 2%). In total, 1219 students participated in the study. As described above, the students were provided with a class code. When filling in the questionnaire, some students misspelt their class code, for example 'DO6' instead of 'D06' or 'a21' instead of 'A21'. In case of misspelling, I corrected the class code and kept the data. Some students did not write an identifiable code, e.g. '819'. In such cases, I could correct the class code when the student started filling in the questionnaire at the same time as all students from a certain class (i.e. the teacher let the students fill in the questionnaire during class). This I also double checked with the background questions about the students' school type and grade. Yet, in five cases, I could not correct the class code. I considered these cases invalid and excluded them from the dataset (cases 954, 2665, 3081, 3141, and 3506). **Table 11** lists some descriptive statistics of the participants in the dataset. There was no missing data since all questions were marked as mandatory. According to their self-report, the study participants were 49% female (*N* = 595) and 44% male (*N* = 529); 7% of the students (*N* = 90) preferred not to specify their sex. 36% of the participating students were aged 14 years (*N* = 441), 47% aged 15 years (*N* = 573), and 10% aged 16 years (*N* = 115); 7% had a different age (*N* = 85).

Table 11. Descriptive statistics of the valid sample | *information provided by the teacher and associated with the class code

	Austria	Switzerland	Germany	Total
Sample size N	798	183	233	1214
Count (% of the sample)				
Grade*				
8	140 (18%)	111 (61%)	-	251 (21%)
9	626 (78%)	72 (39%)	233 (100%)	931 (77%)
mixed (8 and 9)	17 (2%)	-	-	17 (1%)
10	15 (2%)	-	-	15 (1%)
Physics class in current year*				
yes	659 (83%)	116 (63%)	233 (100%)	1008 (83%)
no	128 (16%)	67 (37%)	-	193 (16%)
mixed	11 (1%)	-	-	11 (1%)
Age				
14 years old	236 (30%)	80 (44%)	125 (54%)	441 (36%)
15 years old	403 (50.5%)	79 (43%)	91 (39%)	573 (47%)
16 years old	91 (11%)	14 (8%)	10 (4%)	115 (9%)
other	68 (8.5%)	10 (5%)	7 (3%)	85 (7%)
Sex				
female	368 (46%)	87 (48%)	140 (60%)	595 (49%)
male	373 (47%)	83 (45%)	73 (31%)	529 (44%)
prefer not to specify	57 (7%)	13 (7%)	20 (9%)	90 (7%)
Language spoken most at home				
German	554 (69%)	128 (70%)	189 (81%)	871 (72%)
another language	117 (15%)	34 (19%)	22 (9%)	173 (14%)
both equally	127 (16%)	21 (11%)	22 (9%)	170 (14%)

6.1.3. Preparation of the collected data

First, I prepared the collected data. I decided on excluding students who took less than three minutes to take the test. I argue that it is very unlikely that these students have read the instrument but ticked answers randomly instead. As a result, I excluded 27 students from the data. Thus, the sample size for analysis was *N* = 1187 students.

6.2. Quantitative and qualitative analysis of the collected data

The collected data was analysed quantitatively and qualitatively in seven main steps. Below I first describe these steps, and then I present the different analysis methods in more detail.

As a first step of data analysis, I needed to investigate to what extent the introductory text on the physics content area mechanics affects the students' expressed interest, when using the same items (RQ4, see **Section 4.3.**). I conducted separate mixed Rasch model analyses for the data concerning students' interest in mechanics collected with the two different versions of the questionnaire and checked whether the obtained groups (i.e. their item and person measures) differ in comparison to the ones obtained with the full data set, utilising independent-samples t-tests. Only if there were no statistically significant differences, I would be able to combine both datasets for further analyses.

Second, I investigated into which different types of interest in mechanics and in particle physics the students can be categorised (RQ5, see **Section 4.3.**), utilising mixed Rasch model analyses.

Third, I investigated to what extent the proposed hierarchy of students' levels of interest in particle physics can be applied to the students' types of interest in mechanics and in particle physics (RQ6, see **Section 4.3.**). I have introduced a conceptualisation of interest in particle physics comprising three levels of interest to describe the item hierarchy in my first study based on the context in which different physics contents may be set (see **Section 5.2.1.**). The three levels of interest in particle physics provide an overview of which contexts are more (or less) interesting relative to one another for the students. To investigate whether the item hierarchies of the students' types of interest obtained in the main study are also in line with the proposed hierarchy of levels of interest, I conducted a qualitative analysis of the item hierarchy of each type of interest; that is, I examined the item measures and wordings.

I also investigated whether the students are overall more interested in the content area 'particle physics' or 'mechanics' (RQ7, see **Section 4.3.**), utilising an analysis of the mean individual interest differences. I calculated the individual differences between the Rasch person measures about the students' interest in particle physics and mechanics. I subtracted the mechanics interest measures from the particle physics interest measures. Then, I calculated the mean of the total sample and the different interest types. Positive values would indicate that particle physics is on average more interesting whereas negative values would indicate that it is mechanics. I analysed whether this difference is statistically significant utilising a one-sample t-test.

In the fifth step, I investigated to what extent the students can be categorised into different types of physics-related self-concept (RQ8, see **Section 4.3.**), utilising mixed Rasch model analyses.

Then, I investigated to what extent physics-related self-concept is a better independent variable than sex for distinguishing between different types of interest in mechanics and in particle physics (RQ9, see **Section 4.3.**), utilising independent-samples t-tests, and linear and logistic regression model analyses.

Finally, I investigated to what extent the students' previous experience with the content areas in school affects their expressed interest in these content areas (RQ10, see **Section 4.3.**). First, I conducted a Rasch model analysis separately of the data regarding students' previous experience with mechanics and particle physics, respectively. I decided on utilising a Rasch model (in contrast to a mixed Rasch model) because the focus was on students' different degrees of previous experience with the content areas (in contrast to qualitative differences in their experience profiles). I used the obtained Rasch person measures regarding the students' previous experience with the content areas in school to conduct linear regression model/correlation analyses to investigate the association of experience with and interest in both content areas.

6.2.1. (Mixed) Rasch models

In this section I first summarise the features of Rasch models and then describe the additional features of mixed Rasch models.

I used **Rasch models** to analyse the collected data about students' (1) experience with mechanics and (2) experience with particle physics (RQ10; see **Section 4.3.**). Analysing data using a Rasch model facilitates the computation of linear measures for both persons and items. The person measure reflects the degree of the measured construct that a student has (Wright & Stone, 1979). For (1), that is the student's degree of previous experience with mechanics in school; and for (2), the student's degree of previous experience with particle physics in school. The higher the respective person measure, the higher the person's experience in mechanics (1) and particle physics (2). The item measure reflects for (1), how much previous experience in school students report regarding the mechanics content presented in the item; and for (2), how much previous experience in school students report regarding the particle physics content presented in the item. The lower the respective item measure, the more experience students report. Person and item measures are expressed on the same linear scale and in the unit of logits. To examine person and item measures Wright Maps can be created in Rasch analysis (Wright & Stone, 1979).

I conducted **mixed Rasch model** analyses to investigate the students' types of (3) interest in mechanics, (4) interest in particle physics, and (5) physics-related self-concept (RQs 4, 5, and 8; see **Section 4.3.**). That is, I examined whether students not only differ in their quantitative degrees of the measured constructs but also qualitatively in their answer profiles. If students had qualitatively different answer profiles, I could describe them in terms of different types. Within each type, students have similar answer profiles but different degrees of the measured construct. One aim of my study is to investigate these types with qualitatively different answer profiles and the students' different degrees of the measured construct within each type. To which type a student belongs, is a latent qualitative characteristic of the student. Thus, latent class analysis (LCA; Lazarsfeld & Henry, 1968) could be used to uncover the students' most probable type assignment. Yet, one assumption of LCA is that persons within a certain type all have the same degree of the measured construct. Hence, to additionally investigate the students' different degrees of the measured construct within the different types, another approach must be used that adds the above-described features of Rasch analysis. Mixed Rasch analysis combines the plus-sides of latent class analysis and Rasch analysis (Wright & Stone, 1979; Rost & von Davier, 1995). Hence, I used a mixed Rasch approach to analysing the collected data about students' (3) interest in mechanics, (4) interest in particle physics, and (5) physics-related self-concept.

In a mixed Rasch model, persons are categorised into different classes based on their most probable class assignment like in a latent class analysis. In my study, the classes represent groups of students with qualitatively different answer profiles.[14] For (3), the answer profiles show the mechanics interest profiles, that is, how interesting different items on mechanics are relative to each other; and for (4), the particle physics interest profiles, that is, how interesting different items on particle physics are relative to each other. Within a group of students, the same items are more (or less) interesting relative to each other for the students. For (5), the answer profiles show the physics-related self-concept profiles, that is, how easy to agree with different items on a student's physics-related self-concept are relative to each other. Within a group of students, the same items on physics-related self-concept are more (or less) easy to agree with relative to each other for the students. Within each group of students, a Rasch model is used to describe students' degrees regarding the measured construct as described above. I evaluated and compared the students' and the items' characteristics within and between the different groups of students based on the Wright Maps. For example, for (4), I evaluated and compared the students' different degrees of interest in particle physics and the interestingness of the different items on particle physics relative to each other within and between the different groups.

In sum, there are three principles underlying mixed Rasch analysis (Wright & Stone, 1979; Rost & von Davier, 1995):

1) The sample consists of a mix of several groups of students with qualitatively different answer profiles.
2) The students are categorised into one of these groups based on their most probable group assignment, as in latent class analysis.
3) Within each group, the Rasch model is applied. Hence, each student is characterised by a quantitative person measure, that is, a measure representing the student's degree of the measured construct. Students may differ in their person measures.

There are several mixed Rasch models, and the choice of the right model needs to be based on the nature of the data. Since my measurement instrument consisted of rating scale items, I used the mixed Rasch rating scale model. To understand this model one needs to understand the concept of threshold measures. A threshold measure marks the threshold between one rating scale category to another. For example, threshold 1 is between the categories '*very low*' and '*low*'. Since there are five rating scale categories, there are four thresholds, and hence four threshold measures for each item. In the mixed Rasch rating scale model the differences between the threshold measures are the same across all items within a group but can differ for different threshold. For example, the difference between the threshold measures 1 and 2 is the same for all items within a group. Likewise, the difference between the threshold measures 2 and 3 is the same for all items within the group, but this difference can differ from the difference between the threshold measures 1 and 2. The mixed Rasch rating scale model can be obtained by restricting the mixed Rasch general ordinal model. The mixed Rasch general ordinal model can be expressed with the equation 1[15] (Rost & von Davier, 1995).

---

[14] The term 'group' does not refer to a 'type' (as mentioned in RQ5 and 8). A group may indeed form a type or several groups combined may together form a type.

[15] Σ denotes a summation and Π denotes a product.

$$P(\boldsymbol{X_v} = \boldsymbol{x}) = \sum_{c=1}^{C} \pi_c \prod_{i=1}^{k} \frac{\exp\left(x_i \theta_{vc} - \sum_{l=1}^{x_i} \tau_{ilc}\right)}{\sum_{h=0}^{m} \exp\left(h\theta_{vc} - \sum_{l=1}^{h} \tau_{ilc}\right)} \tag{1}$$

with $\sum_{l=1}^{0} \tau_{ilc} = 0$. v is the person number and n is the total number of persons; i is the item number and k is the total number of items; c is the class number and C is the total number of classes. l is the threshold number, m is the total number of thresholds, $h \in \{1, \dots, m\}$ is the number of passed thresholds. The number of response categories is $m + 1$. A person's response to item I_i is $x_i \in \{0, \dots, m\}$, the response vector to all items is $\boldsymbol{x} = (x_1, \dots, x_k)$, and $\boldsymbol{X_v}$ is the response vector of a person. Equation 1 describes the probability of a person's observed response vector. The threshold measures are τ_{ilc}. The threshold measures within a latent class obey the normalisation condition that they sum up to zero; that is,

$$\sum_{i=1}^{k} \sum_{h=1}^{m} \tau_{ihc} = 0 \tag{2}$$

for all c because the item measure mean is set to zero within a latent class (Rost & von Davier, 1995). The measures of all passed thresholds can be summed up to yield the measure of a certain response category. The person measures are θ_{vc}. The threshold and person measures are class-specific, and thus are estimated for each latent class. The π_c parameterise the class sizes, and thus sum up to one; that is,

$$\sum_{c} \pi_c = 1 \tag{3}$$

By restricting the τ_{ihc} of the mixed Rasch general ordinal model (equation 1), one obtains sub models, such as the mixed Rasch rating scale model. The mixed Rasch rating scale model results from restricting the threshold measures; that is,

$$\tau_{ihc} = \sigma_{ic} + \tau_{hc} \tag{4}$$

with $\sum_h \tau_{hc} = 0$ for all c (Rost & von Davier, 1995). The item measure of item I_i in class c is σ_{ic}.

The Rasch person and item measures are estimated using a conditional maximum likelihood (CML) approach. The observed data are seen as function of unknown Rasch measures. The measures are estimated using the estimation-maximisation (EM) algorithm (Dempster et al., 1977; von Davier, 2001). The measures are first estimated (E-step). Then, the aim is to find the maximum of likelihood (explained below) of having generated the observed data for the estimated measures (M-step). Since the item and person measures are not estimated simultaneously, they are not biased. First, the item measures are estimated. The estimated item measures are then used to estimate the person measures. The Rasch person measures are based on the raw scores. Thus, within a group all persons who achieved the same raw scores also have the same person measures. The person measures used in my study are weighted likelihood estimates (WLE; von Davier, 2001; Warm, 1989) as used also in the PISA studies (OECD, 2007a, 2016).

The likelihood function *L* (i.e. the probability of the complete persons times items data matrix) is the product of the individual probabilities for each person and each item. Since the values of the likelihood function are usually small, the natural logarithm of the likelihood (ln*L*) is used instead. The model describes the data the better, the higher the logarithmic likelihood (*LL*[16]) is. It is important to note that less restrictive models (e.g. a model with more latent classes) usually have higher logarithmic likelihoods.

Hence, as a criterion for model validity, I combine qualitative and quantitative methods. Qualitatively, I analyse whether the retrieved model (e.g. a model consisting of three different groups of students with different particle physics interest profiles) is interpretable in a sound manner. Quantitatively, I use the Bayesian information criterion (BIC) value since previous studies suggest that it is the most reliable when conducting mixed Rasch analysis (Häußler, Lehrke et al., 1998; Preinerstorfer & Formann, 2012; Quandt, 2012; Sen, 2018; Sievers, 1999). The BIC value is calculated from the maximum likelihood *L* of the data with sample size *N* and number of measures to be estimated *k* (equation 5).

$$BIC = -2\ln L + k\ \ln N = -2LL + k\ \ln N \tag{5}$$

The BIC value approximates the Bayes factor for large sample sizes. Hence, I follow Kass and Raftery's (1995) guideline for the Bayes factor; that is, a smaller BIC value indicates that the corresponding model describes the collected data better than other models with higher BIC values. With reference to the lowest BIC value delta-BIC values are calculated. I consider delta-BIC values above 150 to indicate very strong evidence in favour of the model with smaller BIC value, above 20 to 150 strong evidence, and above 3 to 20 positive evidence (Kass & Raftery, 1995).

As outlined above, I conducted the mixed Rasch analysis separately for the data regarding students' (3) interest in mechanics, (4) interest in particle physics, and (5) physics-related self-concept. I calculated different models for each dataset, that is, models with different numbers of groups. In particular, I calculated four different mixed Rasch models for each dataset, from a model with only one group of students, that is, the sample is not divided into groups of students with different answer profiles, to a model with four different groups of students. I did not consider models with more than four groups, since (a) models with relatively more groups are less restrictive and usually have higher logarithmic likelihoods, that is, lower BIC values; (b) my hypotheses are that there are three different types of interest and no different types of self-concept; and (c) having more than four groups would result in relatively small percentual sizes, which reduces the usefulness of the description of types for educational practice.

6.2.2. Evaluation of the obtained Rasch person and item measures

After calculating different (mixed) Rasch models and choosing the best fitting models, the obtained Rasch measures were evaluated. The Rasch measures were evaluated individually for all groups of all chosen models. First, I evaluated the item measures. This is crucial for answering my research questions since (a) the item fit provides further evidence for the quality of the chosen model and (b) the item measures are used in the further analyses. The fit statistics are based on the difference between what is observed and what is expected by the Rasch model. The fit of the items within each

---

[16] The logarithmic likelihood *LL* is the logarithm of the likelihood ln*L*.

group obtained from the mixed Rasch analysis was evaluated using the *Q* index, which is the quotient of two conditional log-likelihood ratios (Rost & von Davier, 1994)[17]. The *Q* index ranges from 0 to 1. A *Q* index of 0 indicates perfect fit, 0.5 indicates independence of the trait and the item (i.e. random response behaviour), and 1 indicates perfect misfit (Rost & von Davier, 1994, 1995). The standardisation of the *Q* index, *ZQ*, can be assumed to be asymptotically normally distributed with a zero mean. To check for item fit, I used the items' *ZQ* values and apply a 2-sigma (95%) confidence interval; that is, I considered items with *ZQ* values ranging from -1.96 to 1.96 as having a good fit.

The Rasch person measures were evaluated in terms of Rasch person fit and multivariate outliers. The Rasch person fit was investigated using the persons' newfit values provided as output by the software 'WINMIRA'. According to the WINMIRA manual the newfit values are 'almost normally distributed' (von Davier, 2001). I excluded persons whose newfit 1 and/or 2 value is $\ll -3$ or $\gg 3$ following the rule-of-thumb guideline provided by Boone et al. (2014). I also excluded persons for who no newfit values were calculated because their person measures are extreme values. To check for multivariate outliers, I performed a linear regression analysis to model the relationships between the dependent variables (interest in mechanics, interest in particle physics, physics-related self-concept; RQ9, see **Section 4.3.**). Since the independent variable is not important for this outlier analysis, I used the students' case ID as independent variable. The regression model was calculated on the Rasch person measures for these variables. Consequently, the resulting regression coefficients were in logit units. Based on the recommendations by Tabachnick and Fidell (2013) to determine the suitability of data for linear regressions I excluded multivariate outliers using the Mahalanobis distance (Mahalanobis, 1936). The Mahalanobis distance is a measure for the distance between one datapoint and the point of the mean of the multivariate dataset. I compared the Mahalanobis distances to a chi-square distribution with the same degrees of freedom, that is, with the number of variables used to calculate the Mahalanobis distances, which is three in my case. I calculated the probability that a value from the chi-square distribution with three degrees of freedom will be less than the Mahalanobis distance of each person. Multivariate outliers are indicated by probability values less than 0.001. Following this cut-off value, the data of these persons were excluded from the further analyses. Moreover, I excluded the data of the students who did not indicate their sex (option: 'prefer not to say') because the reasons why they did not indicate their sex are unknown. For example, students might think that their sex does not matter, or they might not see themselves as either female or male. The reduced sample size following these exclusions is referred to as N_{red}.

6.2.3. Mean, Standard Deviation, and Standard Error

I calculated the mean, standard deviation, and standard error values using the Rasch person measures. In Rasch analysis, the person measures are scaled with respect to the item measures whose mean is set to 0 by default. Person and item measures are expressed on the same linear scale and in the unit of logits. For these calculations I used the reduced dataset with sample size N_{red}; that is, non-fitting persons, multivariate outliers, and persons who did not indicate their sex, were excluded. The mean $\overline{x}$ is the average Rasch person measure of all participants in the reduced dataset for a certain variable x (e.g. the students' degrees of interest in particle physics). The standard deviation s_x is a

---

[17] This is a different fit index as used in the first study. When developing the IPPI, I used a different software (Winsteps) that can only calculate Rasch models (not mixed Rasch models). This software provides MNSQ values for item fit. Rost and von Davier (1994) discuss different fit indices and how they are calculated.

measure for how variable the data for a certain variable *x* of the sample[18] are around the sample mean $\overline{x}$. The standard error SE_x is calculated as the sample's standard deviation divided by the square root of the respective sample size; that is,

$$SE_x = \frac{s_x}{\sqrt{N}} \tag{6}$$

6.2.4. Statistical significance and p-value

In significance testing, the *p-value* is the probability of observing an outcome that is at least as extreme as the outcome actually observed (e.g. interest and self-concept are correlated) under the assumption that the null hypothesis applies. The null hypothesis states that there is no difference or relationship (e.g. interest and self-concept are not correlated). In this study, a p-value of 0.05, that is, a probability of 5% that an outcome was observed under the null hypothesis, was used as α-level. Only if the calculated p-value was smaller than this α-level, the observed outcome was considered to be statistically significant. When multiple significance tests are performed, the α-level is inflated; that is, it becomes more likely that one outcome is significant although it is not. Here, corrections, such as a Bonferroni correction, can be applied (Hox et al., 2017). However, I chose not to apply a correction and instead interpret results with great care.

6.2.5. One-sample t-test

When comparing the mean for a certain variable $\bar{x}$ (e.g. the mean difference in Rasch person measures between interest in particle physics minus mechanics) to an a priori value of the mean μ (e.g. 0, indicating no mean difference in interest), a *one-sample t-test* may be used. A one-sample t-test provides evidence whether the difference between the mean and the a priori value of the mean is statistically significant. The null hypothesis is that the means are equal. For example, I used a *one-sample t-test* to analyse the mean individual differences between mechanics and particle physics interest (RQ7, see **Section 4.3.**).

The basis of a t-test is the *t-distribution*. The t-distribution is a bell-shaped function differing from the standard normal distribution in height and width depending on the number of degrees of freedom (e.g. sample size). Yet, for large degrees of freedom (e.g. sample sizes above 30) the t-distribution more and more looks like the standard normal distribution. From the t-distributions one can tell the *t-value.* In a t-test the t-value defined by the degrees of freedom is investigated. The t-value corresponding to the null hypothesis is 0. t-values significantly different from 0, that is, above (right hand side of the distribution) or below (left hand side of the distribution) a certain critical value, indicate that the null hypothesis is rejected in favour of the alternative hypothesis.

If $\bar{x}$ is the mean, s_x is the corresponding standard deviation, SE_x is the corresponding standard error, μ is the a priori value of the mean, and N is the sample size, the t-value can be calculated as follows in a one-sample t-test:

---

[18] A distinction has to be made between the *sample* that is participating in the research and the *population* with respect to which one wants to draw conclusions. In equations this is done for certain variables, such as the standard deviation, by using the Latin letter (s) to indicate a reference to the sample and the Greek letter (σ) to indicate a reference to the population.

$$t = \frac{\bar{x} - \mu}{\frac{s_x}{\sqrt{N}}} = \frac{\bar{x} - \mu}{SE_x} \tag{7}$$

Some assumptions need to be met for conducting a one-sample t-test, namely the data are normally distributed, independent, continuous, and obtained via a simple random sample.

6.2.6. Independent-samples t-test

When comparing the means for a certain variable *x* between two groups (e.g. groups who were presented with (1) the original IPN or (2) the newly developed introductory text), *independent-samples t-tests* may be used (also referred to as *unpaired-samples t-test*). An independent-samples t-test provides evidence whether the difference between the group means is statistically significant. For example, I used an independent-samples t-test to analyse the mean interest in mechanics of the two groups who were presented with (1) the original IPN or (2) the newly developed introductory text (RQ4, see **Section 4.3.**).

If $\overline{x}_1$ and $\overline{x}_2$ are the means of the two groups, $s_{x,1}$ and $s_{x,2}$ are the corresponding estimated standard deviances, ${s_{x,1}}^2$ and ${s_{x,2}}^2$ the corresponding variances, and $N_1 = N_2 = N$ is the group size of both groups, the t-value can also be calculated as follows:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{{s_{x,1}}^2 + {s_{x,2}}^2}{2}}\sqrt{\frac{2}{N}}} \tag{8}$$

The term $\sqrt{\frac{{s_{x,1}}^2 + {s_{x,2}}^2}{2}}$ is referred to as pooled standard deviation.

Some assumptions need to be met for conducting an independent-samples t-test, namely the two sample sizes are equal, and the data have equal variances, are normally distributed, independent, continuous, and obtained via a simple random sample. Yet, t-tests are quite robust to deviations from these criteria if the sample size is big enough and relatively similar for the two groups (Muijs, 2004). A *Levene's test* provides evidence whether the difference between the variances of the two groups is statistically significant (Levene, 1961). The null hypothesis is that the groups have equal variances. For equal variances but unequal sample sizes the t-value is calculated as follows:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{(N_1 - 1){s_{x,1}}^2 + (N_2 - 1){s_{x,2}}^2}{N_1 + N_2 - 2}}\sqrt{\frac{1}{N_1} + \frac{1}{N_2}}} \tag{9}$$

The term $\sqrt{\frac{(N_1 - 1){s_{x,1}}^2 + (N_2 - 1){s_{x,2}}^2}{N_1 + N_2 - 2}}$ is the pooled standard deviation for unequal group sizes.

6.2.7. Cohen's d_C

If there is a statistically significant difference between the mean and an a priori value of the mean (one-sample t-test) or between the group means (independent-samples t-test), effect sizes indicate the strength of the difference. Effect sizes are commonly reported as Cohen's d_C (Cohen, 1988; equation 10).

$$d_C = \frac{\overline{x}_1 - \overline{x}_2}{\sqrt{\frac{{s_{x,1}}^2 + {s_{x,2}}^2}{2}}} \tag{10}$$

This equation (10) is similar to the ones used for calculation the t-value (equations 7, 8, and 9). Indeed, it is possible to calculate Cohen's d_C from the t-value. The above-described adaption of the pooled standard deviation for unequal sample sizes applies in both cases. Cohen (1988) defines the effect size of the group differences as follows:

- Small effect: $d_C = 0.2 - 0.4$
- Medium effect: $d_C = 0.5 - 0.7$
- Large effect: $d_C \gg 0.8$

In significance testing, the effect size is usually reported together with the respective confidence interval (Cohen, 1988; Wilkinson & American Psychological Association Task Force on Statistical Inference, 1999). Originally, Cohen (1988) defined the effect size as the combined area covered by two equal-sized equally varying normal distributions that is not overlapping. Grice and Barrett (2014) point towards drawbacks of this definition based on the non-overlapping area. For example, following Cohen (1988), $d_C = 0.5$ means that 33% of the area is not overlapping but when looking at the data only 20% of data points are in the non-overlapping area (Grice and Barret, 2014).

6.2.8. Pearson correlation coefficient

The *Pearson correlation coefficient* has several names, such as the *Pearson's r*, the *bivariate correlation* in SPSS, or simply the *correlation coefficient* in research papers. The Pearson correlation coefficient r provides evidence to what extent the two variables x and y are linearly correlated in terms of direction and strength. It ranges from -1 (perfect negative linear correlation) to 1 (perfect positive linear correlation). It is calculated as the ratio between the covariance of the two variables and the product of their standard deviations. Hence, the Pearson's r is basically a standardisation of the covariance. The covariance can be calculated for interval scaled variables (e.g. Rasch person measures in logit units). The standardisation of the covariance to obtain the Pearson's r is necessary since otherwise two different covariances cannot be compared with each other.

If x and y are two different interval scaled variables with mean values $\overline{x}$ and $\overline{y}$, x_i and y_i are the corresponding values of a person *i*, and *N* is the sample size, the *covariance* is calculated as follows:

$$cov(x,y) = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{N-1} \tag{11}$$

If x and y are two different interval scaled variables and s_x and s_y are the corresponding estimated variances, the *Pearson r* is calculated as follows:

$$r = \frac{cov(x, y)}{s_x \cdot s_y} \tag{12}$$

6.2.9. Linear Regression Models

I investigated the relationship between interest and self-concept (RQ9, see **Section 4.3.**) as well as interest and previous experience with the content areas in school (RQ10, see **Section 4.3.**) utilising linear regression models on the person measures obtained from the mixed Rasch analysis. To describe linear relationships between two interval scaled variables, a linear regression model can be used, that is, graphically spoken, fitting a regression line in the scatterplot of the data. The aim is that the sum of the squares of the residuals, that is, the differences between the predicted and the observed y-values, is minimal. This is achieved by calculating the slope of the regression line (i.e. the regression coefficient; equation 13).

$$b = \frac{cov(x, y)}{{s_x}^2} \tag{13}$$

To ensure comparability of the regression coefficients across different models, the *standardised regression coefficient* β (also referred to as *standardised regression weight, beta weight, beta coefficient)* is calculated using the standard deviations of the variables (equation 14).

$$\beta = b\,\frac{s_x}{s_y} \tag{14}$$

To investigate to what extent the predicted and the observed values agree, the coefficient of determination *R²* is calculated using the variances of the predicted and the observed values (equation 15).

$$R^2 = \frac{{s_y}^2_{predicted}}{{s_y}^2} \tag{15}$$

The coefficient of determination ranges from 0 to 1 and indicates to what extent the variance of *y* is correctly predicted by the regression model. Commonly, it is analysed whether the coefficient of determination differs significantly from 0 using the p-value. The rule-of-thumb guideline for interpreting R^2 is that the values 0-0.1 indicate poor improvement in fit over the baseline model, 0.1-0.3 modest improvement, 0.3-0.5 moderate improvement, and >0.5 strong improvement (Muijs, 2004). Muijs (2004) suggest using the *'Adjusted R^2'*, a downwards adjustment of R^2 that takes into account that the model is likely to fit the population less well than the sample.

6.2.10. Logistic Regression Models

I investigated to what extent the interest types are described better, when using self-concept instead of sex as an independent variable (RQ9, see **Section 4.3.**) utilising logistic regression model analyses. In particular, I calculated different logistic regression models using the results from the mixed

Rasch analysis. This enabled me to investigate the correlational relationship between the independent variables (e.g. students' characteristics) and the likelihood of the event (e.g. students being assigned to a certain type). Logistic regression allows for the dependent variable to be categorical and for the independent variables to be both continuous and categorical. In my study, the dependent variables were the student's most probable type assignment obtained from the mixed Rasch analysis for the data about students' (1) interest in mechanics and (2) interest in particle physics. The independent variables were the student's sex (categorical) and physics-related self-concept (continuous Rasch person measure and categorical most probable type assignment obtained from the mixed Rasch analysis). I tested several assumptions, such as adequacy of expected frequencies and absence of outliers in the solution, at the beginning of the analysis to ensure that the data is suitable for logistic regression based on the recommendations by Tabachnick and Fidell (2013).

I calculated several logistic regression models. I examined the likelihood ratios of the baseline model, that is, a model in which the independent variables are 0 (the constant-only model), and the different comparison models to find the models that are significantly better than the baseline model using logarithmic likelihood ratio tests[19]. The logarithmic likelihood ratio is calculated using equation 18.

$$Logarithmic\ Likelihood\ Ratio\ (LLR) = -2 * \ln\left(\frac{L_0}{L_c}\right) = -2 * (LL_0 - LL_c) \qquad (18)$$

L_0 is the likelihood of the baseline model (the constant-only model), LL_0 is the maximised logarithmic likelihood of the baseline model (the constant-only model); L_c is the likelihood of the comparison model; and LL_c is the maximised logarithmic likelihood of the comparison model. In a logarithmic likelihood ratio test, one analyses whether the logarithmic likelihood ratio is significantly different from 0. LLR-values significantly different from 0 indicate that the null hypothesis does not apply; that is, the comparison model is significantly better than the baseline model.

After choosing the best model, I interpreted the odds ratio as commonly done in logistic regression analysis (Niu, 2020). The odds are the ratio of an event occurring to the event not occurring. For example, the **odds of group assignment** are the ratio of the number of students being in type 2 to the number of students not being in type 2, that is, the students are in type 1 instead. The odds ratio is the ratio of the odds and is measured between two individuals differing by one unit on the independent variable. For example, the **odds ratio of group assignment per sex** is the ratio of the odds of male students' type assignment to the odds of female students' type assignment. The odds ratio indicates the direction of the relationship (Niu, 2020). The odds ratio can also be described in terms of probabilities; that is,

$$Odds\ Ratio\ (OR) = \frac{p(x+1)/(1-p(x+1))}{p(x)/(1-p(x))} \qquad (16)$$

where *p* is the probability that the event occurs (e.g. a student being in type 2), *x* is the independent variable (e.g. students' sex with reference category female), and *x+1* is one-unit change in the independent variable (e.g. from female to male). Often, researchers erroneously report the odds ratio as ratio of probabilities (described below; Niu, 2020).

---

[19] They are referred to as chi-square tests in the statistical software 'SPSS'.

The significance of the correlational relationships between the different independent variables and the likelihood of the event was tested by calculating the p-values of the different logistic regression models. Most statistical software provides the odds ratio and the p-value as the standard output of a logistic regression analysis.

After determining the direction and the significance of the relationships, I examined the strength of statistically significant correlational relationships in terms of probabilities (Niu, 2020). The probability is the ratio of a certain event occurring to the total number of events. For example, the **probability of type assignment** is the ratio of the number of students being in type 2 to the total number of students. The ratio of probabilities is also known as relative risk. For example, the **relative risk of type assignment per sex** is the ratio of the probability of male students' type assignment to the probability of female students' type assignment. The strength of a correlational relationship can be investigated by calculating the relative risk (Niu, 2020; equation 17).

$$Relative\ Risk\ (RR) = OR - (OR - 1) * p(x+1) \tag{17}$$

where *p(x+1)* is the probability that the event (e.g. a student being in type 2) occurs for the comparison group (e.g. male students). First, I manually calculated the probability *p(x+1)* for the categorical variables.[20] Then, I manually calculated the relative risk for the best fitting logistic regression model using the odds ratios obtained from the statistical software.

To investigate to what extent the values predicted by the regression model and the observed values agree, I used the *Nagelkerke's* R^2. It is also referred to as *Pseudo* R^2 and ranges from 0 (indicating that the independent variables do not improve the model) to 1 (indicating the absence of prediction errors, that is, the perfect model). In addition, I analysed the number of correctly classified cases.

To find the independent variables that make the model significantly better compared to a model where the same variable is 0, I utilised the Wald statistics (equation 19; Field, 2009). The approach is equivalent to conducting a t-test in linear regression. In a Wald test, the *Wald-value* corresponding to the null hypothesis is 0 and Wald-values significantly different from 0 indicate that the null hypothesis does not apply; that is, the model is significantly better compared to a model where the same variable is 0.

$$Wald = \left(\frac{\ln(OR)}{SE_{\ln(OR)}}\right)^2 \tag{19}$$

6.2.11. Software

I used the software 'WINMIRA' to calculate the (mixed) Rasch models as well as the corresponding model, person, and item fit values. I also created the figures showing the Rasch answer profiles with WINMIRA. To conduct different statistical analyses of the obtained Rasch person measures, such as linear and logistic regression and t-tests, I used the software 'IBM SPSS statistics 28'. Finally, I used the software 'RStudio' for creating visualisations of the different analysis results, such as scatterplots and boxplots.

---

[20] This is to my knowledge not possible for continuous variables.

6.3. Results and discussion

The data collected with my measurement instrument was analysed in seven main steps (described in **Section 6.2.**). In the first step, I investigated to what extent the introductory text on the physics content area mechanics affects the students' expressed interest, when using the same items (RQ4, see **Section 4.3.**). Second, I investigated into which different types of interest in mechanics and in particle physics the students can be categorised (RQ5, see **Section 4.3.**). Third, I investigated to what extent students' types of interest can be described in terms of a hierarchy of students' levels of interest (RQ6, see **Section 4.3.**). I also investigated whether the students are overall more interested in the content area particle physics or mechanics (RQ7, see **Section 4.3.**). In the fifth step, I investigated to what extent the students can be categorised into different types of physics-related self-concept (RQ8, see **Section 4.3.**). Then, I investigated to what extent physics-related self-concept is a better independent variable than sex for distinguishing between different types of interest in mechanics and in particle physics (RQ9, see **Section 4.3.**). Finally, I investigated to what extent the students' previous experience with the content areas affects their expressed interest in these content areas (RQ10, see **Section 4.3.**). Below I present and discuss the results concerning the RQs 4 to 10 with respect to the respective hypotheses. Due to the high number of RQs, I decided to structure the results and discussion individually for each RQ. This approach makes it easier for readers to follow compared to first reporting the results concerning all RQs and then presenting the discussion concerning all RQs.

6.3.1. Influence of the introductory text on students' interest in mechanics

The results concerning my research question about the influence of the introductory text on the students' interest are presented and discussed in this section. My RQ4 asked *'To what extent does the introductory text on the physics content area mechanics affect students' expressed interest, when using the same items?'*. The original IPN introductory text presents mechanics mostly in contexts related to traffic and accidents, whereas the newly developed text presents mechanics in a broader variety of contexts. The hypothesis for this research question was that the students' expressed interest in mechanics changes, when using a different version of the introductory text in combination with the same items. This hypothesis was based on Schwarz' study (1999) which showed that the information provided in a questionnaire influences the participants' responses. It was shown that the framework of an item has a strong influence on the participants' response since they assume that the information provided in a questionnaire is relevant, and thus base their judgement primarily on it (Schwarz, 1999).

Results

To investigate to what extent the introductory text on mechanics affects the students' expressed interest, I separately analysed the data subsets collected with the instrument to measure mechanics interest comprising the (1) original IPN introductory text (*N* = 596) and (2) the newly developed introductory text (*N* = 591). For both datasets (1) and (2), I conducted a mixed Rasch analysis for four different models, from a model with only one group[21] of students, that is, the sample is not divided into groups of students with different interest profiles, to a model with four groups of students with different interest profiles. The delta-BIC values of the four calculated models with respect to the model with the lowest BIC value for both are listed in **Table 12**.

---

[21] The term 'group' does not refer to an 'interest type' (mentioned in RQ5). A group may indeed form an interest type or several groups combined may together form an interest type.

For both datasets, the BIC value of the 3-groups-model was the smallest. For dataset (1), the delta-BIC values indicated very strong evidence against the 1-group-model, strong evidence against the 2-groups-model, and positive evidence against the 4-groups-model. For dataset (2), the delta-BIC values indicated very strong evidence against the 1-group-model, and strong evidence against the 2- and 4-groups-models. For both datasets (1) and (2), the delta-BIC values provided evidence that a model comprising more than one group describes the data better than the 1-group-model. Hence, the sample consisted indeed of a mix of groups with different interest profiles which requires the application of the mixed Rasch model for analysis.

Table 12. Delta-BIC values of the four calculated models to describe the students' interest in mechanics. Data collected with the instrument to measure mechanics interest comprising the (1) original IPN introductory text and (2) newly developed introductory text.

	Delta-BIC			
Number of groups	1	2	3	4
(1) original IPN text	353	82	0	9
(2) newly developed text	249	31	0	85

The 3-groups-model described both datasets (1) and (2) the best. Consequently, I categorised the students into three different groups of interest in mechanics for both datasets (1) and (2). For dataset (1), group 1_M consisted of 45%, group 2_M of 33.5%, and group 3_M of 21.5% of the sample.[22] For dataset (2), group 1_M consisted of 43%, group 2_M of 41%, and group 3_M of 16% of the sample. **Figure 7** shows the answer profiles, that is, how interesting the different items are relative to each other, of all three groups 1_M, 2_M, and 3_M. In the answer profile figures, the items are shown on the x-axis and the items' threshold measures are shown on the y-axis. For each item four measures are given, each of which marks the threshold between one rating scale category to another. For example, threshold 1 is between the categories '*very low*' and '*low*'. Since there are five rating scale categories, there are four thresholds, and hence four threshold measures for each item. One can think of the thresholds as obstacles: The more interesting an item overall is, the lower are all the obstacles for the students. The line between the data points does not have a meaning but helps to observe some patterns.

For both datasets (1) and (2), one can see that the threshold measures of each item for group 1_M are more spread out and further away from each other than the threshold measures of each item for group 2_M. In addition, in group 1_M the distance from the first to the second threshold is bigger than the other threshold distances for both datasets (1) and (2). This is also the case for group 2_M in dataset (2). One can also see that apart from this difference in in-between-thresholds-distances the interest profiles, that is, how interesting the different items are relative to each other, are similar for both groups 1_M and 2_M for both datasets (1) and (2). For example, item M07 (artificial organs and joints in medicine) has the lowest threshold measures, whereas items M05 (calculating the energy of motion based on the speed of a car) and M10 (calculating the braking path based on the speed of a car; set in the context 'quantitative science') have very high threshold measures in both groups.

---

[22] The index M stands for Mechanics.

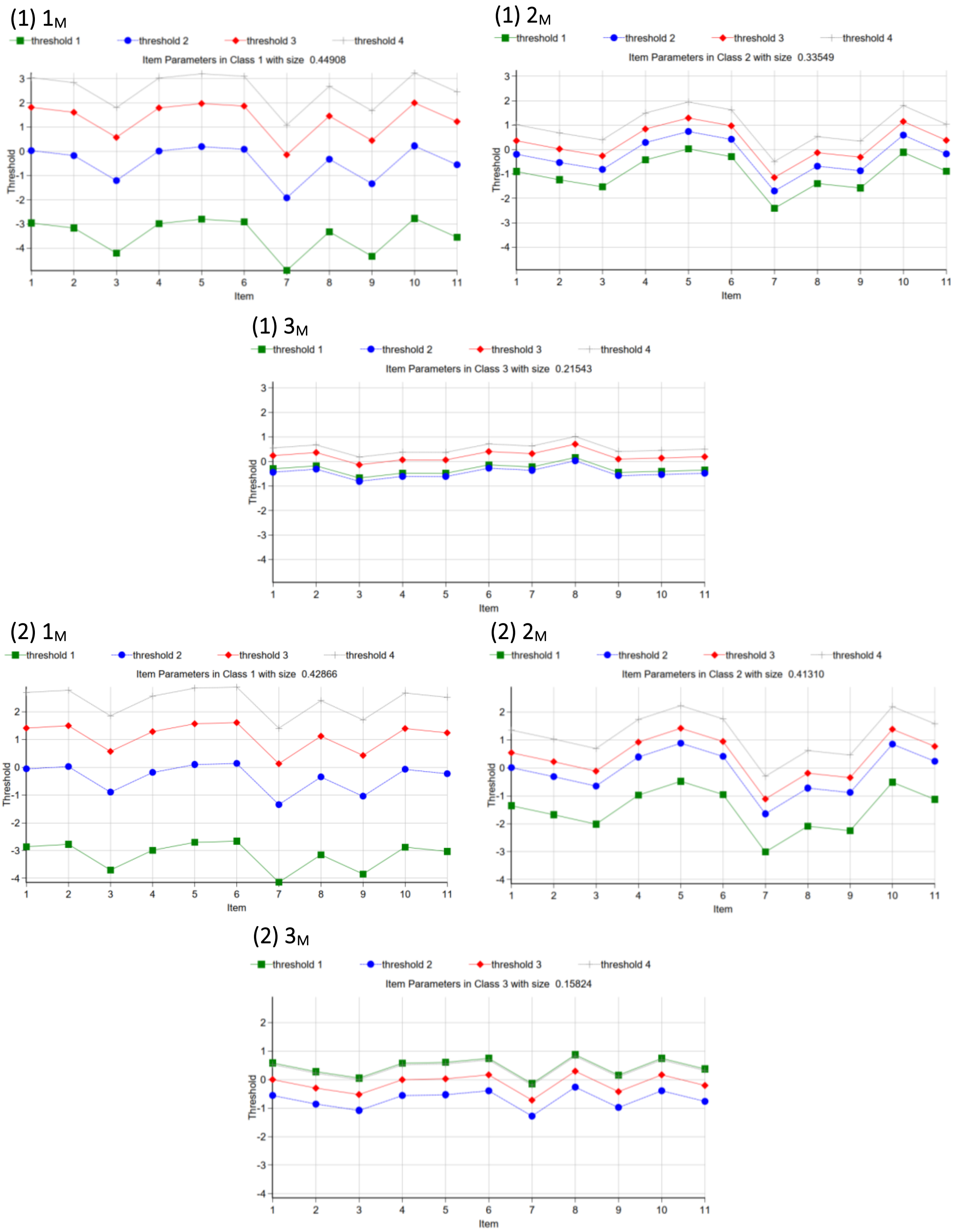

(1) 1_{M} (1) 2_{M} (1) 3_{M} (2) 1_{M} (2) 2_{M} (2) 3_{M}

Figure 7. Interest profiles, that is, how interesting the different items are relative to each other, of the three groups 1_{M}, 2_{M}, and 3_{M} for datasets (1) and (2).

Note: On the x-axis, the items are shown, and on the y-axis, the threshold measures are shown in logit units, each marking the threshold between one rating scale category to another. For example, threshold 1 is between the categories *‘very low’* and *‘low’*. The line between the data points does not have a meaning but helps to observe some patterns.

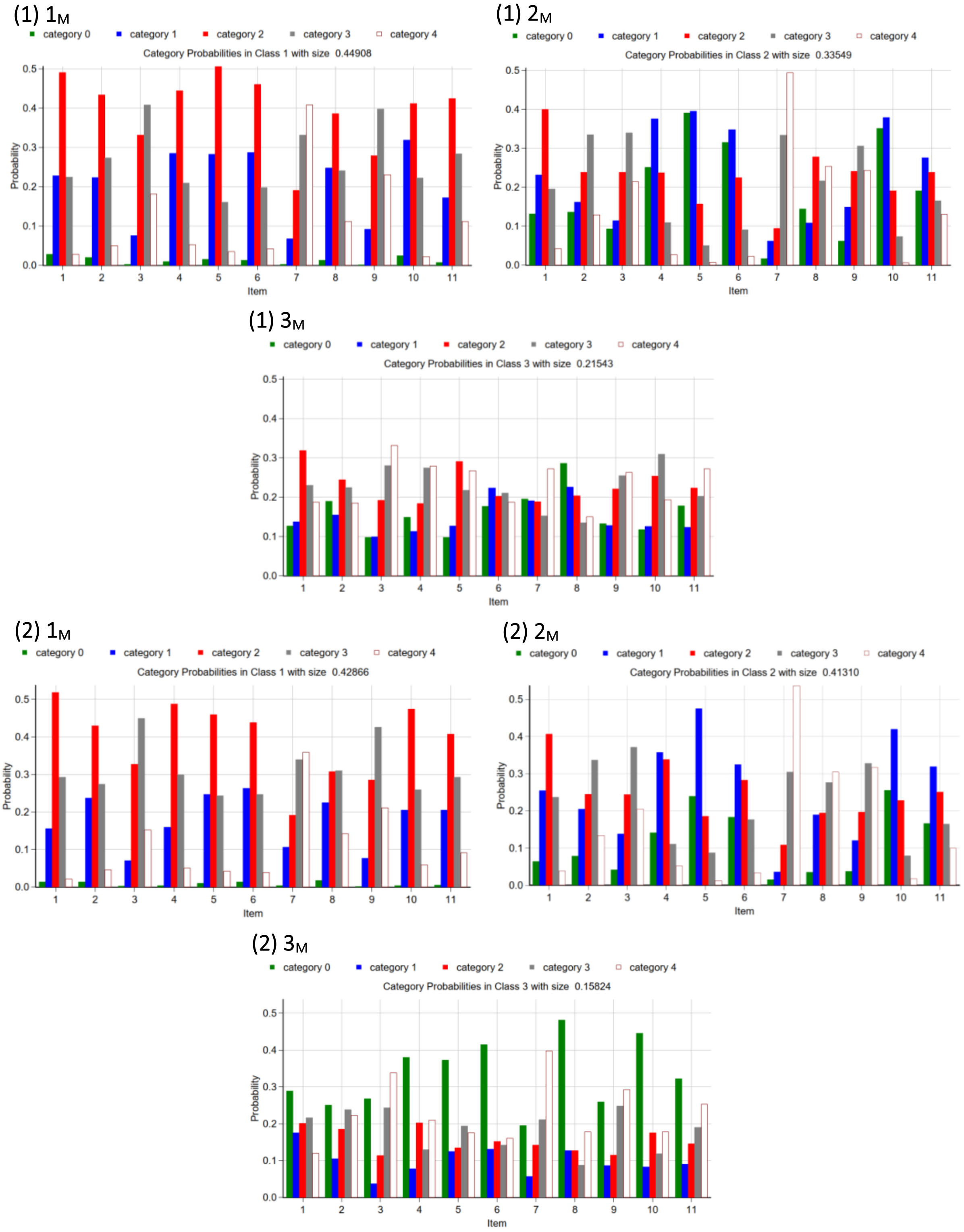

(1) 1_{M} (1) 2_{M} (1) 3_{M} (2) 1_{M} (2) 2_{M} (2) 3_{M}

Figure 8. Category probabilities, that is, the probabilities to choose the different categories of the rating scale, of the groups 1_{M}, 2_{M}, and 3_{M} for datasets (1) and (2).
Note: Dataset (1) was collected using the original IPN introductory text (N = 596) and dataset (2) using the newly developed introductory text (N = 591). The students expressed their degree of interest in each item on a 5-category rating scale (‘My interest in this is …’ *very high* (=4), *high* (=3), *medium* (=2), *low* (=1), or *very low* (=0))

Finally, one can see that the interest profile of group 3_M differs from the other groups for dataset (1) but is similar to the other groups for dataset (2). For dataset (1), in group 3_M the thresholds are not ordered and close together; the items are all similarly interesting relative to each other. For dataset (2), in group 3_M the thresholds are not ordered either; the relative interestingness of the different items is similar to groups 1_M and 2_M but the differences in interestingness are way less pronounced.

To study this difference between the different groups for both datasets (1) and (2), I analysed the students' use of the answer categories of the rating scale. **Figure 8** shows the category probabilities, that is, the probabilities to choose the different categories of the rating scale, of the groups 1_M, 2_M, and 3_M for both datasets (1) and (2). For both datasets (1) and (2), one can see that the groups differ from each other in their use of the rating scale categories. Group 1_M tended to use the non-extreme categories, especially the middle category, across the 11 items. For dataset (1), group 2_M tended to use all the rating scale categories, especially also the extreme categories depending on the item, whereas group 3_M tended to use all the rating scale categories rather equally across the 11 items. For dataset (2), group 2_M also tended to use the non-extreme categories but also the extreme categories depending on the item, whereas group 3_M tended to use the extreme categories across the 11 items. Comparing the group sizes and category probability distributions for both datasets (1) and (2), one can see that group 1_M is similarly sized and has a similar category probability distribution for both datasets (1) and (2). In contrast, for dataset (1) there were a lot of non-extreme responses in group 3_M, and hence the group size is bigger than group 3_M for dataset (2) where there were mostly negative extreme responses. This also resulted in the differences in group sizes and category probability distributions of group 2_M between the datasets (1) and (2).

The observations based on the answer profiles and category probability distributions were further supported when analysing the items based on their item measures, that is, their interestingness. **Table 13** lists the item hierarchies within the different groups for both datasets (1) and (2). The colour scale supported findings based on the answer profiles. First, one can see that for both datasets (1) and (2), the same items were more (or less) interesting relative to each other within the groups. However, for some groups the differences in interestingness were more pronounced than in other groups as can be seen using the colour scale. Moreover, the group 3_M in dataset (1) is an exception because all items were similarly interesting for the students, and hence there was not a distinct item hierarchy. Remembering that for dataset (1), group 3_M tended to use all the rating scale categories rather equally across the 11 items, the absence of a distinct hierarchy makes sense.

In Rasch analysis, the person measures are scaled with respect to the item measures whose mean is set to 0 by default. The three groups had different mean person measures, that is, degrees of interest as listed in **Table 14**. For both datasets, group 1_M had the highest mean interest; for dataset (1), group 3_M had a higher mean than group 2_M; and for dataset (2), group 2_M had a higher mean than group 3_M. This finding is not surprising, remembering the category probability distributions, especially that for group 3_M of dataset (2) there were mostly negative extreme responses.

Table 13. Item IDs, corresponding Rasch item measures (reflecting their interestingness), and SEs for the different groups and for both datasets (1) and (2).

	(1) original IPN introductory text						(2) new introductory text					
Item	Group 1_M	SE	Group 2_M	SE	Group 3_M	SE	Group 1_M	SE	Group 2_M	SE	Group 3_M	SE
M10	0.67	0.09	0.86	0.08	-0.08	0.08	0.28	0.08	0.98	0.08	0.31	0.10
M05	0.64	0.09	1.00	0.08	-0.16	0.09	0.45	0.09	1.02	0.08	0.16	0.10
M06	0.54	0.09	0.68	0.08	0.18	0.08	0.49	0.09	0.54	0.07	0.30	0.10
M01	0.49	0.09	0.07	0.07	0.02	0.08	0.30	0.08	0.14	0.07	0.14	0.10
M04	0.47	0.09	0.55	0.08	-0.16	0.09	0.17	0.08	0.52	0.07	0.13	0.10
M02	0.28	0.09	-0.26	0.07	0.14	0.08	0.38	0.08	-0.18	0.07	-0.16	0.10
M08	0.13	0.09	-0.42	0.07	0.48	0.08	0.01	0.08	-0.59	0.07	0.43	0.10
M11	-0.10	0.09	0.09	0.07	-0.03	0.08	0.13	0.08	0.37	0.07	-0.07	0.10
M03	-0.75	0.09	-0.55	0.07	-0.36	0.09	-0.54	0.08	-0.52	-0.07	-0.39	0.10
M09	-0.88	0.09	-0.60	0.07	-0.13	0.08	-0.69	0.08	-0.75	0.07	-0.29	0.10
M07	-1.47	0.09	-1.43	0.09	0.10	0.08	-0.99	0.09	-1.52	0.09	-0.58	0.11

Note: The Rasch item measures are listed in logit units. Lower and higher item measures represent more and less interesting items, respectively. Model SE refers to the standard error of the item measure in logit units. The mean item measure of each group is set to zero in mixed Rasch analysis. The items are ordered according to their measures in group 1_M of dataset (1). The colour scale represents the interestingness across all groups and both datasets; dark blue is the most interesting item and dark red is the least interesting item.

Table 14. Person measure means and standard deviations, minimum and maximum values for all groups 1_M, 2_M, and 3_M for both datasets (1; *N* = 596) and (2; *N* = 591)

Person Measures				
	Group	1_M	2_M	3_M
(1) original IPN text	Mean (Standard Deviation)	0.78 (0.99)	-0.17 (0.61)	0.18 (0.99)
	Min – Max	-2.55 – 3.46	-2.37 – 1.26	-3.39 – 3.64
(2) newly developed text	Mean (Standard Deviation)	0.84 (0.74)	0.08 (0.65)	-0.11 (1.03)
	Min – Max	-2.06 – 2.74	-1.59 – 2.09	-2.61 – 3.52

Note: The Rasch person measure means and standard deviations are listed in logit units. Lower and higher person measures represent less and more interested persons, respectively. The person measures are estimated based on the item measures in Rasch analysis.

In sum, the analysis of the answer profiles, category probability distributions, and mean person measures indicated that overall, for most of the students in both datasets the same items were more (or less) interesting relative to each other. However, there was some evidence that there is a group of students in dataset (1) who rated all the items similarly interesting. Since this group would still be distinguishable when analysing both datasets combined, I argue that for further analyses such a combination may be possible. I checked this by conducting a combined analysis and by comparing the results of the separate analyses and the combined analysis.

First, I combined both datasets (1) and (2). Then, I conducted the mixed Rasch analysis for four different models, from a model with only one group[23] of students, that is, the sample is not divided into groups of students with different interest profiles, to a model with four groups of students with different interest profiles. The delta-BIC values of the four calculated models with respect to the model with the lowest BIC value are listed in **Table 15**. The BIC value of the 4-groups-model was the smallest. The delta-BIC values indicated very strong evidence against the 1-group-model and the 2-groups-model and strong evidence against the 3-groups-model.

Table 15. Delta BIC-values of the four calculated models to describe the students' interest in mechanics

	Delta-BIC			
Number of groups	1	2	3	4
Delta-BIC	812	277	66	0

Qualitative analysis of the groups described by the different models showed that a distinction into four different groups is sufficient to describe the students' interest in mechanics. The distinction into more groups did not improve the description of students' interest because then the groups' sizes would be very small and/or several groups would have very similar interest profiles. Consequently, I categorised the students of this sample into four different groups of interest in mechanics. Group 1_M consisted of 37% of the sample, group 2_M of 31%, group 3_M of 18%, and group 4_M of 14%. **Figure 9** shows the interest profiles, that is, how interesting the different items are relative to each other, of all four groups 1_M, 2_M, 3_M, and 4_M.

One can see that the threshold measures of each item for group 1_M are more spread out and further away from each other than the threshold measures of each item for group 2_M, which in turn are further away from each other than the threshold measures of each item for group 3_M. In groups 1_M and 2_M the distance from the first to the second threshold is bigger than the other threshold distances. In group 3_M the thresholds are not ordered, which may be a result of low response frequencies in one or more rating scale categories (as discussed below). One can also see that apart from this difference in in-between-thresholds-distances the interest profiles, that is, how interesting the different items were relative to each other, are similar for all three groups 1_M, 2_M, and 3_M. For example, item M07 (artificial organs and joints in medicine) has the lowest threshold measures, whereas items M05 (calculating the energy of motion based on the speed of a car) and M10 (calculating the braking path based on the speed of a car; set in the context 'quantitative science') have relatively high threshold measures in all three groups. Finally, one can see that group 4_M differs in its interest profile from the other groups. The thresholds are close together. The items were all similarly interesting relative to each other, although item M07 (artificial organs and joints in medicine) was the least interesting relative to the others.

---

[23] The term 'group' does not refer to a 'type' (as mentioned in RQ5 and RQ8). A group may indeed form a type or several groups combined may together form a type.

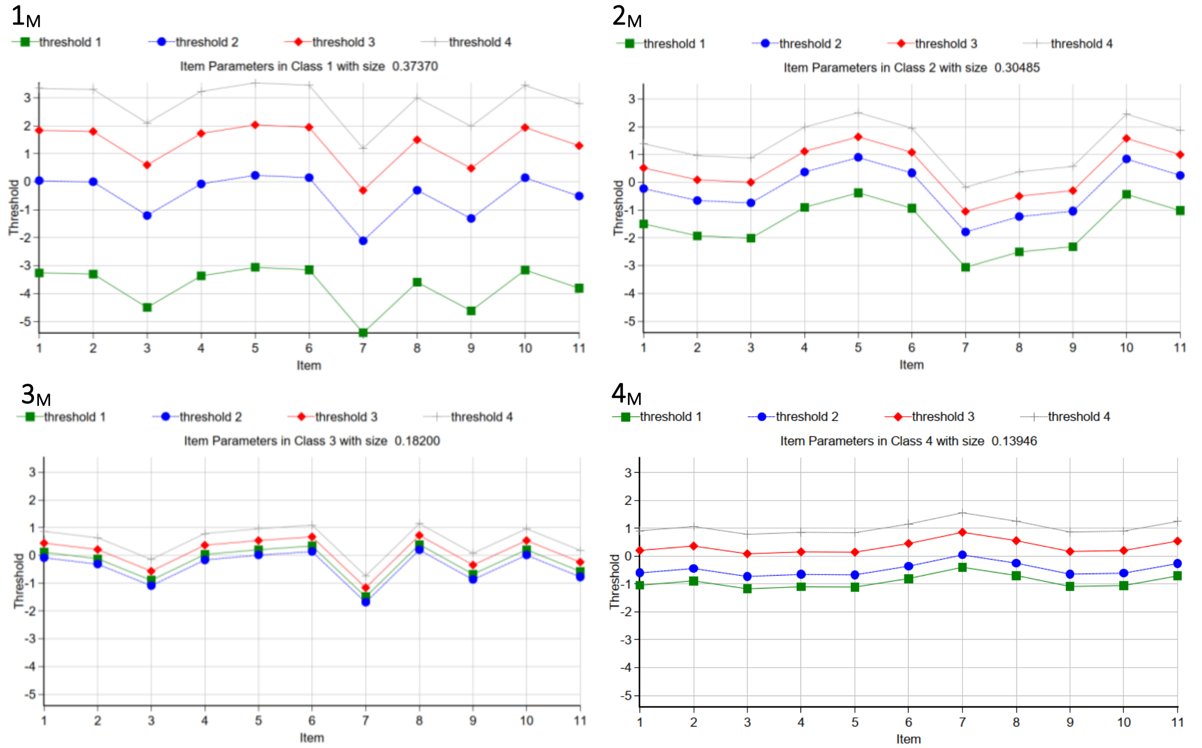


Figure 9. Interest profiles, that is, how interesting the different items are relative to each other, of all groups 1_M, 2_M, 3_M, and 4_M.

Note: On the x-axis, the items are shown, and on the y-axis, the threshold measures are shown in logit units, each marking the threshold between one rating scale category to another. For example, threshold 1 is between the categories *'very low'* and *'low'*. The line between the data points does not have a meaning but helps to observe some patterns.

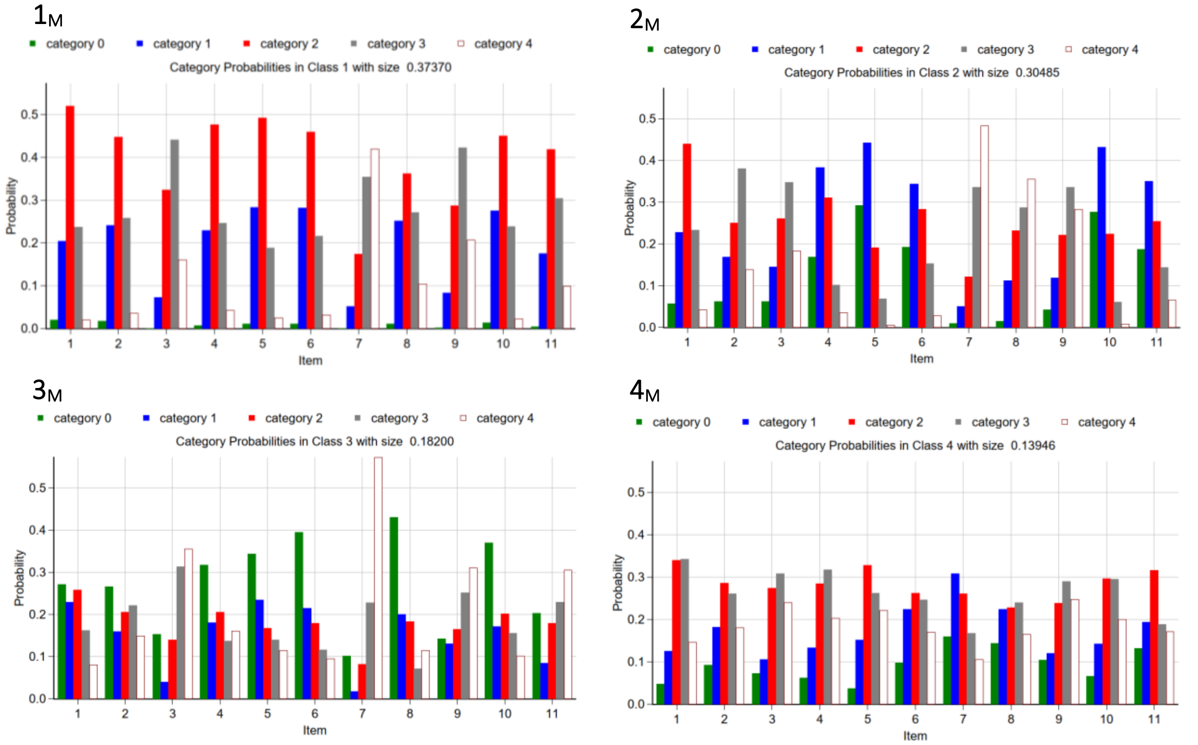


Figure 10. Category probabilities, that is, the probabilities to choose the different categories of the rating scale, of all groups 1_M, 2_M, 3_M, and 4_M.

Note: The students expressed their degree of interest in each item on a 5-category rating scale ('My interest in this is ...' *very high* (=4), *high* (=3), *medium* (=2), *low* (=1), or *very low* (=0))

To study the difference between the groups 1_M, 2_M, 3_M, and 4_M I analysed how the students used the different answer categories of the rating scale. **Figure 10** shows the category probabilities, that is, the probabilities to choose the different categories of the rating scale, of all groups. One can see that the groups differ from each other in their use of the rating scale categories. Group 1_M tended to use the non-extreme categories, especially the middle category and not the negative extreme categories, across the 11 items. Group 2_M also tended to use the non-extreme rating scale categories, but also the negative and positive extreme categories depending on the item. Group 3_M tended to use the negative and positive extreme rating scale categories depending on the item. Group 4_M tended to use all the rating scale categories rather equally across the 11 items.

The observations based on the answer profiles and category probability distributions were confirmed when analysing the items based on their item measures, that is, their interestingness. **Table 16** lists the items and their item measures for all groups 1_M, 2_M, 3_M, and 4_M. The lower the item measure of an item is, the more interesting was the item rated by the students. One can see that the hierarchy of items based on their interestingness is very similar for the groups 1_M, 2_M, and 3_M. Lower and higher item measures represent more and less interesting items, respectively.

Table 16. Item IDs, corresponding Rasch item measures (reflecting their interestingness), and standard errors (SE) for all groups 1_M, 2_M, 3_M, and 4_M.

Item	Group 1_M	SE	Group 2_M	SE	Group 3_M	SE	Group 4_M	SE
M05	0.68	0.07	1.17	0.07	0.43	0.07	-0.20	0.08
M10	0.60	0.07	1.11	0.07	0.43	0.07	-0.14	0.08
M06	0.60	0.07	0.61	0.06	0.56	0.07	0.11	0.08
M01	0.49	0.07	0.05	0.06	0.34	0.07	-0.13	0.08
M02	0.45	0.07	-0.38	0.06	0.11	0.07	0.03	0.08
M04	0.38	0.07	0.64	0.06	0.26	0.07	-0.18	0.08
M08	0.15	0.07	-0.97	0.06	0.62	0.09	0.22	0.08
M11	-0.06	0.07	0.53	0.06	-0.35	0.07	0.21	0.08
M03	-0.75	0.07	-0.47	0.06	-0.67	0.07	-0.26	0.08
M09	-0.87	0.07	-0.77	0.06	-0.45	0.07	-0.17	0.08
M07	-1.66	0.07	-1.52	0.07	-1.27	0.09	0.52	0.08

Note: The Rasch item measures and standard errors are listed in logit units. Lower and higher item measures represent more and less interesting items, respectively. The mean item measure of each group is set to zero in mixed Rasch analysis. The items are ordered according to their measures in group 1_M. The colour scale represents the interestingness across all groups and both datasets; dark blue is the most interesting item and dark red is the least interesting item. More detailed item tables and Wright Maps for each group are presented in the **Appendix 10.2.7.** and **10.2.8.** of this thesis.

To compare this combined analysis with the separate analysis of the datasets (1) and (2), I investigated to what extent the students have the same most probable group assignment in both analyses. **Table 17** lists the most probable group assignment in the combined analysis per most probable group assignment in the separate analyses.

Table 17. Most probable group assignments for all groups 1_M, 2_M, 3_M, and 4_M of the combined analysis and the groups 1_M, 2_M, and 3_M for datasets (1) and (2).

		Combined analysis			
		Count (row %, column %)			
	Group	1_M	2_M	3_M	4_M
(1) original IPN text	1_M	**238 (84.7%, 49.8%)**	25 (8.9%, 7,1%)	11 (3.9%, 5.5%)	7 (2.5%, 4.5%)
	2_M	1 (0.5%, 0.2%)	**135 (71.4%, 38.5%)**	52 (27.5%, 25.9%)	1 (0.5%, 0.6%)
	3_M	1 (0.8%, 0.2%)	-	**43 (34.1%, 21.4%)**	**82 (65.1%, 52.2%)**
(2) newly developed text	1_M	**218 (82.0%, 45,6%)**	5 (1.9%, 1.4%)	6 (2.3%, 3.0%)	37 (13.9%, 23.6%)
	2_M	17 (7.2%, 3.6%)	**183 (77.2%, 52,1%)**	25 (10.5%, 12.4%)	12 (5.1%, 7.6%)
	3_M	3 (3.4%, 0.6%)	3 (3.4%, 0.9%)	**64 (72.7%, 31.8%)**	**18 (20.5%, 11.5%)**

Table 17 shows that for the students with a most probable group assignment to groups 1_M and 2_M the assignment was the same in the combined and the separate analyses. Yet, for the students with a most probable group assignment to group 3_M in the separate analyses, their assignment in the combined analysis differs for the datasets (1) and (2). Whereas most students from group 3_M (dataset 1) were assigned to group 4_M (combined datasets), most students from group 3_M (dataset 2) were assigned to group 3_M (combined datasets). Since the group 3_M (dataset 1) differed in the item hierarchy from the other groups, it is not surprising that their most probable group assignment was group 4_M (combined datasets) which also differed in the item hierarchy from the other groups. Accordingly, group 4_M (combined datasets) mostly comprises students that were assigned to group 3_M in the separate analysis of dataset (1). In sum, both datasets (1) and (2) were best described by a 3-group model each. When combining these two datasets, the data was best described by a 4-group model because the overlap between the respective subgroups was not perfect.

So far, I compared the students' most probable group assignment obtained from the separate analyses (a) and the combined analysis (b) of the datasets. This means that my focus was on the qualitative aspect of the students' interest in mechanics. Next, I compared their quantitative degree of interest obtained from the two analyses.

Figure 11 shows the boxplots with jitter of the students' Rasch person measures reflecting their degrees of interest in mechanics comparing both versions of the questionnaire obtained from the separate analyses (a) and the combined analysis (b) of the datasets. One can see that both distributions are very similar. Yet, for the dataset (2), collected using the new introductory text, the range of person measures is wider and the distribution is more spread in the combined analysis than in the separate analysis.

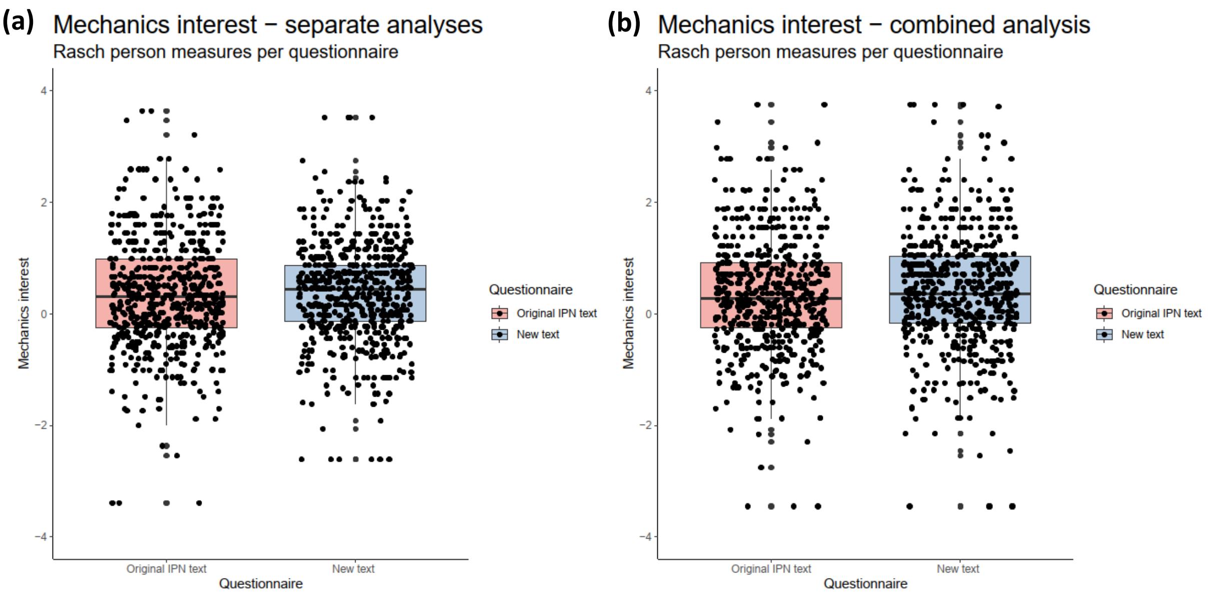


Figure 11. Boxplots with jitter of the students' Rasch person measures in logit units reflecting their degrees of interest in mechanics comparing both versions of the questionnaire obtained from separate analyses (a) and the combined analysis (b) of the datasets.

In addition, I performed an independent-samples t-test to check whether there is a significant difference in mean person measures reflecting the students' degrees of mechanics interest obtained from the separate analyses of both datasets between the students who were presented with the original IPN introductory text and the students presented with the newly developed introductory text. The Levene's test was not significant; that is, equal variances for both groups are assumed. I found that there was no statistically significant difference in mean mechanics interest (separate analyses) between the students who were presented with the IPN introductory text (0.35 ± 0.98 logits) compared to those presented with the new introductory text (0.40 ± 0.91 logits), t(1185)=-0.772, p=0.440.

I conducted the same t-test for the data obtained from the combined analysis of both datasets. The Levene's test was not significant, and hence equal variances are assumed. There was no statistically significant difference in mean mechanics interest (combined analysis) between the students who were presented with the IPN introductory text (0.35 ± 1.00 logits) compared to those presented with the new introductory text (0.41 ± 1.07 logits), t(1185)=1.01, p=0.311.

Discussion

The results presented in this section provide evidence that a combined analysis of both datasets (1) and (2) is possible. First, the t-tests provided evidence that there was no significant difference in mean interest between the students who were presented with the original IPN introductory text (1) and those who were presented with the newly developed introductory text (2) regardless of the analysis method (separate analyses vs. combined analysis). However, when calculating the mixed Rasch model separately for both datasets, the group 3_M for dataset (1) had a slightly differing item hierarchy compared to the other groups in both datasets. This group seems to be rather similarly interested in all items compared to the other groups in both datasets which all have a similar item hierarchy. Especially different is the relatively high interest in the items M05 (calculating the energy of motion based on the speed of a car) and M10 (calculating the braking path based on the speed of a car; both presenting science in the context ‘car accidents’) of the group 3_M for dataset (1). However, this can be explained by the specific rating scale use uncovered by the mixed Rasch models. For dataset (1) there were a lot of non-extreme responses in group 3_M, and hence the group size was bigger than group 3_M for dataset (2) where there were mostly negative extreme responses. Another possible additional explanation is that the students who were presented with the newly developed introductory text are aware of a broader range of possible contexts in which the mechanics contents may be set. Hence, they might have been more sensitive to the contexts presented in the items. In contrast, the group 3_M for dataset (1) may not have been aware that there is a broad range of possible contexts for mechanics, and hence these students showed a different item hierarchy. Another possible explanation is that because the original IPN text rather focuses on contexts related to traffic and accidents, the students would like to learn more about this, which resulted in the unexpected relatively higher interest in the items M05 (calculating the energy of motion based on the speed of a car) and M10 (calculating the braking path based on the speed of a car). Second, using a mixed Rasch approach the group 3_M (dataset 1) with a differing item hierarchy was still distinguishable in the combined analysis. In particular, most students assigned to the group 3_M (dataset 1) were assigned to the group 4_M in the combined analysis (see **Table 17**). The best fitting mixed Rasch rating scale model when conducting an analysis of the combined dataset identified four distinct mechanics interest groups.

In sum, three factors were relevant to investigate to what extent the introductory text on mechanics affects the students’ expressed interest. First, the analysis method based on different mixed Rasch rating scale models might have resulted in slightly different groups. Here, one reason is that the models for both datasets were based on roughly half the data in comparison to calculating models for the full data. One indication for the analysis method as influential factor was that the group 3_M is bigger for dataset (1) than (2). Second, the new introductory text (dataset 2) presented a broader variety of contexts, and hence the students might have been more ‘sensitive’ to the contexts mentioned in the items. Third, the original IPN introductory text (dataset 1) mostly focused on the contexts traffic and accidents, and hence the students’ interest in contents set in these contexts (items M05: calculating the energy of motion based on the speed of a car; M10: calculating the braking path based on the speed of a car) might have been triggered.

Overall, my study provided evidence that there was no quantitative difference in the students’ interest as indicated by the t-test. That is, I found that the mean interest in mechanics is independent from the used version of the introductory text. Hence, I rejected the hypothesis that the students’ expressed interest in mechanics changes, when using a different version of the introductory text in combination with the same items. However, my study provided some evidence that there are qualitative differences in interest, that is, certain items seem to have been relatively more interesting for some students in dataset (1). These slight qualitative differences in the item hierarchy could be explained by the different foci of the different versions of the introductory text, but also by the analysis method. If there really was a group of students with a differing item hierarchy, this group would also be distinguishable in the mixed Rasch rating scale model, when combining the datasets for analysis. Hence, I decided to combine both datasets for further analyses.

6.3.2. Students’ types of interest in mechanics and particle physics

The results concerning my research question about the students’ categorisation into different types of interest are presented and discussed in this section. RQ5 asked *‘Into which different types of interest in physics can German-speaking students aged 14 to 16 years be categorised, while comparing a classical and a modern physics content area (namely mechanics and particle physics)?’*. The hypothesis for this research question was that the interest types (A, C, NG) found in the IPN study are still valid for today’s students and for both content areas, mechanics and particle physics (Langeheine et al., 2001; Rost et al.,1999; Sievers, 1999). The first type (A) describes students that are generally and highly interested in physics; the second type (C) describes students that are only interests in physics when set in a context that relates to humans and nature, applications, and relevance for society; and the third type (NG) is similar to type A or type C depending on the content area. To answer this research question and describe interest types, I compared the item hierarchies of the different groups obtained from a mixed Rasch analysis for both content areas. Below I will present the results and discussion regarding the types of interest in mechanics and particle physics, as well as a comparison.

6.3.2.i. Mechanics

First, I investigated students’ types of interest in mechanics. To investigate the students’ types of interest in mechanics I combined both datasets (1) with the original introductory text and (2) with the newly developed introductory text.

Results

The results of the mixed Rasch analysis of the combined mechanics dataset were presented in the previous section (**Section 6.3.1**). The best fitting mixed Rasch rating scale model when conducting an analysis of the combined dataset identified four distinct mechanics interest groups[24].

When conducting (mixed) Rasch analyses, an evaluation of the obtained item and person measures is crucial as it provides further evidence for the quality of the chosen model. First, I evaluated the Rasch item measures. When checking for item fit, I found that all items fit well in all mechanics groups obtained in the combined analysis, that is, in group 1_M (*ZQ* values ranging from -1.01 to 1.43), in group 2_M (*ZQ* values ranging from -1.34 to 1.63), in group 3_M (*ZQ* values ranging from -1.26 to 1.78), and in group 4_M (*ZQ* values ranging from -0.88 to 1.30). Based on the item fit statistics, I argue that the data supports one trait, interest in mechanics, for all four groups, that is, that the IPN instrument to measure interest in mechanics is unidimensional. Second, I evaluated the Rasch person measures for the students’ interest in mechanics. I investigated the Rasch person fit and excluded those students from further analysis, whose newfit 1 and/or 2 value was $\ll -3$ or $\gg 3$ or for who no newfit values were calculated. Hence, I excluded 28 students from further analysis, which will in combination with exclusion because of the person fit evaluation regarding interest in particle physics and physics-related self-concept result in the reduced sample N_{red} (N_{red} = 1001, see **Section 6.3.6.**).

---

[24] The term ‘group’ does not refer to an ‘interest type’ (as mentioned in RQ5). A group may indeed form a type or several groups combined may together form a type.

Table 18. Person measure means and standard deviations, minimum and maximum values of the students' interest in mechanics, previous experience with mechanics, and physics-related self-concept, and count per self-concept group and sex for all groups 1_M, 2_M, 3_M, and 4_M (combined analysis).

Mechanics interest groups						
		Group	1_M	2_M	3_M	4_M
Interest in mechanics *Rasch person measures*		*Mean (Standard Deviation)*	0.91 (0.97)	0.07 (0.68)	-0.04 (0.73)	0.31 (0.75)
		Min – Max	-1.87 – 3.71	-2.07 – 2.23	-2.15 – 1.89	-1.26 – 3.06
Previous experience with mechanics *Rasch person measures*		*Mean (Standard Deviation)*	-0.25 (0.84)	-0.41 (0.70)	-0.21 (1.08)	-0.13 (1.01)
		Min – Max	-3.57 – 3.29	-3.57 – 1.28	-3.57 – 4.52	-2.73 – 4.52
Physics-related self-concept *Rasch person measures*		*Mean (Standard Deviation)*	-0.48 (1.94)	-0.90 (1.99)	-1.06 (2.24)	0.06 (2.31)
		Min – Max	-6.16 – 4.41	-6.16 – 5.10	-6.16 – 4.41	-6.16 – 4.41
Self-concept groups	1_{SC}	*Count (row %, column %)*	202 (40.2%, 47.8%)	167 (33.2%, 53.9%)	76 (15.1%, 53.5%)	58 (11.5%, 46.0%)
	2_{SC}		169 (43.3%, 40.0%)	106 (27.2%, 34.2%)	53 (13.6%, 37.3%)	62 (15.9%, 49.2%)
	3_{SC}		52 (48.1%, 12.3%)	37 (34.3%, 11.9%)	13 (12.0%, 9.2%)	6 (5.6%, 4.8%)
Sex	Female	*Count (row %, column %)*	232 (43.0%, 54.8%)	182 (33.7%, 58.7%)	89 (16.5%, 62.7%)	37 (6.9%, 29.4%)
	Male		191 (41.4%, 45.2%)	128 (27.8%, 41.3%)	53 (11.5%, 37.3%)	89 (19.3%, 70.6%)

Note: The descriptive statistics presented in this table were calculated using the reduced sample size following the exclusion of persons based on person misfit regarding mechanics interest, particle physics interest, and physics-related self-concept (N_{red} = 1001, see **Section 6.3.6.**). The Rasch person measures are listed in logit units. Lower and higher person measures represent persons with less and more interest/experience/self-concept, respectively. The person measures are estimated based on the item measures in Rasch analysis. The analysis of the data regarding students' physics-related self-concept and their previous experience with mechanics in school is presented in **Section 6.3.5.** and **6.3.7.**, respectively.

To learn more about the students' different groups, I calculated corresponding descriptive statistics for the reduced sample size N_{red}. **Table 18** lists the descriptive statistics for the mechanics interest groups (1_M, 2_M, 3_M, and 4_M). The mean interest in mechanics was the highest for group 1_M students, followed by group 4_M students and group 2_M students, which in turn had higher mean interest in mechanics than group 3_M students. This can also be seen in **Figure 12 (a)** which shows the boxplots of the students' Rasch person measures indicating their degrees of interest in mechanics for all mechanics interest groups. The mean experience with mechanics in school was rather similar for all groups. However, the mean physics-related self-concept was the highest for group 4_M students followed by group 1_M students and group 2_M students, which in turn had higher mean self-concept than group 3_M students.

This can also be seen in **Figure 12 (b)** which shows the students' Rasch person measures indicating their degrees of physics-related self-concept for all mechanics interest groups. The groups 1_M, 2_M, and 3_M consisted of slightly more female students than male students, whereas it was the opposite for group 4_M.

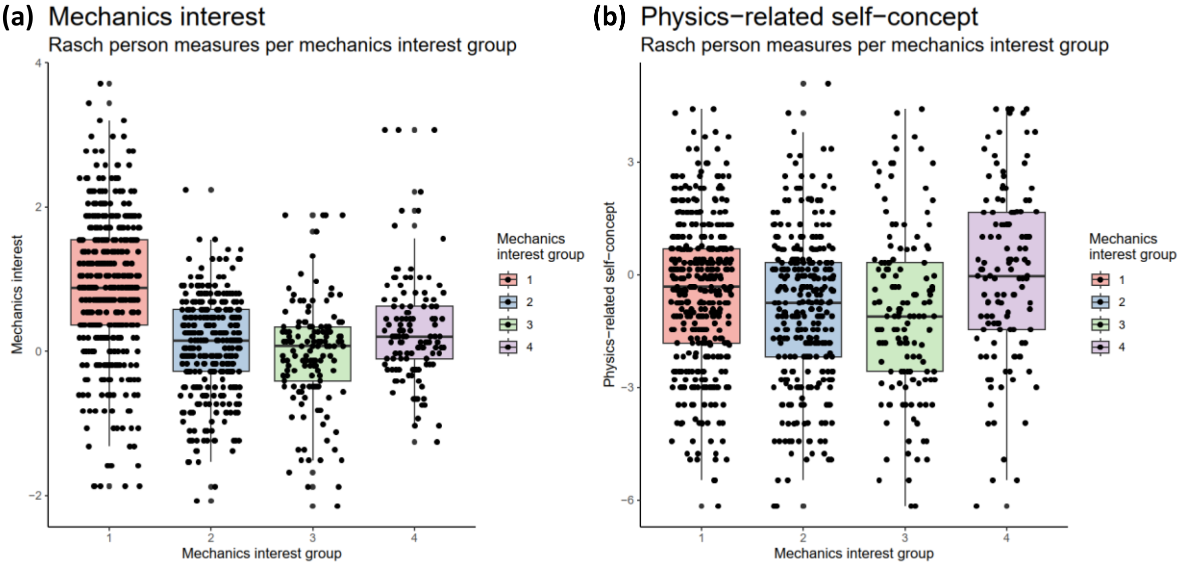


Figure 12. Boxplot with jitter of the students' Rasch person measures in logit units for all mechanics interest groups indicating their (a) degrees of interest in mechanics and (b) physics-related self-concept. The figure is based on the reduced sample size following the exclusion of persons based on person misfit regarding mechanics interest, particle physics interest, and physics-related self-concept (N_{red} = 1001, see **Section 6.3.6.**).

To describe the students' types of interest in mechanics, I compared the item hierarchies of these four groups (**Table 16**). I found that the groups 1_M, 2_M, and 3_M, that is, 86% of the students, have very similar item hierarchies. For the students assigned[25] to these groups certain items were relatively more interesting than other items. For example, the item M07 (artificial organs and joints in medicine) was very interesting for these students. However, there were some small differences in the item hierarchies. For the students assigned to the group 2_M, the items M11 (accident statistics and speed limits) and M04 (transformation of energy of motion into other forms of energy) were less interesting compared to group 1_M and 3_M students. In contrast to the groups 1_M, 2_M, and 3_M, for group 4_M students, all items were similarly interesting. There was no clear item hierarchy, except for the item M07 (artificial organs and joints in medicine) and M11 (accident statistics and speed limits), which were the least interesting items for these students. I will discuss the similarities and differences regarding the item hierarchies of the different groups below.

Overall, I found that the students assigned to the groups 1_M, 2_M, and 3_M, that is, 86% of the students, can be described with one single type of interest in mechanics M_1 because of their equivalent item hierarches. The students with most probable group assignment to group 4_M, that is, 14% of the students, differ in their item hierarchy, and thus form a second type of interest in mechanics M_2.

---

[25] The students are assigned to the group, for which they have the most probable assignment; that is, the assignment is based on the highest probability. In the following, when writing 'assigned to' or 'categorised into' the most probable group assignment is meant.

To learn more about the students' different types of interest in mechanics, I calculated corresponding descriptive statistics for the reduced sample size N_{red}. **Table 19** lists the descriptive statistics for the mechanics interest types (M_1 comprising groups 1_M, 2_M, and 3_M; M_2 comprising group 4_M). The mean interest in mechanics was higher for type M_1 students than for type M_2 students. This can also be seen in **Figure 13 (b)** which shows the boxplots of the students' Rasch person measures indicating their degrees of interest in mechanics for all mechanics interest types. The mean experience with mechanics in school was lower for type M_1 students than for type M_2 students. The mean physics-related self-concept was lower for type M_1 students than for type M_2 students. This can also be seen in **Figure 13 (b)** which shows the students' Rasch person measures indicating their degrees of physics-related self-concept for all mechanics interest types. Type M_2 consisted of 71% male and 29% female students, and 7% of all female and 19% of all male students belonged to type M_2. In sum, type M_2 (comprising group 4_M) students were relatively more male students or students with relatively higher physics-related self-concept compared to type M_1 (comprising groups 1_M, 2_M, and 3_M) students.

Table 19. Person measure means and standard deviations, minimum and maximum values of the students' interest in mechanics, previous experience with mechanics, and physics-related self-concept, and count per self-concept type and sex for both types M_1 and M_2 (combined analysis)

Mechanics interest types				
		Type	M_1	M_2
		Group	1_M, 2_M, and 3_M	4_M
Interest in mechanics *Rasch person measures*		*Mean (Standard Deviation)*	0.46 (0.95)	0.31 (0.75)
		Min – Max	-2.15 – 3.71	-1.26 – 3.06
Previous experience with mechanics *Rasch person measures*		*Mean (Standard Deviation)*	-0.30 (0.84)	-0.13 (1.01)
		Min – Max	-3.57 – 4.52	-2.73 – 4.52
Physics-related self-concept *Rasch person measures*		*Mean (Standard Deviation)*	-0.72 (2.02)	0.06 (2.31)
		Min – Max	-6.16 – 5.10	-6.16 – 4.41
Self-concept types	SC_1	*Count (row %, column %)*	547 (89.5%, 62.5%)	64 (10.5%, 50.8%)
	SC_2		328 (84.1%, 37.5%)	62 (15.9%, 49.2%)
Sex	Female	*Count (row %, column %)*	503 (93.1%, 57.5%)	37 (6.9%, 29.4%)
	Male		372 (80.7%, 42.5%)	89 (19.3%, 70.6%)

Note: The descriptive statistics presented in this table were calculated using the reduced sample size following the exclusion of persons based on person misfit regarding mechanics interest, particle physics interest, and physics-related self-concept (N_{red} = 1001, see **Section 6.3.6.**). The Rasch person measures are listed in logit units. Lower and higher person measures represent persons with less and more interest/experience/self-concept, respectively. The person measures are estimated based on the item measures in Rasch analysis. The analysis of the data regarding students' physics-related self-concept and their previous experience with mechanics in school is presented in **Section 6.3.5.** and **6.3.7.**, respectively.

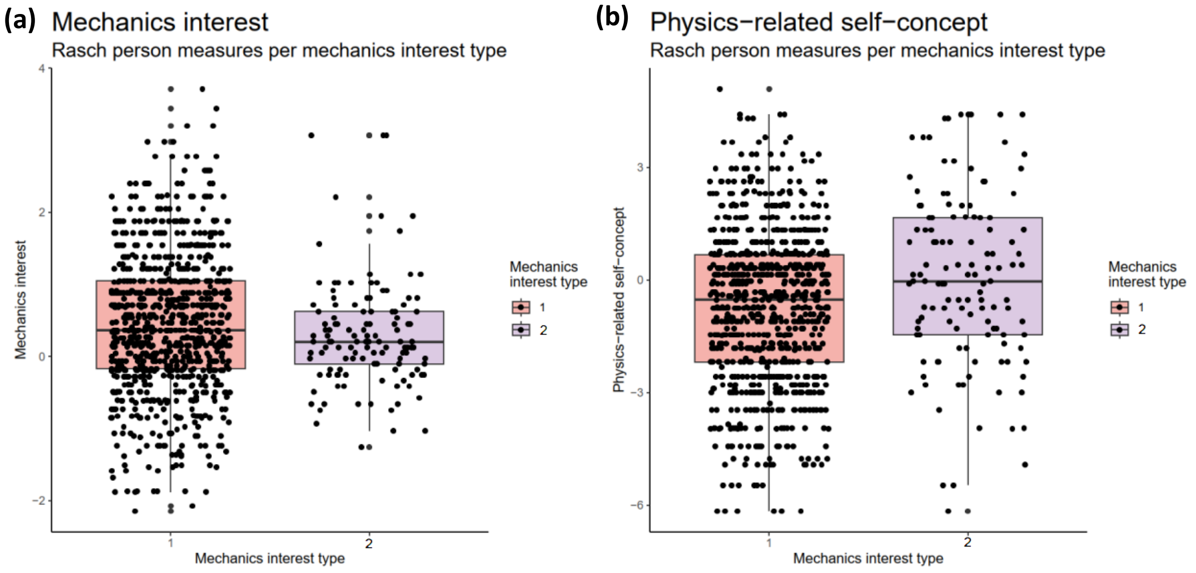


Figure 13. Boxplot with jitter of the students' Rasch person measures in logit units for both mechanics interest types indicating their (a) degrees of interest in mechanics and (b) their degrees of physics-related self-concept. The figure is based on the reduced sample size following the exclusion of persons based on person misfit regarding mechanics interest, particle physics interest, and physics-related self-concept (N_{red} = 1001, see **Section 6.3.6.**).

Discussion

I found that the groups 1_M, 2_M, and 3_M have very similar item hierarchies and can be described with one single type of interest in mechanics M_1. However, there were some small differences in their item hierarchies. For the students assigned to the group 2_M, the items M11 (accident statistics and speed limits) and M04 (transformation of energy of motion into other forms of energy) were less interesting compared to group 1_M and 3_M students. These items are set in the context 'car accidents'. In contrast, the items M09 (designing a safety car for severe accidents) and M03 (increase of probability and consequences of a car accident with speed), which are also both set in the context 'car accidents', were some of the most interesting for group 2_M students. This was also the case for the other groups 1_M and 3_M. I argue that the task 'discussion' of item M11 (accident statistics and speed limits) might have been less interesting than other tasks for the students of group 2_M. Moreover, the context 'car accidents' is formulated rather unspectacularly in the items M11 (accident statistics and speed limits). This contrasts with the other items set in the same context, that is, items M03 (increase of probability and consequences of a car accident with speed) and M09 (designing a safety car for severe accidents). Hence, the item M11 (accident statistics and speed limits) might have been less interesting. In addition, the context 'car accidents' is not mentioned directly in item M04 (transformation of energy of motion into other forms of energy) but can only be assumed indirectly, and hence this item might have been less interesting. In addition, the item M04 (transformation of energy of motion into other forms of energy) presents the content 'qualitative physics', whereas item M03 (increase of probability and consequences of a car accident with speed) presents the content 'quantitative physics'. Although both items, M03 (increase of probability and consequences of a car accident with speed) and M04 (transformation of energy of motion into other forms of energy), are set in the context 'car accident', the content of item M04 might have been less interesting in this context than the content of item M03 because the quantitative aspect of item M03 is a measure for the severity of the consequences of a car accident. The item M08 (constructing different pulleys and trying them out) was relatively more interesting for group 1_M and 2_M students. I argue that the students were comparatively more interested in the task 'doing an experiment' presented in item M08 (constructing different pulleys and trying them out). Yet, for the group 3_M this is the least interesting item. Here, I argue that the students were either not interested in the task 'doing an experiment' or in the content 'pulley'.

Apart from the described differences the item hierarchies were equivalent. When conducting a mixed Rasch analysis, one may at first glance interpret all the different groups as different types of students, as done, for example, in the initial publications of the IPN study (Häußler, 1987; Häußler et al., 1996; Häußler, Hoffmann et al., 1998; Häußler, Lehrke et al., 1998). However, I argue that the students were categorised into three groups in the mixed Rasch model despite their similar interest profiles because of their different response styles.

The category probability plots (**Figure 10**) show that group 1_M tended to use the non-extreme categories, especially the middle category and not the negative extreme category, across the 11 items. Group 2_M also tended to use the non-extreme rating scale categories, but also the negative and positive extreme categories depending on the item. Group 3_M tended to use the negative and positive extreme rating scale categories depending on the item. These differences in the rating scale category use also explain why the distances between the thresholds in the interest profiles (**Figure 9**) are different for the groups 1_M, 2_M, and 3_M. Non-extreme response style (NERS) is characterised 'by widely spaced first and fourth thresholds while the second and third thresholds are close together', and extreme response stye (ERS) by thresholds that 'are nearby each other, sometimes even overlapping' (Wetzel et al., 2013). For group 1_M, the first threshold is widely spaced while the others are close together, that is, the distance from the first to the second threshold is bigger than the other threshold distances, which resembles a NERS. This is because the students used the extreme category *'very high'* (=5) in addition to the non-extreme categories and not the category *'very low'* (=1). For group 2_M, this pattern is similar, although all thresholds are closer together because the students also used all the extreme categories in addition to the non-extreme categories, which rather resembles the intended use of the rating scale (i.e. a 'normal' response style). For group 3_M, all thresholds are very close together because the students used all the rating scale categories, but especially the extreme categories, which resembles an ERS. Moreover, the thresholds 1 and 2 are not ordered, that is, threshold 2 is lower than threshold 1. Non-ordered threshold measures indicate that some categories of the rating scale are never the most probable. Following Andrich (1978, 2005), some researchers argue that non-ordered threshold measures indicate a violation of the intended category ordering. However, following Linacre (1991) and Masters (1982), other researchers (e.g. Adams et al., 2012) show that non-ordered threshold measures result from a low frequency of responses in one or more categories. Adams et al. (2012) show in their study that there is 'no necessary connection' between the ordering of the threshold measures and the ordering of the categories. Hence, having groups with non-ordered threshold measures in a model indicates that these groups differ in their response style, that is, their use of the rating scale. This differing category use of group 3_M can be observed from the category probability plots (**Figure 10**). The group 3_M uses the category 'very low' more probably than 'low', and hence the thresholds are not ordered.

In contrast to the groups 1_M, 2_M, and 3_M, for group 4_M students, all items were similarly interesting. There was no clear item hierarchy, except for the item M07 (artificial organs and joints in medicine) and M11 (accident statistics and speed limits), which were the least interesting items for these students. For these students the context 'car' was the most interesting, not only when referring to accidents. Hence, I argue that the students with most probable group assignment to group 4_M, 14% of the sample, are considered to form a second type of interest in mechanics M_2, comprising students interested in physics relating to the motion of cars. However, this is a rather small fraction of the students. Group 4_M tended to use all the rating scale categories, especially the non-extreme categories, rather equally across the 11 items, which resembles a 'normal' response style. However, for the items M07 (artificial organs and joints in medicine) and M08 (constructing different pulleys and trying them out) rather the two negative categories were used. It is interesting that exactly these items on a medical context and a hands-on experiment task were rated the least interesting for these students. Apart from these students forming a second type of interest, it may also be that these students read the rating scale wrong, that is, the category 'very high' becomes 'very low' and so on. However, in the in total 24 student think-aloud interview, in which the students had to rate their interest using this rating scale, no such problems were observed. Hence, I considered it unlikely that 14% of the students interpret the rating scale wrong in the online questionnaire. It might also have been a combination of both reasons, first, a wrong interpretation of the rating scale, and second, a second type of interest in mechanics.

The analysis of the answer profiles and category probability distributions showed that overall for most students (i.e. type M_1 comprising the students assigned to the groups 1_M, 2_M, and 3_M) the same items were more or less interesting relative to each other, whereas only a small fraction of the students (i.e. type M_2 comprising the students assigned to the group 4_M) differed in their interest. Since the students forming type M_2 were interested in items set in the context 'car', even when not referring to 'accidents', I referred to them as the 'students interested in physics relating to the motion of cars'. Further discussions on which contexts were more (or less) interesting relative to each other for the types of interest in mechanics are presented in **Section 6.3.3.** which is on RQ6.

6.3.2.ii. Particle Physics

Second, I investigated students' types of interest in particle physics. That is, I conducted the same type of group analysis described in the previous sections but for the students' interest in particle physics (instead of their interest in mechanics).

Results

As for the students' interest in mechanics, I calculated and compared four different models with one to four groups[26]. The BIC value of the 3-groups-model was the smallest. The delta-BIC values of the four calculated models with respect to the model with the lowest BIC value are listed in

Table 20. The delta-BIC values indicated very strong evidence against the 1-group-model and the 2-groups-model and strong evidence against the 4-groups-model.

Table 20. Delta BIC-values of the four calculated models to describe students' interest in particle physics

	Delta-BIC			
Number of groups	1	2	3	4
Delta-BIC	905	230	0	41

---

[26] The term 'group' does not refer to an 'interest type' (as mentioned in RQ5). A group may indeed form a type or several groups combined may together form a type.

Qualitative analysis of the groups described by the different models also showed that a distinction into three different groups is sufficient to describe the students' interest in particle physics. The distinction into more groups did not improve the description of students' interest because then the groups' sizes would be very small and several groups would have very similar interest profiles. Consequently, I categorised the students into three different groups of interest in particle physics. Group 1_{PP} consisted of 46% of the sample, group 2_{PP} of 33%, and group 3_{PP} of 21%.[27,28]

Figure 14 shows the interest profiles of all the groups 1_{PP}, 2_{PP}, and 3_{PP}. One can see that the threshold measures of each item for group 1_{PP} are more spread out and further away from each other than the threshold measures of each item for group 2_{PP}. In group 1_{PP} the distance from the first to the second threshold is bigger than the other threshold distances. In group 2_{M} the thresholds are not ordered, which is a result of low response frequencies in one or more rating scale categories (as discussed below). Apart from this difference in in-between-thresholds-distances, the interest profiles (i.e. how interesting the different items are relative to each other) were similar for both groups 1_{PP} and 2_{PP}.

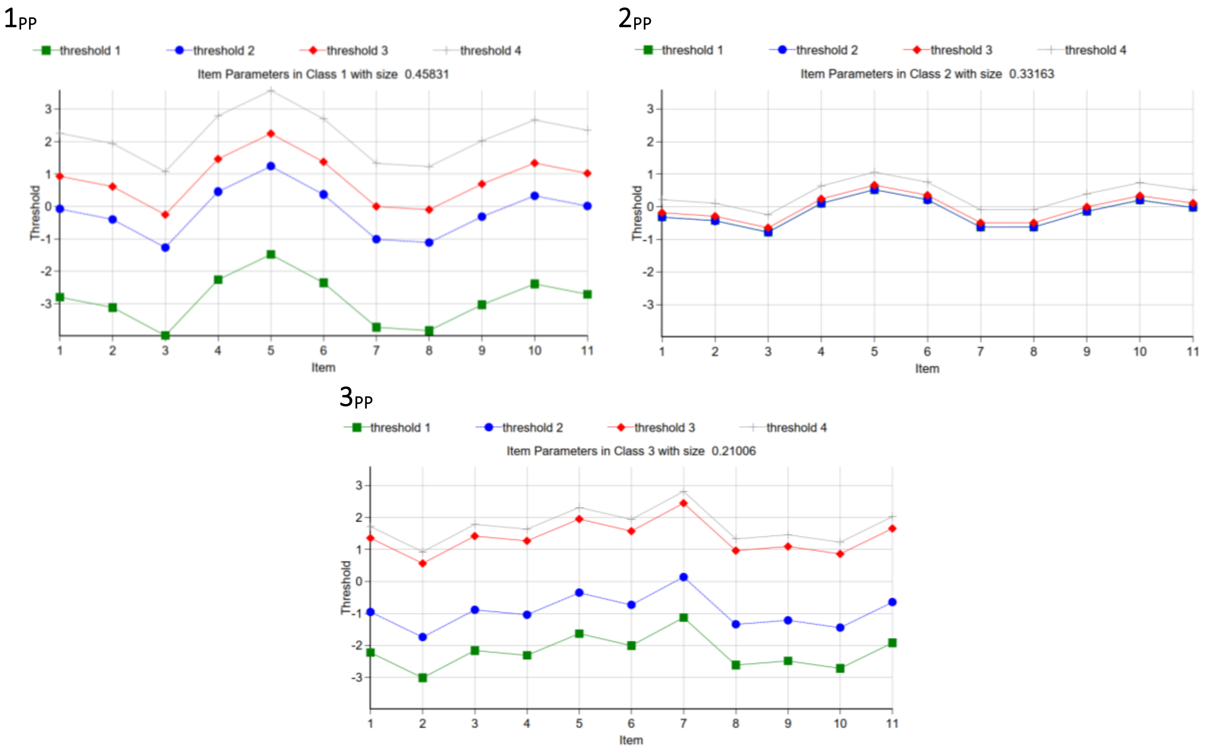


Figure 14. Interest profiles, that is, how interesting the different items are relative to each other, of the three groups 1_{PP}, 2_{PP} and 3_{PP}.

Note: On the x-axis, the items are shown, and on the y-axis, the threshold measures are shown in logit units. For example, threshold 1 is between the categories *'very low'* and *'low'*. The line between the data points does not have a meaning but helps to observe some patterns.

---

[27] The index PP stands for Particle Physics.

[28] The students are assigned to the group, for which they have the most probable assignment; that is, the assignment is based on the highest probability. In the following, when writing 'assigned to' or 'categorised into' the most probable group assignment is meant.

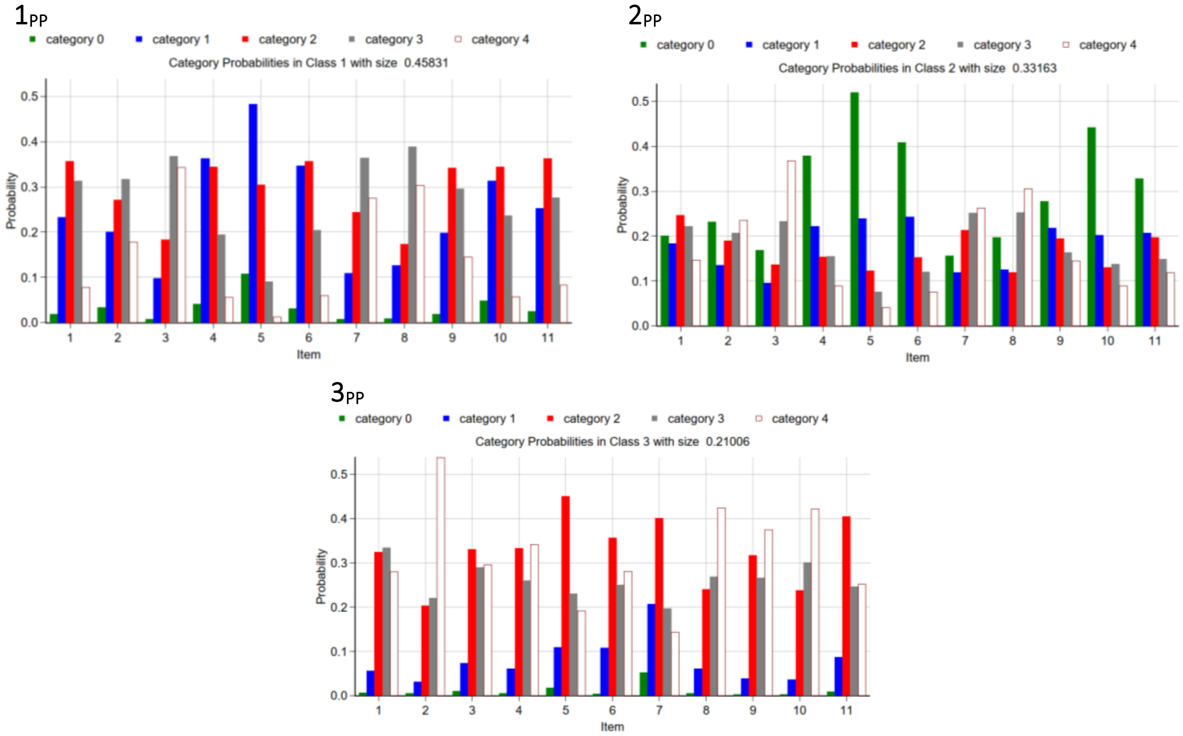


Figure 15. Category probabilities, that is, the probabilities to choose the different categories of the rating scale, of the three groups 1_{PP}, 2_{PP} and 3_{PP}.
Note: The students expressed their degree of interest in each item on a 5-category rating scale ('My interest in this is …' *very high* (=4), *high* (=3), *medium* (=2), *low* (=1), or *very low* (=0))

For example, item PP07 (medical diagnostics, set in the context 'medicine') has very low threshold measures, whereas items PP05 (masses of particles) and PP10 (calculating the energy of a particle collision at the speed of light; set in the context 'quantitative science') have very high threshold measures. Finally, one can see that the interest profile of group 3_{PP} differs from the other groups. For example, item PP07 (medical diagnostics) has very high threshold measures. The thresholds 1, 2, and 3 are further away from each other, whereas the thresholds 3 and 4 are very close together.

To study the difference between groups 1_{PP}, 2_{PP} and 3_{PP}, I analysed how the students used the different answer categories of the rating scale. **Figure 15** shows the category probabilities. One can see that the groups differed in their use of the rating scale categories. Group 1_{PP} tended to use the non-extreme, middle categories of the rating scale across all items and additionally the extreme category 'very high'. Group 2_{PP} tended to use all the rating scale categories, especially also the extreme categories depending on the item. Group 3_{PP} tended to use the categories 'medium', 'high', and 'very high'.

When conducting (mixed) Rasch analyses, an evaluation of the obtained item and person measures is crucial as it provides further evidence for the quality of the chosen model. When checking for item fit for particle physics, all items fit well in in group 1_{PP} (*ZQ* values ranging from -1.27 to 1.18) and group 3_{PP} (*ZQ* values ranging from -0.82 to 1.47). In group 2_{PP}, all items fit well (*ZQ* values ranging from -1.91 to 0.61) except for the item PP07 (medical diagnostics), of which the *ZQ* value is 3.57. I argue that this is because item PP07 (medical diagnostics) is set in a medical context unlike the other items, and hence, might have appeared to be less related to the measured latent trait, that is, interest in particle physics, or to relate to different aspects of it. Based on the item fit statistics, I argue that the data supports one trait, interest in particle physics, in all three groups. Yet, interest in particle physics is a broad trait because particle physics may be set in many different contexts (cf. Linacre (2020) about mathematics as a broad domain). Hence, based on the item fit statistics, the data supported one trait, interest in particle physics, for all groups, that is, that the IPPI is unidimensional as already shown in my first study (**Chapter 5**). Then, I evaluated the Rasch person measures for the students' interest in particle physics. I excluded those students from further analysis, whose newfit 1 and/or 2 value was $\ll -3$ or $\gg 3$ or for who no newfit values were calculated. In total, I excluded 39 students from further analysis because of misfit regarding their interest in particle physics, which will in combination with exclusion because of the person fit evaluation regarding interest in particle physics and physics-related self-concept result in the reduced sample N_{red} (N_{red} = 1001, see **Section 6.3.6.**).

To learn more about the students' different groups, I calculated corresponding descriptive statistics for the reduced sample size N_{red}. **Table 21** lists the descriptive statistics for the particle physics interest groups (1_{PP}, 2_{PP}, and 3_{PP}). The mean interest in particle physics was the highest for group 3_{pp} students, followed by group 1_{PP} students, which in turn had higher mean interest than group 2_{PP} students. This can also be seen in **Figure 16 (a)** which shows the boxplots of the students' Rasch person measures indicating their degrees of interest in particle physics for all particle physics interest groups. The mean experience with particle physics in school was rather similar for all groups. However, the mean physics-related self-concept was the highest for group 3_{PP} students followed by group 1_{PP} students and group 2_{PP} students. This can also be seen in **Figure 16 (b)** which shows the students' Rasch person measures indicating their degrees of physics-related self-concept for all particle physics interest groups. The groups 1_{PP} and 2_{PP} consisted of slightly more female students than male students, whereas it is the opposite for group 3_{PP}.

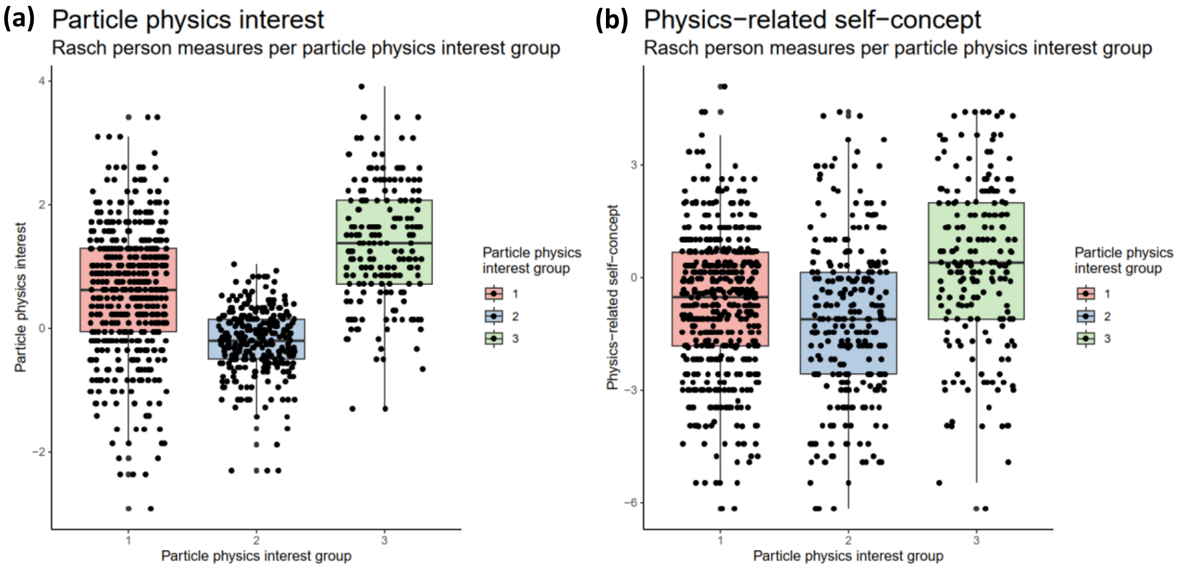


Figure 16. Boxplot with jitter of the students' Rasch person measures in logit units for all particle physics interest groups indicating their (a) degrees of interest in particle physics and (b) their physics related-self-concept. The figure is based on the reduced sample size following the exclusion of persons based on person misfit regarding mechanics interest, particle physics interest, and physics-related self-concept (N_{red} = 1001, see **Section 6.3.6.**).

Table 21. Person measure means and standard deviations, minimum and maximum values of the students' interest in particle physics, previous experience with particle physics, and physics-related self-concept, and count per self-concept groups and sex for all groups 1_{PP}, 2_{PP}, and 3_{PP}

Particle physics interest groups					
		Group	1_{PP}	2_{PP}	3_{PP}
Interest in particle physics *Rasch person measures*		*Mean (Standard Deviation)*	0.58 (1.01)	-0.21 (0.54)	1.37 (0.95)
		Min – Max	-2.92 – 3.42	-2.30 – 1.04	-1.30 – 3.92
Previous experience with particle physics *Rasch person measures*		*Mean (Standard Deviation)*	-0.65 (0.88)	-0.77 (1.08)	-0.57 (1.16)
		Min – Max	-3.38 – 2.27	-3.38 – 4.00	-3.38 – 4.00
Physics-related self-concept *Rasch person measures*		*Mean (Standard Deviation)*	-0.71 (1.86)	-1.17 (2.15)	0.35 (2.18)
		Min – Max	-6.16 – 5.10	-6.16 – 4.41	-6.16 – 4.41
Self-concept groups	1_{SC}	*Count (row %, column %)*	263 (52.3%, 51.3%)	166 (33.0%, 57.8%)	74 (14.7%, 36.8%)
	2_{SC}		186 (47.7%, 36.3%)	98 (25.1%, 34.1%)	106 (27.2%, 52.7%)
	3_{SC}		64 (59.3%, 12.5%)	23 (21.3%, 8.0%)	21 (19.4%, 10.4%)
Sex	Female	*Count (row %, column %)*	306 (56.7%, 59.6%)	174 (32.2%, 60.6%)	60 (11.1%, 29.9%)
	Male		207 (44.9%, 40.4%	113 (24.5%, 39.4%)	141 (30.6%, 70.1%)

Note: The descriptive statistics presented in this table were calculated using the reduced sample size (N_{red} = 1001, see **Section 6.3.6.**). The Rasch person measure means and standard deviations are listed in logit units. Lower and higher person measures represent persons with less and more interest/experience/self-concept, respectively. The person measures are estimated based on the item measures in Rasch analysis. The analysis of the students' physics-related self-concept and their previous experience with particle physics in school is presented in **Section 6.3.5.** and **6.3.7.**, respectively.

To describe the students' types of interest in particle physics, I compared the item hierarchies of the three groups, in which the items are sorted by their item measure, that is, by their interestingness (**Table 22**). I found that the groups 1_{PP} and 2_{PP}, had similar item hierarchies. This means that for both groups the relative interestingness of the different items was the same. For example, for both groups the most interesting items were PP03 (particle detectors and smuggled arms), PP08 (transforming a mobile phone into a particle detector), PP07 (medical diagnostics), and PP02 (particle physics and the big bang). However, the item hierarchy of the group 3_{PP} differed. For example, item PP07 (medical diagnostics) was the least interesting in the group 3_{PP}, whereas it was among the most interesting for the other groups, and item PP10 (calculating the energy of a particle collision at the speed of light) was among the most interesting in group 3_{PP}, whereas it was among the least interesting for the other groups. I will discuss the similarities and differences regarding the item hierarchies of the different groups below.

Table 22. Item IDs, corresponding Rasch item measures (reflecting their interestingness), and standard errors (SE) for all three groups 1_{PP}, 2_{PP} and 3_{PP}.

Item	Group 1_{PP}	SE	Group 2_{PP}	SE	Group 3_{PP}	SE
PP05	1.40	0.06	0.70	0.05	0.57	0.09
PP04	0.62	0.06	0.28	0.05	-0.11	0.09
PP06	0.53	0.06	0.39	0.05	0.19	0.09
PP10	0.49	0.06	0.37	0.05	-0.52	0.09
PP11	0.17	0.06	0.15	0.05	0.28	0.09
PP01	0.08	0.06	-0.15	0.05	-0.02	0.09
PP09	-0.15	0.06	0.03	0.05	-0.28	0.09
PP02	-0.24	0.06	-0.26	0.05	-0.81	0.09
PP07	-0.85	0.06	-0.46	0.05	1.07	0.09
PP08	-0.95	0.06	-0.45	0.05	-0.41	0.09
PP03	-1.10	0.06	-0.61	0.05	0.04	0.09

Note: The Rasch item measures are listed in logit units. Lower and higher item measures represent more and less interesting items, respectively. The mean item measure of each group is set to zero in mixed Rasch analysis. The items are ordered according to their measures in group 1_{PP}. The colour scale represents the interestingness across all groups; dark blue is the most interesting item and dark red is the least interesting item. More detailed item tables and Wright Maps for each group are presented in the **Appendix 10.2.7.** and **10.2.8.** of this thesis.

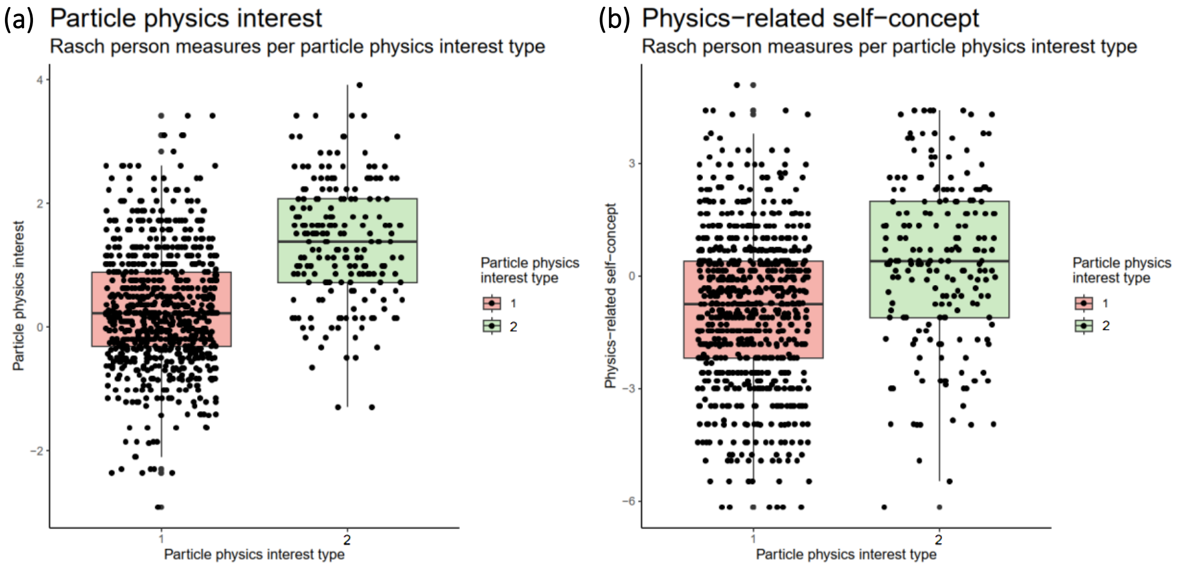


Figure 17. Boxplot with jitter of the students' Rasch person measures in logit units indicating their degrees of (a) interest in particle physics and (b) physics-related self-concept for both particle physics interest types. The figure is based on the reduced sample size following the exclusion of persons based on person misfit regarding mechanics interest, particle physics interest, and physics-related self-concept (N_{red} = 1001, see **Section 6.3.6.**).

Since the item hierarchies were equivalent, I argue that the students of the groups 1_{PP} and 2_{PP}, that is, 79% of the students, can be described with one single type of interest in particle physics (PP_1). The students with most probable group assignment to group 3_{PP}, that is, 21% of the students, differ in their item hierarchy and are considered to form a second type of interest in particle physics (PP_2).

To learn more about the students' different types of interest in particle physics, I calculated corresponding descriptive statistics for the reduced sample size N_{red}. **Table 23** lists the descriptive statistics for the particle physics interest types (PP_1 comprising groups 1_{PP} and 2_{PP} and PP_2 comprising group 3_{PP}). The mean interest in particle physics was lower for type PP_1 students than for type PP_2 students. This can also be seen in **Figure 17 (a)** which shows the boxplots of the students' Rasch person measures indicating their degrees of interest in particle physics for all particle physics interest types. The mean experience with particle physics in school was similar for type PP_1 and PP_2 students. The mean physics-related self-concept is lower for type PP_1 students than for type PP_2 students. This can be seen in **Figure 17 (b)** which shows the students' Rasch person measures indicating their degrees of physics-related self-concept for all particle physics interest types. Type PP_2 consisted of 70% male and 30% female

Table 23. Person measure means and standard deviations, minimum and maximum values of the students' interest in particle physics, previous experience with particle physics, and physics-related self-concept, and count per self-concept types and sex for all types PP_1 and PP_2

Particle physics interest types				
		Type	PP_1	PP_2
		Group	1_{PP} and 2_{PP}	3_{PP}
Interest in particle physics *Rasch person measures*		*Mean (Standard Deviation)*	0.30 (0.95)	1.37 (0.95)
		Min – Max	-2.92 – 3.42	-1.30 – 3.92
Previous experience with particle physics *Rasch person measures*		*Mean (Standard Deviation)*	-0.69 (0.96)	-0.57 (1.61)
		Min – Max	-3.38 – 4.00	-3.38 – 4.00
Physics-related self-concept *Rasch person measures*		*Mean (Standard Deviation)*	-0.87 (1.98)	0.35 (2.18)
		Min – Max	-6.16 – 5.10	-6.16 – 4.41
Self-concept types	SC_1	*Count (row %, column %)*	516 (84.5%, 64.5%)	95 (15.5%, 47.3%)
	SC_2		284 (72.8%, 35.5%)	106 (27.2%, 52.7%)
Sex	Female	*Count (row %, column %)*	480 (88.9%, 60.0%)	60 (11.1%, 29.9%)
	Male		320 (69.4%, 40.0%)	141 (30.6%, 70.1%)

Note: The descriptive statistics presented in this table were calculated using the reduced sample size following the exclusion of persons based on person misfit regarding mechanics interest, particle physics interest, and physics-related self-concept (N_{red} = 1001, see **Section 6.3.6.**). The Rasch person measures are listed in logit units. Lower and higher person measures represent persons with less and more interest/experience/self-concept, respectively. The person measures are estimated based on the item measures in Rasch analysis. The analysis of the data regarding students' physics-related self-concept and their previous experience with particle physics in school is presented in **Section 6.3.5.** and **6.3.7.**, respectively.

students, and 11% of all female and 31% of all male students belonged to type PP_2. In sum, type PP_2 (comprising group 3_{PP}) students were relatively more male students or students with relatively higher physics-related self-concept compared to type PP_1 (comprising groups 1_{PP} and 2_{PP}) students. This is similar to the findings concerning the mechanics interest types.

Discussion

For the content area particle physics, I found that 79% of the sample, that is, the groups 1_{PP} and 2_{PP}, have a similar interest profile. At the bottom of the hierarchy are the most interesting items, PP03 (particle detectors and smuggled arms), PP08 (transforming a mobile phone into a particle detector), PP07 (medical diagnostics), and PP02 (particle physics and the big bang). These items are either set in an interesting context (PP03: particle detectors and smuggled arms, the context ‘smuggled arms’ suggests danger similar to the context ‘car accidents’ in mechanics; PP02: particle physics and the big bang, the context ‘the big bang’ is spectacular/fascinating/unique; PP07: medical diagnostics, the context ‘medicine’ is related to one’s own body) or present an interesting task (PP08: transforming a mobile phone into a particle detector, task ‘hands-on experiment’). The least interesting items present particle physics contents (e.g. the item PP05: masses of particles) without setting these contents in a context.

As for mechanics interest, I argue that the students of groups 1_{PP} and 2_{PP} were categorised into different groups of particle physics interest despite their equivalent item hierarchies because they had different response styles. This can be seen in the category probability plots (**Figure 15**) and the interest profiles (**Figure 14**). Group 1_{PP} students used the extreme category *‘very high’* (=5) in addition to the non-extreme categories, and thus, only the first threshold is widely spaced, while the others are close together, which resembles a NERS. Group 2_{PP} students used all the categories, and hence, all thresholds are very close together. In addition, group 2_{PP} students tended to use the extreme category ‘very low’, and hence, the thresholds 1 and 2 are even overlapping, which resembles an ERS.

In comparison to groups 1_{PP} and 2_{PP}, the item hierarchy of the group 3_{PP} differed. For example, item PP07 (medical diagnostics) was the least interesting for the group 3_{PP}, whereas it was among the most interesting for the other groups. Item PP07 (medical diagnostics) does not present a content but only the context ‘medical diagnostics’. How this context relates to particle physics can be read only in the introductory text. Item PP10 (calculating the energy of a particle collision at the speed of light) was among the most interesting in group 3_{PP}, whereas it was among the least interesting for the other groups. Although the items PP05 (masses of particles) and PP10 (calculating the energy of a particle collision at the speed of light) both present quantitative physics, item PP10 (calculating the energy of a particle collision at the speed of light) was more interesting than item PP05 (masses of particles). I argue that the quantitative physics might have been less interesting in item PP05 (masses of particles), which is not set in a context, than in item PP10 (calculating the energy of a particle collision at the speed of light), which is set in the spectacular context ‘particle collisions at almost the speed of light’. In addition, the item PP04 (particles in the nucleus space of an atom) that presents qualitative physics and somewhat tackles the context ‘existential questions of humankind’ (‘*What are we made of?*’) was relatively interesting. Overall, one can see that some items on the special aspects of particle physics were the most interesting. These items tackle the context ‘existential questions of humankind’ (PP02: particle physics and the big bang, ‘*Where do we come from?*’; PP04: particles in the nucleus space of an atom, ‘*What are we made of?*’), are set in the context ‘spectacular experiment’ (PP10: calculating the energy of a particle collision at the speed of light) or present the task ‘experiment’ (PP08: transforming a mobile phone into a particle detector; PP09: planning an experiment on particle acceleration). In contrast, the item that just presents the context ‘medicine’ but no content (PP07: medical diagnostics) is the least interesting.

Since the item hierarchy of group 3_{PP} differed from the other groups, I considered the students assigned to group 3_{PP}, that is, 21% of the students, to form a second type of interest in particle physics PP_2. Since these students were interested in the special aspects of particle physics as a content area and a scientific endeavour, that is, its focus on the existential questions of humankind (PP02: particle physics and the big bang; PP04: particles in the nucleus space of an atom) and on experiments (PP08: transforming a mobile phone into a particle detector; PP09: planning an experiment on particle acceleration; PP10: calculating the energy of a particle collision at the speed of light), I referred to them as the ‘particle physics lovers’.

As can be seen in category probability plots (**Figure 15**) and the interest profiles (**Figure 14**), group 3_{PP} tends to use the rating scale categories ‘*medium’* (=2), ‘*high’* (=3), and ‘*very high’* (=4). Hence, their thresholds 3 and 4 are very close together and the thresholds 1 and 2 are widely spaced. This resembles their tendency to only use the three most positive response categories. Again, it is somewhat surprising that item PP07 (medical diagnostics) is the least interesting for this group. As already discussed for mechanics, it may also be that these students read the rating scale wrong, that is, the category ‘very high’ becomes ‘very low’ and so on. However, I consider it unlikely that 21% of the students interpret the rating scale wrong in the online questionnaire. I argue that a combination of both reasons seems most likely, first, a wrong interpretation of the rating scale, and second, a second type of interest in particle physics PP_2, referred to as the ‘particle physics lovers’.

The analysis of the answer profiles and category probability distributions showed that overall for most students (i.e. type PP_1 comprising the students assigned to the groups 1_{PP} and 2_{PP}) the same items were more (or less) interesting relative to each other, whereas only a small fraction of the students (i.e. type PP_2 comprising the students assigned to the group 3_{PP}) differed in their interest. Further discussions on which contexts were more (or less) interesting relative to each other for the types of interest in particle physics are presented in **Section 6.3.3.** which is on RQ6.

6.3.2.iii. Comparison

In this section I will compare the students’ types of interest in mechanics and particle physics, respectively, and discuss common aspects.

Results

First, I compared the most probable group assignments for the students regarding their interest in mechanics and particle physics, respectively (**Table 24**). Reading the table in rows, one can see that the students assigned to group 1_M were also rather assigned to group 1_{PP}. Yet, the students assigned to the group 2_M were assigned to both 1_{PP} and 2_{PP}. Students assigned to group 3_M were rather assigned to group 2_{PP}. Students assigned to group 4_M were equally assigned to all particle physics groups, but especially to the group 3_{PP}. Reading the table in columns, one can see that group 1_{PP} mainly consisted of group 1_M and 2_M students. Group 2_{PP} mainly consisted of group 2_M and 3_M students. Group 3_{PP} mainly consisted of group 1_M and also of group 4_M students.

Table 24. Most probable group assignments for all groups 1_M, 2_M, 3_M, and 4_M (combined analysis) and the groups 1_{PP}, 2_{PP}, and 3_{PP}.

	Interest in Particle Physics			
	Count (row %, column %)			
	Group	1_{PP}	2_{PP}	3_{PP}
Interest in Mechanics	1_M	290 (60.7%, 51.1%)	69 (14.4%, 18.6%)	119 (24.9%, 47.6%)
	2_M	182 (51.9%, 32.1%)	132 (37.6%, 35.7%)	37 (10.5%, 14.8%)
	3_M	50 (24.9%, 8.8%)	117 (58.2%, 31.6%)	34 (16.9%, 13.6%)
	4_M	45 (28.7%, 7.9%)	52 (33.1%, 14.1%)	60 (38.2%, 24.0%)

Second, I compared the most probable type assignments for the students in mechanics and particle physics (**Table 25**). Reading the table in rows, one can see that the students assigned to type M_1 were also rather assigned to type PP_1. Yet, the students assigned to the type M_2 were also rather assigned to type PP_1. Reading the table in columns, one can see that type PP_1 mainly consisted of type M_1 students. Type PP_2 also mainly consisted of type M_1 students. Only 60 students were assigned to both types M_2 and PP_2.

Table 25. Most probable type assignments for all types M_1 and M_2 (combined analysis) and the types PP_1 and PP_2.

	Interest in Particle Physics		
	Count (row %, column %)		
	Type	PP_1	PP_2
Interest in Mechanics	M_1	840 (81.6%, 89.6%)	190 (18.4%, 76.0%)
	M_2	97 (61.8%, 10.4%)	60 (38.2%, 24.0%)

To gain additional evidence for my explanation based not only on the students' degree of interest but also on their response style, I analysed whether the same students showed certain response styles, such as a NERS or an ERS, across both content areas mechanics and particle physics. Thus, I investigated whether the students of groups 1_M (NERS), 2_M, and 4_M (both rather normal response style) are also in group 1_{PP} (NERS) and 3_{PP} (three most positive categories; see **Table 24**). Moreover, I investigated whether the students of groups 3_M (ERS) are also in group 2_{PP} (ERS; see **Table 24**). I found that 61% of the students assigned to group 1_M were also rather assigned to group 1_{PP}; 52% of the students assigned to group 2_M were assigned to group 1_{PP} but 38% were assigned to group 2_{PP}; and 38% of the students assigned to group 4_M were assigned to group 3_{PP} and the rest rather equally to group 1_{PP} and group 2_{PP}. Indeed, 58% of the students assigned to group 3_M were assigned to group 2_{PP}. On the other hand, 48% of the group 3_{PP} students were assigned to group 1_M which is not surprising considering their rather similar response category use (3_{PP}: three most positive categories; 1_M: NERS plus most positive category).

Discussion

This analysis provided additional evidence for the explanation of the categorisation into different groups based on response style rather than qualitative differences in interest. The item hierarchies were similar for students in groups 1_M, 2_M, and 3_M, and for the students in groups 1_{PP} and 2_{PP}, that is, there were little qualitative differences in the interestingness of different items relative to each other. Instead, the students differed in their use of the rating scale. Additional evidence for our response style explanation is that the students with NERS (groups 1_M and 1_{PP}, respectively) have a higher mean interest than the students with ERS (groups 3_M and 2_{PP}, respectively), which is counterintuitive (see **Table 18** and **Table 21**).

This can also be seen in the boxplots with jitter for the interest groups of both content areas (see **Figure 12 (a)** and **Figure 16 (a)**). For mechanics, the NERS of group 1_M can be seen because the students' person measures are evenly and widely spread around the mean, whereas the ERS of group 3_M can be seen because the students' person measures are relatively close around the mean plus there are extreme positive and negative outliers. The response styles of groups 2_M and 4_M were overall found to be normal. For 2_M, the students' person measures are evenly but narrowly spread around the mean, which rather resembles a NERS group, whereas for 4_M, the students' person measures are relatively close around the mean plus there are extreme positive outliers, which rather resembles an ERS group. For particle physics, the distribution of person measures for group 1_{PP} students resembles a NERS group, whereas the distribution of person measures for group 2_{PP} students resembles an ERS groups. However, the person measures of group 3_{PP} are evenly and widely spread around the mean. This also reflects their use of the rating scale (i.e. mostly using the three most positive categories).

In sum, I argue that since the students differ in their response style, they were categorised into the different groups 1_M, 2_M, and 3_M and 1_{PP} and 2_{PP}, respectively, although their item hierarchies were equivalent, that is, the same items are more (or less) interesting for all three groups. Hence, I argue that 86% of the students in mechanics and 79% of the students in particle physics can be described in terms of one single type of interest in physics, the type of students that is only interested in physics set in certain contexts, type M_1 and PP_1, respectively. However, for mechanics, I found that 14% of the students, that is, group 4_M, had a different item hierarchy. For these students all items were rather similarly interesting whereby the ones that present the context 'car' were slightly more interesting than other items. Hence, I found that this second type of interest in mechanics (M_2) comprised students who are interested in physics relating to the motion of cars. For particle physics, I found that 21% of the sample, that is, the group 3_{PP}, had a different item hierarchy compared to the other particle physics interest groups. These students were rather interested in the special aspects of particle physics as a scientific endeavour, that is, its focus on the existential questions of humankind (PP02, particle physics and the big bang) and on experiments (PP08, transforming a mobile phone into a particle detector; PP09, planning an experiment on particle acceleration; PP10, calculating the energy of a particle collision at the speed of light). Since the group 3_{PP} differed from the other groups, these students can be described in terms of a second type of interest in particle physics PP_2.

Table 26 presents the main points of the discussion concerning RQ5. For both content areas two different types of interest in physics were distinguished. The first one is named 'Physics? Only in certain contexts!' and comprises students who are only interested in physics contents when set in certain contexts, that is, the type M_1 concerning the content area mechanics and the type PP_1 concerning the content area particle physics. Most students (71%) were categorised into this type of interest concerning both content areas. Whereas the interest type ‘Physics? Only in certain contexts!’ is universal, the second type of interest differs for the two content areas. The second type of interest in mechanics comprised students who were interested in physics relating to the motion of cars and accordingly named; and the second type of interest in particle physics is named ‘particle physics lovers’. Only a small fraction of students (60, i.e. 5%) were categorised into the type interested in physics relating to the motion of cars concerning mechanics and into the particle physics lover type concerning particle physics. This indicates that these interest types were rather content and context specific, in contrast to the content overarching type of students that were only interested in physics set in certain contexts.

Table 26. Categorisation of students into different groups of interest for mechanics and particle physics and overarching interest types, and number of students assigned to the same type for both content areas.

Interest types	Mechanics *Groups (count, column %)*	Particle physics *Groups (count, column %)*	Students Count (total %)
Physics? Only in certain contexts! *(M_1 and PP_1)*	1_{M}, 2_{M}, and 3_{M} (1030, 86%)	1_{PP} and 2_{PP} (937, 79%)	840 (71%)
Students interested in physics relating to the motion of cars *(M_2)*	4_{M} (157, 14%)	-	60 (5%)
Particle physics lovers *(PP_2)*	-	3_{PP} (250, 21%)	

My description of two different interest types concerning both content areas based on the groups unrevealed by the mixed Rasch analysis, is supported by the findings of the past empirical studies (Rost et al., 1999). The students’ interest type ‘Physics? Only in certain contexts’ concerning both content areas is in line with type C, which describes students who are highly interested in physics as it relates to humans, nature, applications, and society, found in the IPN study (Langeheine et al., 2001; Rost et al.,1999; Sievers, 1999); on category A found in the HOPE study, which describes students’ ‘curiosity to understand the world, natural phenomena and the universe’ (Levrini et al., 2017); and on the interest category ‘living systems’ found by Drechsel et al. (2011) for PISA 2006 data. The type of students only interested in physics set in certain contexts comprised slightly more female than male students.

Concerning the content area ‘mechanics’, the type interested in physics relating to the motion of cars is only somewhat in line with the type of students that are generally and highly interested in physics, referred to as type A in the IPN study (Langeheine et al., 2001; Rost et al.,1999; Sievers, 1999). In line with IPN type A, these students were rather equally interested in all items. Yet, they did not have the highest mean interest in mechanics, which does not align with the IPN interest type A. However, it somewhat resembles the group of students referred to as ‘practical inquirers’ (1.5%) by Radišić et al. (2021) resulting from their analysis of PISA 2015 data obtained from students in Italy. These ‘practical inquirers’ had a high self-efficacy and instrumental motivation, engaged in leisure science activities but had the lowest interest in science topics. Similarly, the type of students interested in physics relating to the motion of cars found in my study had the highest mean physics-related self-concept. In addition, these students were slightly more interested in mechanics contents set in the context ‘car’ than in other contexts. This is in line with the interest category ‘physical/technology systems’ found by Drechsel et al. (2011) for PISA 2006 data. Here, Olsen and Lie (2011) found that within the cluster ‘physical/technological systems’, the items related to fires and pure or frontier science are crucial. Here, I argue that fires and car accidents may arouse similar emotions in students as both are suggesting danger in relation to one’s own body.

Concerning the content area ‘particle physics’, the particle physics lover type of students is in line with the type of students that are generally and highly interested in physics, referred to as type A in the IPN study (Langeheine et al., 2001; Rost et al.,1999; Sievers, 1999). However, not all items were equally and highly interesting for the particle physics lover type of students. For example, the item presenting quantitative particle physics without any context was rated relatively less interesting. This does not align with the IPN type A. In addition, the particle physics lover type of students were mostly interested in the special aspects of particle physics as a scientific endeavour. This aligns with the category B found in the HOPE study, which describes students’ ‘interest in physics knowledge as a special way of knowing, investigating, questioning and thinking’ (Levrini et al., 2017); and with the interest category ‘physical/technology systems’ found by Drechsel et al. (2011). Again, Olsen and Lie (2011) found that within the cluster ‘physical/technological systems’, the items related to fires and pure or frontier science are crucial. Here, I argue that particle physics may be a prime example for pure or frontier science. The particle physics lover type of students also somewhat resembles the group of students referred to as ‘scientists’ (12%) by Radišić et al. (2021) resulting from their analysis of PISA 2015 data obtained from students in Italy. These ‘scientists’ had the highest mean self-efficacy, instrumental motivation, enjoyment of and interest in school science and in leisure science activities. The particle physics lover type of students also had the highest mean interest in particle physics. This also aligns with the finding by Habig et al. (2018) that students with high individual interest express higher interest in unique contexts. Here, I argue that particle physics as a scientific endeavour resembles a unique context as used in Habig et al.’s study (2018) well. In addition, the particle physics lover type of students were mostly male. This is also in line with past empirical finding from the ASPIRES study on students aged 12 to 14 years. Here, notably fewer female (37%) than male (63%) students were classified as being ‘science keen’ (Archer et al., 2013).

The resulting description of the students’ types of interest in mechanics and particle physics, respectively, does not fully align with my hypothesis that the IPN interest types (A, C, NG; Langeheine et al., 2001; Rost et al.,1999; Sievers, 1999) are still valid for today’s students and for mechanics and particle physics. First, in the IPN study all groups resulting from the mixed Rasch rating scale model were described in terms of interest types. In contrast, I considered the item hierarchies of the different groups when describing interest types. Here, my focus was on the hierarchy of interesting contexts presented in the items. My idea was to outbalance the effect of the students’ different response styles because the mixed Rasch model differentiates between different groups not only because of qualitatively different item hierarchies (i.e. qualitative differences in the item hierarchies) but also because of different response styles. Given an equivalence of their item hierarchies (i.e. given the absence of qualitative differences in item hierarchies), I combined different groups to form interest types. Only if a group had a unique item hierarchy, I considered this group as a separate interest type. Using this approach, I could replicate the IPN interest type C (Langeheine et al., 2001; Rost et al.,1999; Sievers, 1999) concerning both content areas. However, using my approach I could not replicate the IPN interest type NG (Langeheine et al., 2001; Rost et al., 1999; Sievers, 1999). Concerning the type NG, Sievers (1999) had already concluded that (1) it was rather interpreted based on structural than on content-related considerations and (2) it could have been excluded from further analyses. Here my study provides a different approach than suggested by Sievers (1999), that is, to consider all students in further analyses. Moreover, I could only partially replicate the IPN interest type A (Langeheine et al., 2001; Rost et al.,1999; Sievers, 1999) separately for the content areas 'mechanics' and 'particle physics'. That is, type A is not universal, but depends on the specific content area. Overall, in line with my hypothesis I found that within the different types of interest, different contexts are more (or less) interesting relative to each other. This might also provide evidence aligning with my hypothesis that the interest types comprise groups of students on different levels of interest, each associated with different aspects of physics. Further discussions on whether these types of interest in mechanics can be described in terms of a hierarchy of students' levels of interest are presented in **Section 6.3.3.**.

6.3.3. The hierarchy of levels of interest in physics (HOLIP)

In this section the results related to my research question about the applicability of the hierarchy of levels of interest in particle physics to the different types of interest in mechanics and particle physics are presented and discussed. RQ6 asked '*To what extent can students' types of interest be described in terms of a hierarchy of students' levels of interest?*'. I introduced a conceptualisation of students' interest in particle physics comprising three levels of interest to describe the item hierarchy in my first study based on the storyline of a learning activity, that is, the context in which physics contents may be set (see **Section 5.3.3.**). I characterised the level of focused interest as being interested in particle physics when it is set in a context that is related to (1) one's own body, for example, 'medical diagnostics' (PP07, medical diagnostics); (2) socio-scientific issues, for example, 'smuggled arms' (PP03, particle detectors and smuggled arms); or (3) existential questions of humankind, for example, 'the big bang' (PP02, particle physics and the big bang). The level of open interest comprises students that are additionally interested in particle physics set in the broad context 'everyday life'. Finally, I defined the level of broad interest as being interested in particle physics, even when it is set in a technical or purely scientific context. Here, scientific context means that the item only presents particle physics as a scientific content area or refers to particle physics as a scientific endeavour.

Results

To investigate to what extent students' types of interest can be described in terms of such a hierarchy of levels of interest, I conducted an analysis of the item hierarchies of the different interest types, that is, I examined the item measures and wording. If the instrument to measure mechanics interest and the IPPI, respectively, display the proposed hierarchy of levels of interest for an interest type, item measures should depend on the context that the item in question represents. Items set in an everyday context, such as the human body, should be among the most interesting, whereas those set in a purely scientific context, such as qualitative or quantitative science, should be among the least interesting.

For mechanics, I found that in the groups 1_M, 2_M, and 3_M had an equivalent item hierarchy, and hence I combined them into the type of interest in mechanics M_1. I found that the conceptualisation of students' interest as a hierarchy of levels of interest, originally developed for particle physics, can also be applied to describe students' interest in mechanics. Overall, the conceptualisation suits to describe the item hierarchies of the groups 1_M, 2_M, and 3_M. These groups were jointly described as the type M_1, that is, students who are only interested in physics set in certain contexts. The most interesting items present mechanics in contexts related to one's own body (M07: artificial organs and joints in medicine;

M09: designing a safety car for severe accidents) and socio-scientific issues (M03: increase of probability and consequences of a car accident with speed; M09: designing a safety car for severe accidents; M11: accident statistics and speed limits). Here, the context ‘car accidents’ is predominant, and I argue that it may arouse emotions because it suggests danger. This potential arousal of emotions caused by the contexts is in line with the characterisation of the level of focused interest. However, the aspect ‘existential questions of humankind’, that is also part of the level of focused interest in the conceptualisation, was not tackled in the mechanics items, and thus, it is not possible to tell whether this aspect would also have been very interesting for the students in relation to the content area mechanics. Similarly, the context ‘everyday life’, that defines the level of open interest in the conceptualisation, was mostly tackled by referring to cars, which I consider to be a very specific context. Thus, it is not possible to tell whether the broad field of ‘everyday life’ mentioned in the conceptualisation would also have been very interesting for the students in relation to the content area mechanics. The focus on cars as a context might reinforce student conceptions that physics is about ‘using’ and ‘dominating’ nature instead of understanding it (Muckenfuß, 1995). Nowadays, the context ‘car’ might be less interesting for students compared to the 1980s considering that climate change is a very concerning and omnipresent topic (e.g. ‘Fridays For Future’ movement). Hence, I argue that it is not surprising that the context ‘car’, which does not question the use of individual transport or the emission of greenhouse gases nor relates to accidents (i.e. items M04: transformation of energy of motion into other forms of energy; M05: calculating the energy of motion based on the speed of a car; M06: force-saving devices in a car repair shop; M10: calculating the braking path based on the speed of a car), was the least interesting for the students. Similarly, the context ‘crude oil’ was – not surprisingly – one of the least interesting for the students. Hence, I argue that the contexts ‘car’ and ‘crude oil’ are rather resembling the aspect ‘technology’ (than the aspect ‘everyday life’) mentioned in the conceptualisation. It is well in line with the conceptualisation that these ‘technology’ contexts were the least interesting for the students. However, the context ‘science’ was not tackled in the mechanics questionnaire as there is no item that presents the context ‘science’; that is, there was no item that only presents mechanics as a scientific content area or refers to mechanics as a scientific endeavour. Thus, it is not possible to tell whether the context ‘science’ would also have been very interesting for the students in relation to the content area mechanics. Overall, the conceptualisation of students’ interest as a hierarchy of levels of interest fits well to describe the item hierarchies of the type M_1, that is, the groups 1_M, 2_M, and 3_M. Yet, some contexts (e.g. existential questions, everyday life, science) that are mentioned in the original conceptualisation were not addressed in the mechanics items.

I found that for the students of the type M_2, that is, the group 4_M, all items were rather similarly interesting. Yet, the items M07 (artificial organs and joints in medicine), M08 (constructing different pulleys and trying them out), and M11 (accident statistics and speed limits) were the least interesting, although they are very interesting for the other groups. However, the relatively most interesting items were set in contexts related to accidents and cars (M01: devices that amplify forces (e.g. pulley, lifting platform); M03: increase of probability and consequences of a car accident with speed; M04: transformation of energy of motion into other forms of energy; M05: calculating the energy of motion based on the speed of a car; M07: artificial organs and joints in medicine; M10: calculating the braking path based on the speed of a car). The items M02 (pumping oil from great depths, context ‘crude oil’) and M06 (force-saving devices in a car repair shop, context ‘car repair shop’) were less interesting. It is in line with the conceptualisation that the context ‘car accidents’ was very interesting as it resembles the aspect ‘socio-scientific issues’ but also the aspect ‘one’s own body’, both of which may arouse emotions. The items M08 (constructing different pulleys and trying them out) and M11 (accident statistics and speed limits) were relatively uninteresting, which is not in line with other past empirical studies indicating that experiments and discussions are interesting for students. However, this is similar to group 2_M and 3_M, respectively (both described as type M_1), for which M11 (accident statistics and speed limits) and M08 (constructing different pulleys and trying them out), respectively, are also rather uninteresting. In all, I found that the conceptualisation of interest does not fit to describe these students and that they were interested in physics relating to the motion of cars since the context 'car' was the most interesting for them, not only when referring to accidents.

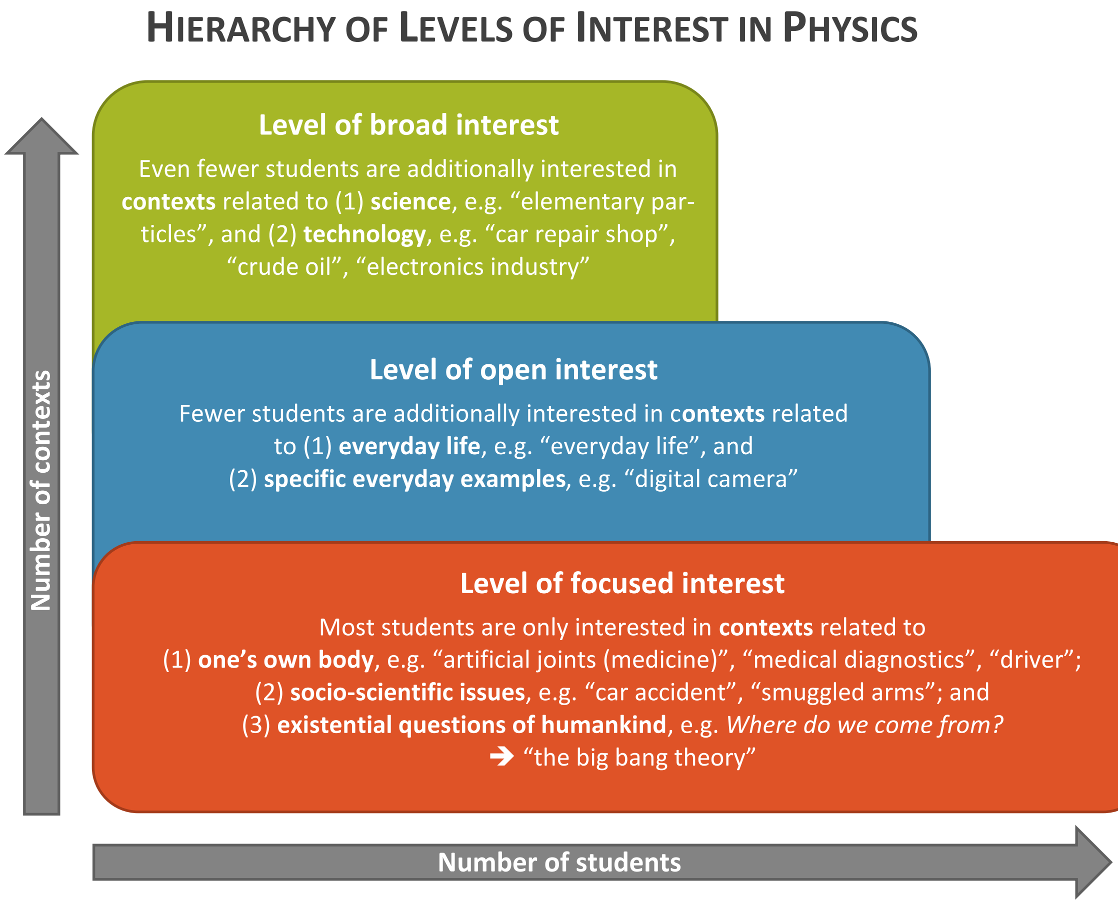
For particle physics, I found that the conceptualisation of students' interest as a hierarchy of levels of interest also perfectly suits to describe the item hierarchies of the type PP_1, that is, the groups 1_{PP} and 2_{PP}. For these students, the four items PP02 (particle physics and the big bang), PP03 (particle detectors and smuggled arms), PP07 (medical diagnostics), and PP08 (transforming a mobile phone into a particle detector) were the most interesting, which form the level of focused interest in the original conceptualisation; the three items PP01 (devices that detect particles (e.g. digital camera)), PP09 (planning an experiment on particle acceleration), and PP11 (particle physics has changed our daily life) were in the middle of the item hierarchy, which together with the previous items formed the level of open interest; and the items PP04 (particles in the nucleus space of an atom), PP05 (masses of particles), PP06 (particle accelerators in the electronics industry), and PP10 (calculating the energy of a particle collision at the speed of light) were the least interesting, which formed the level of broad interest together with all other items in the first study.[29] Hence, I found that the interest type PP_1, that is, most students, can be described with my conceptualisation of interest.

I found that my conceptualisation of interest does not describe the item hierarchy of group 3_{PP}, that is, the physics lover type of students (PP_2). Most students of this type were interested in the special aspects of particle physics as a content area and a scientific endeavour, that is, its focus on the existential questions of humankind (PP02: particle physics and the big bang; PP04: particles in the nucleus space of an atom) and on experiments (PP08: transforming a mobile phone into a particle detector; PP09: planning an experiment on particle acceleration; PP10: calculating the energy of a particle collision at the speed of light). It is in line with the conceptualisation (which is based on the hierarchy of interesting contexts) that most students of this type were interested in contexts related to the existential questions of humankind. However, items presenting physics set in contexts related to socio-scientific issues, everyday life, technology, and one's own body were less interesting for these students, which contrasts with my conceptualisation.

In sum, I found that the item hierarchy of the type of students that is only interested in physics set in certain contexts, that is, types M_1 and PP_1 comprising most students' interest in mechanics and particle physics, respectively, can be described with my conceptualisation of interest as a hierarchy of levels of interest. **Figure 18** shows the conceptualisation comprising three levels of interest adapted to both content areas mechanics and particle physics.

---

[29] Concerning the items PP08 and PP09, I described the effect of the task in the first study; that is, experiments as generally interesting, and doing the experiment (PP08) as more interesting than planning an experiment (PP09).



Level of broad interest

Even fewer students are additionally interested in **contexts** related to (1) **science**, e.g. “elementary particles”, and (2) **technology**, e.g. “car repair shop”, “crude oil”, “electronics industry”

Level of open interest

Fewer students are additionally interested in **contexts** related to (1) **everyday life**, e.g. “everyday life”, and (2) **specific everyday examples**, e.g. “digital camera”

Level of focused interest

Most students are only interested in **contexts** related to (1) **one’s own body**, e.g. “artificial joints (medicine)”, “medical diagnostics”, “driver”; (2) **socio-scientific issues**, e.g. “car accident”, “smuggled arms”; and (3) **existential questions of humankind**, e.g. *Where do we come from?* ➔ “the big bang theory”

Figure 18. Conceptualisation of interest in physics as a hierarchy of three levels of interest in physics (HOLIP). The horizontal axis represents the number of students. The vertical axis represents the number of contexts. Note that the levels are cumulative, that is, the level of broad interest includes the level of open interest which in turn includes the level of focused interest.

Discussion

I found that the item hierarchy of the type of students that is only interested in physics set in certain contexts, that is, types M_1 and PP_1 comprising most students’ interest in mechanics and particle physics, respectively, can be described with my conceptualisation of interest. The applicability of the conceptualisation to particle physics type PP_1 shows that (1) already based on the small sample of the first study ($N = 99$) valid conclusion about the students’ interest were made; and (2) based on a large sample ($N = 1214$) groups with differing item hierarchies, such as group 3_{PP}, can be successfully unveiled when using a mixed Rasch model. Moreover, past empirical studies may explain why I could apply the conceptualisation of students’ interest as a hierarchy of levels of interest – consistently with the results of my first study – to particle physics type PP_1 and why I could additionally apply the conceptualisation to mechanics type M_1 in the second study. As argued above, Fechner et al. (2015) showed that the effect of a context on interest is more distinct, if it can be distinguished clearly from the content and that the positive effect of a context on interest is higher, the less interesting the content itself is. Based on this, I argue that it may be that for this first type of students (type M_1 and PP_1, respectively, called ‘Physics? Only in certain context!’), (1) there was a clear distinction between the content area ‘mechanics’ and ‘particle physics’, respectively, and the contexts in which it was set; and (2) the content area 'mechanics' and 'particle physics', respectively, was less interesting than the contexts in which it was set.

In contrast, I found that only for the interest types M_2 and PP_2, representing a very small fraction of students (14% in mechanics and 21% in particle physics), the above-described conceptualisation does not fit. For mechanics, the second type of interest (M_2) was mostly interested in the context 'car', even when not referring to accidents, called 'students interested in physics relating to the motion of cars'. For particle physics, the second type of interest (PP_2) was mostly interested in the special aspects of particle physics as a content area and a scientific endeavour, called 'particle physics lovers'. The conceptualisation could not be applied to these second types of interest. The difference in applicability of the second interest type (M_2/PP_2) compared to the first interest type (M_1/PP_1) in mechanics and particle physics, respectively, may also be explained based on past empirical studies (Fechner et al., 2015). That is, it may be that there was no clear distinction between the content area 'mechanics' and the context 'car' for the students forming type M_2. Similarly, for the students forming type PP_2, it may be that the content area 'particle physics' could not be distinguished clearly from the contexts in which it was set and/or that the content area 'particle physics' was more interesting than the contexts in which it was set. Hence, the positive effect of contexts are less pronounced (Fechner et al., 2015).

In sum, the conceptualisation of interest as a hierarchy of students' levels of interest, originally introduced in the first study regarding particle physics, was successfully applied to describe the relative interestingness of different contexts for the first type of students (called 'Physics? Only in certain contexts!') regarding both content areas, that is, type M_1 (84% of the students regarding mechanics) and type PP_1 (79% of the students regarding particle physics). Hence, the conceptualisation is renamed 'hierarchy of students' levels of interest in physics' (HOLIP, see **Figure 18**) and suggested as a guideline for physics in general, if the specific physics content area can be set in the recommended contexts. It comprises three levels of interest: I characterised the level of focused interest in particle physics as being interested in particle physics when it is set in a context that is related to one's own body, socio-scientific issues, or existential questions of humankind. I found that aspect 'existential question of humankind' caused a high interest in particle physics, although this context is theoretical, like purely scientific contexts. I believe that these three aspects can be assigned to the same level of interest because they all sparked interest by arousing positive or negative emotions. This aligns with the first phase in interest development, 'triggered situational interest', in which the emotional component prevails (Hidi & Renninger, 2006). Second, students at the level of open interest were additionally interested in physics when it was set in the broad context of 'everyday life'. My definition of the level of open interest aligns with the second phase of interest development, namely 'maintained situational interest' as proposed by Hidi and Renninger (2006). In this phase, a person begins to recognise personal value based on already existing positive feelings, that is, the value-related component of interest prevails. Third, I defined the level of broad interest in particle physics as being interested in physics, even when it is set in a purely scientific or technical context. My definition of the broad level aligns with the third and fourth phases in interest development, namely 'emerging' and 'well-developed individual interest', where the cognitive-epistemic component of interest prevails (Hidi & Renninger, 2006). When interpreting the hierarchy of students' levels of interest, it is crucial to consider that these levels are cumulative. This means that the level of broad interest includes the level of open interest, which further includes the level of focused interest. Moreover, previous studies (e.g. OECD, 2007a, 2016) have demonstrated that students' interest in physics content is low compared to, for example, biology and chemistry content.

Hence, based on Hidi and Renninger's 'Four-phase-model of interest development' (2006), one could argue that many students have not developed a more stable form of interest in physics content. Consequently, when aiming to foster interest in physics, many students might benefit from activities targeting the emotional component of interest, whereas only a few students might benefit from activities targeting the value-related component, and even fewer students might appreciate activities targeting the cognitive-epistemic component. This is also in line with the hierarchy of students' levels of interest in physics (HOLIP).

6.3.4. Which physics content area is more interesting overall?

In this section the results related to my research question about the interest in the two different content areas are presented and discussed. RQ7 asked *'In which physics content area are the students overall more interested, mechanics or particle physics?'*. The hypothesis for this research question was that modern physics content areas are more interesting than classical ones as suggested by past empirical studies (e.g. Häußler, Lehrke, et al., 1998; OECD, 2016; Sjøberg & Schreiner, 2012).

Results

I analysed which content area mechanics or particle physics is overall more interesting by investigating the individual differences in Rasch person measures and one-sample t-tests. **Table 27** lists the means of the individual differences in Rasch person measures (interest in particle physics minus interest in mechanics). One can see that the differences were small.

Table 27. Mean differences in Rasch person measures in logit units (interest in particle physics minus interest in mechanics), corresponding standard deviations, and minimum and maximum values of the individual differences.

Mean of the individual differences in Rasch person measures *Interest in particle physics minus interest in mechanics*		
Mechanics interest type	M_1	M_2
Mean (Standard Deviation)	0.03 (0.85)	0.41 (0.91)
Min – Max	-3.09 – 3.21	-2.07 – 3.13
Particle physics interest type	PP_1	PP_2
Mean (Standard Deviation)	-0.06 (0.81)	0.60 (0.89)
Min – Max	-3.09 – 2.79	-1.96 – 3.21
Full sample		
Mean (Standard Deviation)	0.07 (0.87)	
Min – Max	-3.09 – 3.21	

To check whether the means of the individual differences in Rasch person measures are significantly different from 0, I conducted a one-sample t-test. First, I conducted this analysis for the full sample. The mean difference (M=0.0748, 95% CI = [0.0209, 0.1287]) was significantly higher than 0, t(1000)=2.722, p=0.007. However, the effect size is too low to be interpretable in a sound manner (d_C=0.09, 95% CI [0.024, 0.148]).

Then, I conducted this analysis separately for the mechanics interest types. For type M_1, the mean difference (M=0.0266, 95% CI = [0.8526, 0.0288]) was not significantly different from 0, t(874)=0.925, p=0.355. For type M_2, the mean difference (M=0.4090, 95% CI = [0.2482, 0.5698]) was significantly higher than 0, t(125)=5.034, p=<0.001. This difference translates into a small effect (d_C=-0.45, 95% CI [0.264, 0.631]).

Finally, I conducted this analysis separately for the particle physics interest types. For type PP_1, the mean difference (M=-0.0575, 95% CI = [-0.1141, -0.0013]) was significantly lower than 0, t(799)=-2.008, p=0.045. However, the effect size is too low to be interpretable in a sound manner (d_C=-0.07, 95% CI [-0.140, -0.002]). For type PP_2, the mean difference (M=0.6020, 95% CI = [0.4784, 0.7256]) was significantly higher than 0, t(200)=9.606, p=<0.001. This difference translates into a medium-sized effect (d_C=0.68, 95% CI [0.524, 0.830]).

The one-sample t-test showed that for all students the mean differences in interest were significantly different from 0 with particle physics being more interesting than mechanics. Yet, the effect size was too small to be interpreted in a sound manner. For mechanics, there was no significant difference for type M_1 students (86% of students, ‘Physics? Only in certain contexts!’) but for the type M_2 students (14% of students, ‘students interested in physics relating to the motion of cars’) with particle physics being more interesting translating into a small effect. For particle physics, there was a significant difference for type PP_1 students (79% of students, ‘Physics? Only in certain contexts!’) with mechanics being more interesting. Yet, the effect size was too small to be interpreted in a sound manner. For type PP_2 students (21% of students, ‘particle physics lovers’), particle physics was significantly more interesting than mechanics translating into a medium-sized effect.

Discussion

In sum, I argue that for most students both content areas were equally interesting considering the extremely low effect sizes. However, for those students assigned to the type M_2 (14% of students, ‘students interested in physics relating to the motion of cars’) and the type PP_2 (21% of students, ‘particle physics lovers’), particle physics was found to be more interesting than mechanics. I argue that particle physics was more interesting than mechanics for the students assigned to these interest types (M_2 and PP_2) because they have a higher mean physics-related self-concept (further discussions on the association between self-concept and interested are presented in **Section 6.3.6.**). These findings show that for most students both content areas are equally interesting. Only for a small fraction of students (depending on the applied typology) particle physics was more interesting than mechanics. This finding contrasts with our hypothesis that modern physics contents are more interesting than classical ones as suggested by past empirical studies which show that modern physics contents, such as the history of the universe, are more interesting than classical ones (e.g. Häußler, Lehrke, et al., 1998; OECD, 2016; Sjøberg & Schreiner, 2012). It provides evidence that such generalised observations may result from the differences in interest of a small fraction of students and cannot be applied to most students. Most importantly, these findings emphasise the importance of analysing data concerning students’ interest using methods, such as mixed Rasch analyses, that do not rely on pre-defined categorisations (e.g. sex) but instead distinguish groups based on their responses to the interest instrument used in the study.

6.3.5. Students' types of physics-related self-concept

In this section the results concerning my research question about the students' categorisation into different types of physics-related self-concept are presented and discussed. RQ8 asked *'To what extent can the students be categorised into different types of physics-related self-concept?'*. In my study physics-related self-concept was defined as comprising both aspects, self-perception of ability and beliefs about the perception as a physics person by others. The hypothesis for this research question was that the students cannot be categorised into different types of physics-related self-concept. This was based on the physics identity studies which suggested that the beliefs about the perception as a physics person by others (perceived recognition) and the self-perception of ability (competency beliefs) are strongly associated with one another (Godwin et al., 2016; Kalender et al., 2019b).

Results

To investigate the students' physics-related self-concept I conducted the same mixed Rasch analyses as for students' interest in mechanics and in particle physics. I calculated and compared four different models, from a model with only one group[30] of students, that is, the sample is not divided into groups of students, to a model with four groups of students with different self-concept profiles. The BIC values of the four calculated models with respect to the model with the lowest BIC value are listed in **Table 28**. The BIC value of the 3-groups-model was the smallest. The delta-BIC values indicated very strong evidence against the 1-group-model, and strong evidence against the 2- and 4-groups-models.

Table 28. Delta BIC-values of the four calculated models to describe students' physics-related self-concept

	Delta-BIC			
Number of groups	1	2	3	4
Delta-BIC	1116	104	0	61

Qualitative analysis of the groups confirmed that a distinction into three different groups is sufficient to describe students' physics-related self-concept. The distinction into more groups did not improve the description of students' physics-related self-concept because then the groups' sizes would be very small and several groups would have very similar interest profiles. Consequently, the students were categorised into three different groups of physics-related self-concept. Group 1_{SC} consisted of 49% of the sample, group 2_{SC} of 41%, and group 3_{SC} of 10%.[31] **Figure 19** shows the self-concept profiles of the three groups 1_{SC}, 2_{SC}, and 3_{SC}. The items are shown on the x-axis and the threshold measures of each item are shown on the y-axis. Since there are four rating scale categories, there are three thresholds, and hence three threshold measures for each item. The lower the threshold measures of an item are, the more agreement was expressed by the students. The line between the data points does not have a meaning but helps to observe some patterns.

---

[30] The term 'group' does not refer to a 'self-concept type' (as mentioned in RQ8). A group may indeed form a type or several groups combined may together form a type.

[31] The index SC stands for physics-related Self-Concept.

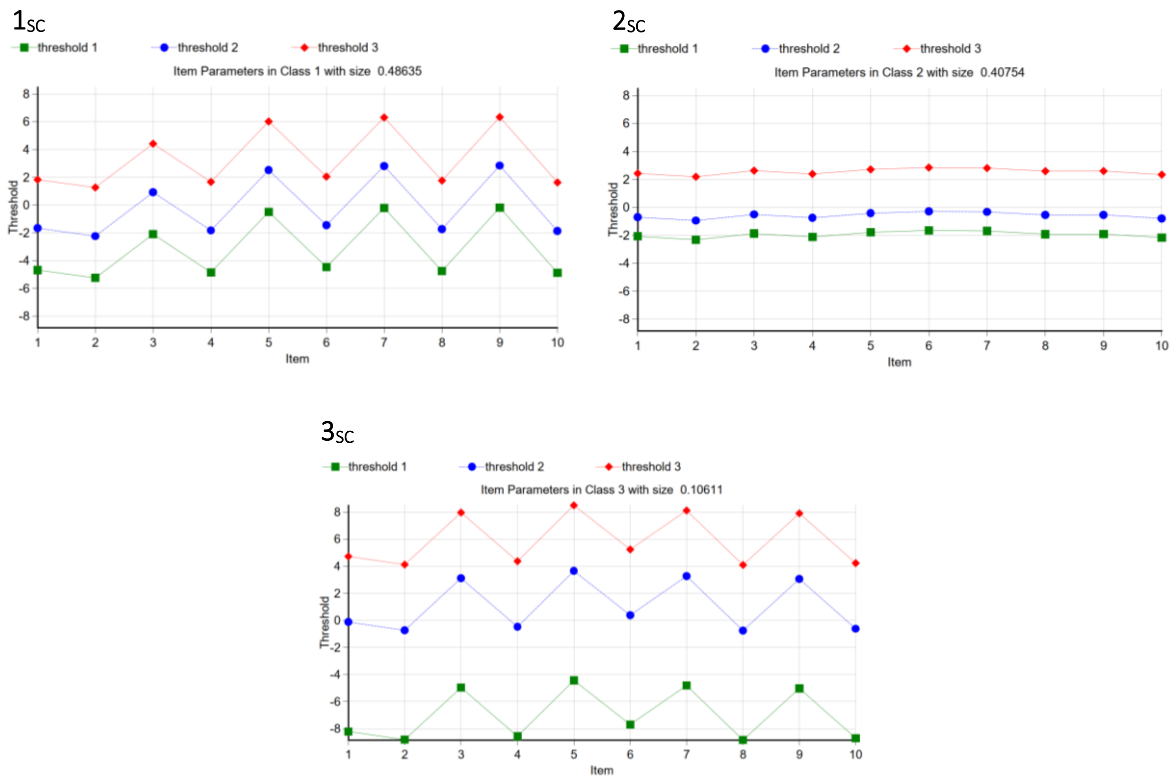


Figure 19. Self-concept profiles, that is, how much agreement was expressed by the students to the different items relative to each other, of the three groups 1_{SC}, 2_{SC}, and 3_{SC}.
Note: On the x-axis, the items are shown, and on the y-axis, the threshold measures are shown in logit units, each marking the threshold between one rating scale category to another. For example, threshold 1 is between the categories '*fully disagree*' and '*rather disagree*'. The line between the data points does not have a meaning but helps to observe some patterns.

Firstly, one can see that the self-concept profiles, that is, how much agreement was expressed by the students to the different items relative to each other, of the three groups 1_{SC}, 2_{SC}, and 3_{SC} differ. Students assigned to[32] the groups 1_{SC} and 3_{SC} expressed more agreement to items SC01 (Learning advanced physics topics would be easy for me), SC02 (I can usually give good answers to test questions on physics topics), SC04 (I can learn physics topics quickly), SC06 (Physics topics are easy for me), SC08 (When I am being taught physics, I can understand the concepts very well), and SC010 than to the items SC03 (My TA or Instructor see me as physics person), SC05 (My parents see me as physics person), SC07 (My classmates see me as physics person), and SC09 (My friends see me as physics person). The distance between the thresholds is smaller in group 1_{SC} than in group 3_{SC}. However, for group 3_{SC} the distance between thresholds 1 and 2 is bigger than between thresholds 2 and 3. Group 2_{SC} students expressed similar agreement to all items. The distance between the thresholds 1 and 2 is smaller than between the thresholds 2 and 3.

---

[32] The students are assigned to the group, for which they have the most probable assignment; that is, the assignment is based on the highest probability. In the following, when writing 'assigned to' or 'categorised into' the most probable group assignment is meant.

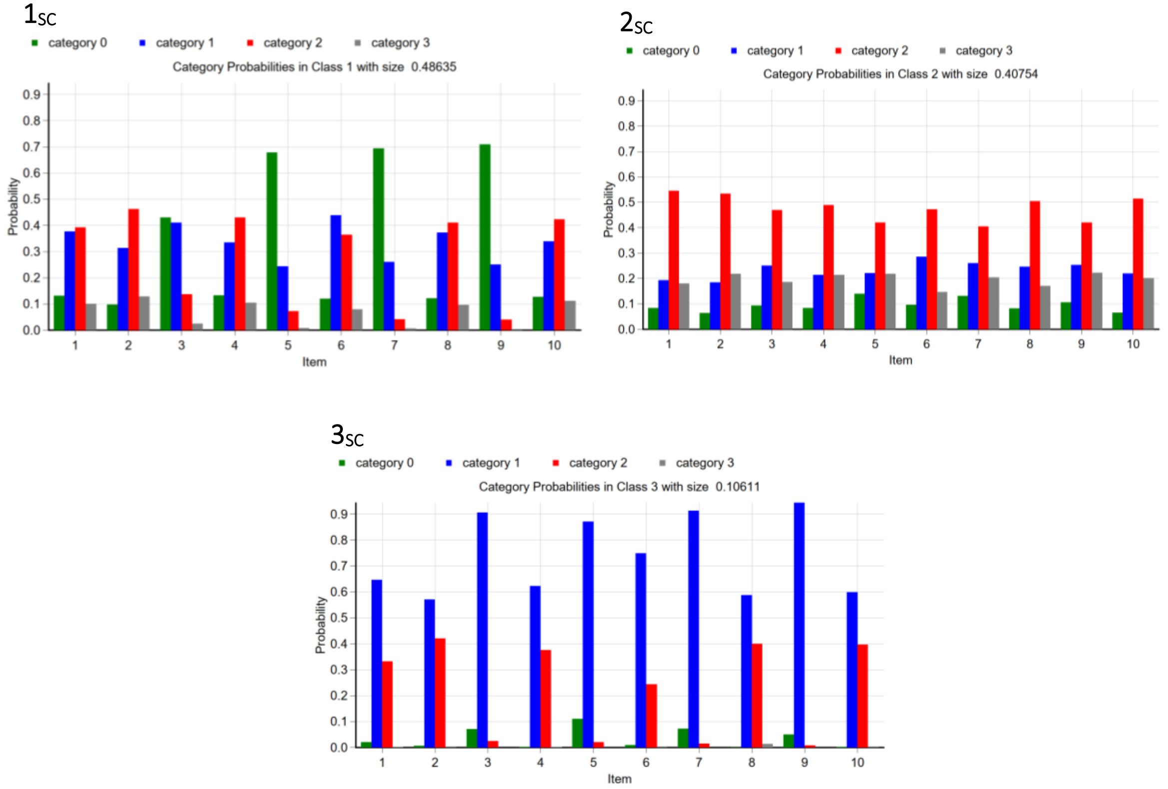


Figure 20. Category probabilities, that is, the probabilities to choose the different categories of the rating scale, of the three groups 1_{SC}, 2_{SC}, and 3_{SC}.
Note: The students expressed their degree of agreement in each item on a 4-category rating scale (*fully agree* (=3), *rather agree* (=2), *rather disagree* (=1), or *fully disagree* (=0)).

To study the difference between groups 1_{SC}, 2_{SC}, and 3_{SC} I analysed how the students used the different answer categories of the rating scale. **Figure 20** shows the category probabilities of the three groups 1_{SC}, 2_{SC}, and 3_{SC}. One can see that the groups differ in their use of the rating scale categories. Group 1_{SC} students tended to use all the categories of the rating scale, especially the two middle categories, but for the items SC03 (My TA or Instructor see me as physics person), SC05 (My parents see me as physics person), SC07 (My classmates see me as physics person), and SC09 (My friends see me as physics person) rather the category '*fully disagree'* (=0). Similarly, group 3_{SC} students tended to use the two middle categories of the rating scale but for the items SC03 (My TA or Instructor see me as physics person), SC05 (My parents see me as physics person), SC07 (My classmates see me as physics person), and SC09 (My friends see me as physics person) only the category '*rather disagree'* (=1). In contrast, group 2_{SC} students used the categories of the rating scale equally across all items, and they tended to use the category '*rather agree'* (=2).

When conducting (mixed) Rasch analyses, an evaluation of the obtained item and person measures is crucial as it provides further evidence for the quality of the chosen model. I evaluated the Rasch item measures in physics-related self-concept. When checking for item fit for the self-concept groups, I found that all items fit well in group 1_{SC} (*ZQ* values ranging from -1.67 to 1.58) except for the item SC05 (*ZQ* = -2.63). I argue that this is because the item SC05 (My parents see me as physics person) was one of four items reflecting the aspect 'perceived recognition' (by the parents), and hence, might have appeared to be less related to the measured latent trait, that is, physics-related self-concept, or to relate to different aspects of it. All items fit well in group 2_{SC} (*ZQ* values ranging from -1.40 to 1.89) and in group 3_{SC} (*ZQ* values ranging from -0.83 to 1.09). Based on the item fit statistics, I argue that the data supported one trait, physics-related self-concept, in all groups, that is, that the measurement instrument about students' physics-related self-concept was found to be unidimensional. Yet, physics-related self-concept is a broad trait because it covers both aspects, ability self-concept and perceived recognition. For the groups, these different aspects differed in their positive (or negative) effect on the student's physics-related self-concept. For group 2_{SC} students, both aspects were contributing in a similar positive (or negative) way to the students' physics-related self-concept. Yet, for group 1_{SC} and 3_{SC} students, the ability self-concept contributed more positively to the students' physics-related self-concept compared to their perceived recognition. Then, I evaluated the Rasch person measures for the students' self-concept. As for students' interest, I investigated the Rasch person fit and excluded those students from further analysis, whose newfit 1 and/or 2 value was $\ll -3$ or $\gg 3$ or for who no newfit values were calculated. Concerning physics-related self-concept, I excluded 73 students from further analysis.

To learn more about the students' different groups, I calculated corresponding descriptive statistics for the reduced sample size N_{red}. **Table 29** lists the descriptive statistics for the physics-related self-concept groups (1_{SC}, 2_{SC}, and 3_{SC}). The mean physics-related self-concept was the highest for group 2_{SC} students, followed by group 1_{SC} students, which in turn had higher mean interest than group 3_{SC} students. This can also be seen in **Figure 21** which shows the students' Rasch person measures indicating their degrees of physics-related self-concept for all self-concept groups. Group 2_{SC} had the highest mean interest in mechanics and particle physics, respectively. The mean interest in mechanics and particle physics, respectively, was rather similar for groups 1_{SC} and 3_{SC}. The groups 1_{SC} and 3_{SC} consisted of slightly more female students than male students, whereas it is the opposite for group 2_{SC}.

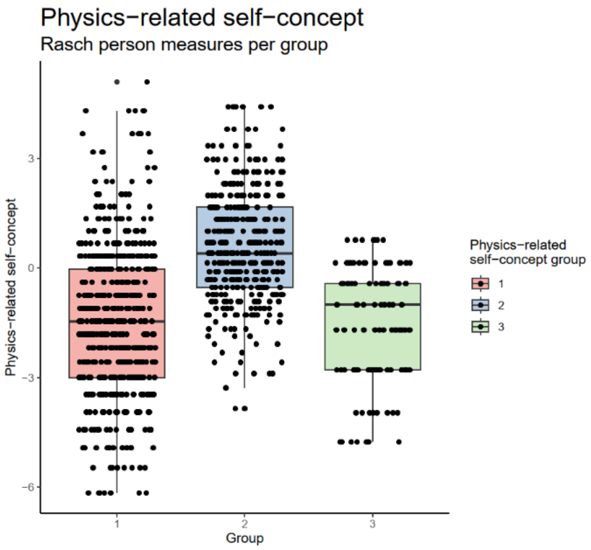


Figure 21. Boxplot with jitter of the students' Rasch person measures in logit units indicating their degrees of physics-related self-concept for all physics-related self-concept groups

Table 29. Person measure means and standard deviations, minimum and maximum values of the students' physics-related self-concept, interest in mechanics, interest in particle physics for all groups 1_{SC}, 2_{SC}, and 3_{SC}

Physics-related self-concept groups					
		Group	1_{SC}	2_{SC}	3_{SC}
Physics-related self-concept *Rasch person measures*		*Mean (Standard Deviation)*	-1.38 (2.07)	0.59 (1.56)	-1.50 (1.59)
		Min – Max	-6.16 – 5.10	-3.85 – 4.41	-4.76 – 0.76
Interest in mechanics *Rasch person measures*		*Mean (Standard Deviation)*	0.32 (0.90)	0.63 (0.94)	0.29 (0.88)
		Min – Max	-2.15 – 3.06	-1.88 – 3.71	-2.07 – 3.20
Interest in particle physics *Rasch person measures*		*Mean (Standard Deviation)*	0.35 (1.02)	0.78 (1.04)	0.28 (0.96)
		Min – Max	-2.92 – 3.42	-2.36 – 3.92	-2.36 – 3.42
Sex	Female	*Count (row %, column %)*	278 (51.5%, 55.3%)	194 (35.9%, 49.7%)	68 (12.6%, 63.0%)
	Male		225 (48.8%, 44.7%)	196 (42.5%, 50.3%)	40 (8.7%, 37.0%)

Note: The descriptive statistics presented in this table were calculated using the reduced sample size (N_{red} = 1001, see **Section 6.3.6.**). The Rasch person measures are listed in logit units. Lower and higher person measures represent persons with less and more self-concept/interest, respectively. The person measures are estimated based on the item measures in Rasch analysis.

To investigate the students' types of physics-related self-concept, I compared the item hierarchies of the three groups, in which the items are sorted by their item measure, that is, by how much agreement the students expressed to each item (**Table 30**). One can see that for groups 1_{SC} and 3_{SC}, there is a clear distinction between items that were easy to agree with, that is, items SC02 (I can usually give good answers to test questions on physics topics), SC10 (I can easily understand new ideas in physics), SC04 (I can learn physics topics quickly), SC08 (When I am being taught physics, I can understand the concepts very well), SC01 (Learning advanced physics topics would be easy for me), SC06 (Physics topics are easy for me), and items that were difficult to agree with, that is, items SC03 (My TA or Instructor see me as physics person), SC05 (My parents see me as physics person), SC07 (My classmates see me as physics person), and SC09 (My friends see me as physics person). In contrast, for group 2_{SC} all items were similarly easy to agree with. The distinction between items that were easy or difficult to agree with is less pronounced. Interestingly, item SC06 (Physics topics are easy for me) was the most difficult to agree with for these students.

I found that the students have been categorised into the different groups 1_{SC} and 3_{SC}, although their item hierarchies are equivalent; that is, the students express more (or less) agreement to the same items. Hence, I argue that 60% of the students can be described in terms of one type of physics-related self-concept SC_1 (comprising groups 1_{SC} and 3_{SC}). In contrast, students assigned to group 2_{SC} differ in their item hierarchy. These students can be described in terms of a second type of physics-related self-concept SC_2. I will discuss the similarities and differences regarding the item hierarchies of the different groups below.

Table 30. Item IDs, corresponding Rasch item measures (reflecting the expressed agreement), and standard errors (SE) for all three groups 1_{SC}, 2_{SC} and 3_{SC}.

Item	Group 1_{SC}	SE	Group 2_{SC}	SE	Group 3_{SC}	SE
SC09	3.00	0.11	0.05	0.08	1.98	0.34
SC07	2.97	0.11	0.26	0.08	2.19	0.34
SC05	2.68	0.1	0.16	0.08	2.57	0.33
SC03	1.08	0.09	0.08	0.08	2.04	0.34
SC06	-1.29	0.08	0.29	0.08	-0.69	0.25
SC01	-1.51	0.08	-0.13	0.08	-1.20	0.24
SC08	-1.57	0.08	0.04	0.08	-1.83	0.23
SC04	-1.67	0.08	-0.17	0.08	-1.55	0.23
SC10	-1.71	0.08	-0.21	0.08	-1.69	0.23
SC02	-2.08	0.08	-0.37	0.08	-1.81	0.23

Note: The Rasch item measures and standard errors are listed in logit units. Lower and higher item measures represent items that are more and less easy to agree with, respectively. The mean item measure of each group is set to zero in mixed Rasch analysis. The items are ordered according to their measures in group 1_{SC}. The colour scale represents the degree of agreement expressed in each item across all groups; dark blue is the item most easy to agree with and dark red is the item most difficult to agree with.

To learn more about the students' different types of physics-related self-concept, I calculated corresponding descriptive statistics for the reduced sample size N_{red}. **Table 31** shows the descriptive statistics for the physics-related self-concept types. The mean self-concept is lower for type SC_1 students than for type SC_2 students. This can also be seen in **Figure 22 (a)** which shows the boxplots of the students' Rasch person measures indicating their degrees of physics-related self-concept for both self-concept types. Type SC_2 consists of 50% male and 50% female students, and 36% of all female and 43% of all male students belong to type SC_2. The mean interest in mechanics and particle physics, respectively, was lower for type SC_1 students than for type SC_2 students. **Figure 22 (b)** shows the students' Rasch person measures indicating their degrees of physics-related self-concept for the different sexes. The difference in mean self-concept between the sexes is less pronounced than between the self-concept types (**Figure 22 (a)**).

In sum, type SC_2 (comprising group 2_{SC}) students had relatively higher physics-related self-concept and more interest in mechanics and particle physics, respectively, compared to type SC_1 (comprising groups 1_{SC} and 3_{SC}) students.

Table 31. Person measure means and standard deviations, minimum and maximum values of the students' physics-related self-concept for all types SC_1 and SC_2

Physics-related self-concept types				
		Type	SC_1	SC_2
		Group	1_{SC} and 3_{SC}	2_{SC}
Physics-related self-concept *Rasch person measures*		*Mean (Standard Deviation)*	-1.40 (2.00)	0.59 (1.56)
		Min – Max	-6.16 – 5.10	-3.85 – 4.41
Interest in mechanics *Rasch person measures*		*Mean (Standard Deviation)*	0.32 (0.90)	0.63 (0.94)
		Min – Max	-2.15 – 3.20	-1.88 – 3.71
Interest in particle physics *Rasch person measures*		*Mean (Standard Deviation)*	0.34 (1.00)	0.78 (1.04)
		Min – Max	-2.92 – 3.42	-2.36 – 3.92
Sex	Female	*Count (row %, column %)*	346 (64.1%, 56.6%)	194 (35.9%, 49.7%)
	Male		265 (57.5%, 43.4%)	196 (42.5%, 50.3%)

Note: The Rasch person measures are listed in logit units. Lower and higher person measures represent persons with less and more self-concept/interest, respectively. The person measures are estimated based on the item measures in Rasch analysis.

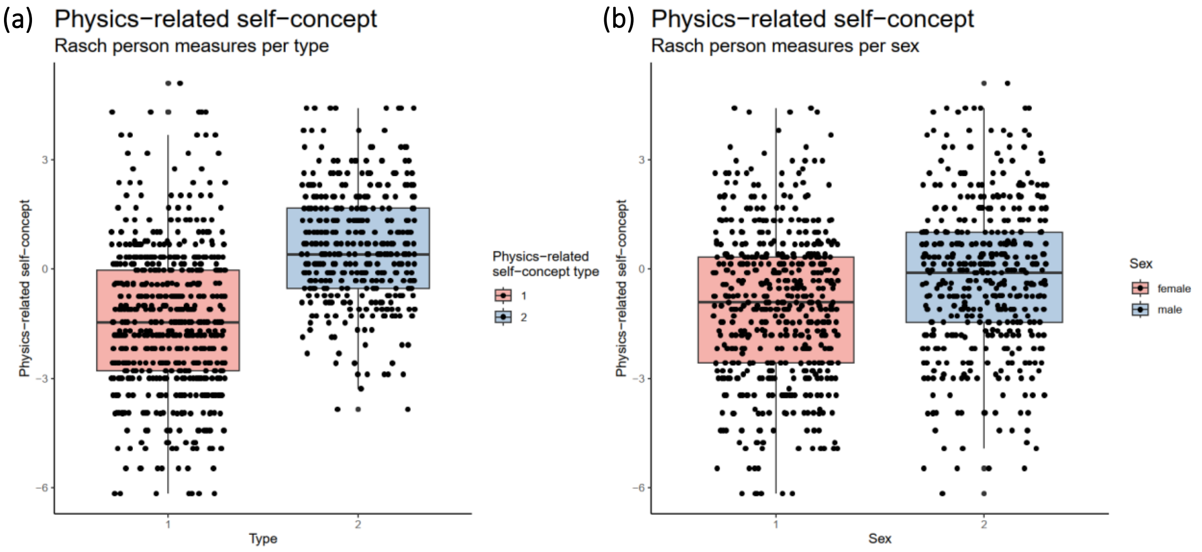


Figure 22. Boxplot with jitter of the students' Rasch person measures in logit units indicating their degrees of physics-related self-concept for both (a) physics-related self-concept types and (b) female and male students.

Discussion

Comparing the different calculated mixed Rasch models, I found that the students were best described when categorised into three different groups. The two assessed aspects 'self-perception of ability' and 'beliefs about the perception as a physics person by others' differed in their (positive or negative) effect on the student's physics-related self-concept for the different groups. For group 2_{SC} students, both aspects were contributing in the same (positive or negative) way to the students' physics-related self-concept; that is, the students expressed similar degrees of agreement to all items. Yet, for group 1_{SC} and 3_{SC} students, the self-perception of ability contributed more positively to the students' physics-related self-concept compared to their beliefs about the perception as a physics person by others; that is, both groups had an equivalent item hierarchy and expressed higher degrees of agreement to ability self-concept items compared to perceived recognition items. As for the mechanics and particle physics interest groups in the previous section, I argue that the students of groups 1_{SC} and 3_{SC} were categorised into different groups despite their equivalent item hierarchies because they have different response styles, for which their self-concept profiles (**Figure 19**) and their category probabilities (**Figure 20**) provide evidence. Group 1_{SC} students used all the rating scale categories depending on the item, especially the non-extreme categories, and thus, the thresholds are rather close together. Group 2_{SC} students tend to use rather the non-extreme categories and especially the category '*agree*' equally across all items. Hence, the thresholds are rather close together. Group 3_{SC} students used the non-extreme rating scale categories across all items, especially the category '*disagree*'. Hence, all thresholds are very widely spaced, which resembles a NERS. The different response styles can also be seen in the boxplots with jitter for the physics-related self-concept groups (**Figure 21**). The NERS of group 3_{SC} can be seen because the students' person measures are evenly and widely spread around the mean. In comparison, the person measures of groups 1_{SC} and 2_{SC} are also evenly spread but closer together.

In sum, I argue that since the students differ in their response style, they have been categorised into the different groups 1_{SC} and 3_{SC}, although their item hierarchies are equivalent; that is, the students express more (or less) agreement to the same items. Hence, I argue that 60% of the students can be described in terms of one type of physics-related self-concept SC_1, the type of students whose self-perception of ability is more positive than their beliefs about the perception as a physics person by others. In contrast, for students assigned to group 2_{SC} both assessed aspects contribute equally to their physics-related self-concept. These students can be described in terms of a second type of physics-related self-concept SC_2, the type of students who have matching self-perception of ability and beliefs about the perception as a physics person by others. Type SC_2 students have higher mean physics-related self-concept and interest in mechanics and particle physics, respectively, compared to type SC_1 students.

For assessing self-concept, a 4-category rating scale was used, whereas a 5-category rating scale was used to assess interest. However, as for the interest groups it was possible to distinguish different response styles using a mixed Rasch model. In contrast to the interest groups in mechanics and particle physics, respectively, the ERS was not found for the self-concept groups. This may either be caused by the different foci of the items (interest vs. self-concept) or by the different rating scale formats (5 vs 4 categories).

The definition of self-concept used in this study is broad and the assessed aspects are two-fold, namely the self-perception of ability and the beliefs about the perception as a physics person by others. The physics identity studies suggested that the perceived recognition (i.e. the beliefs about the perception as a physics person by others) has a strong effect on the competency beliefs (i.e. the self-perception of ability; Kalender et al., 2019b). My study adds to this finding by showing that for 60% of the students, the type SC_1, both aspects contribute differently to their physics-related self-concept; that is, the students have a more positive self-perception of ability than beliefs about the perception as a physics person by others. The remaining 40% of the students, the type SC_2, have equivalent self-perception of ability and beliefs about the perception as a physics person by others. Despite the different contributions of the two aspects to the self-concept of the students assigned to the two self-concept types, the Rasch model could be applied for all students. Hence, my instrument to measure self-concept was found to be unidimensional. This supports a broad definition of self-concept and further contributes to the empirical background for assessing self-concept in future studies.

6.3.6. Physics-related self-concept – a better independent variable than sex?

In this section the results related to my research question about the comparison of physics-related self-concept with sex as an independent variable when analysing students' interest are discussed. RQ9 asked *'To what extent is physics-related self-concept a better independent variable than sex for distinguishing between different types of interest in mechanics and in particle physics?'*. The hypothesis for this research question is that the interest types are described better, when using physics-related self-concept instead of gender sex as an independent variable. This hypothesis is based on previous studies which suggested that interest and physics-related self-concept are associated with one another (Godwin et al., 2016; Häußler, Lehrke, et al., 1998; Kalender et al., 2019a). To answer this research question, I investigated the relationship between the students' interest in physics, physics-related self-concept, and sex using independent-samples t-tests, as well as linear and logistic regression analyses.

Results

At first, I prepared the dataset for such multivariate analyses. I had already excluded 28 students because of person measure misfit during the mechanics interest analysis (**Section 6.3.2.i.**), 39 students because of person measure misfit during the particle physics interest analysis (**Section 6.3.2.ii.**), and 73 students because of person measure misfit during the physics-related self-concept analysis (**Section 6.3.5.**). Since there was some overlap (i.e. some students were excluded because of more than one fit evaluation), I had excluded in total 115 students (i.e. the sample size for further analysis was N = 1072 students). Then, I evaluated multivariate outliers of this sample using the Mahalanobis distances. A value less than 0.001 for the probability that a value from the chi-square distribution with the same degrees of freedom will be less than the Mahalanobis distance indicates a multivariate outlier. Following this cut-off value, I excluded three more students from further analysis (i.e. the sample size for further analysis was N = 1069 students). Finally, I excluded 68 students from the data because they did not indicate their sex. The sample size for further analyses is N_{red} = 1001 students (540 female, 54%; 461 male, 46%).

Moreover, I added the students' most probable interest type in mechanics and particle physics, respectively, as well as the most probable self-concept type as new variables to the dataset. These new variables are a recoding of the most probable interest group. For mechanics, students coded as most probable interest group 1_M, 2_M, or 3_M were combined to one interest type and recoded as most probable interest type M_1, and students coded as 4_M were recoded as most probable interest type M_2. For particle physics, students coded as most probable interest group 1_{PP} or 2_{PP} were combined to one interest type and recoded as most probable interest type PP_1, and students coded as 3_{PP} were recoded as most probable interest type PP_2. For physics-related self-concept, students coded as most probable self-concept group 1_{SC} or 3_{SC} were combined to one self-concept type and recoded as most probable self-concept type SC_1, and students coded as 2_{SC} were recoded as most probable self-concept type SC_2.

Following this data preparation, I could calculate the different logistic regression models. The dependent variable is the student's most probable type assignment. The independent variables are the

Table 32. Regression models analyses of the interest types in mechanics (2 types, reference type is 1_M). ***p < 0.001, **p < 0.01, *p < 0.05 | *B* is the correlation coefficient, *SE* is the standard error, *OR* is the odds ratio, *CI* is confidence interval, and *df* is the degrees of freedom.

Model	A (df=2)		B (df=1)		C (df=1)	
Measure	B (SE)	OR [95% CI]	B (SE)	OR [95% CI]	B (SE)	OR [95% CI]
Constant	2.778*** (0.191)	16.079	2.146*** (0.132)	8.547	2.610*** (0.170)	13.595
Person measure self-concept	-	-	-	-	-	-
Self-concept type (1)	-0.418* (0.195)	0.658 [0.449, 0.965]	-0.480* (0.191)	0.619 [0.425, 0.901]	-	-
Sex (female)	-1.157*** (0.208)	0.315 [0.209, 0.473]	-	-	-1.179*** (0.207)	0.307 [0.205, 0.461]
Correctly classified cases	87.4%		87.4%		87.4%	
Nagelkerke's R^2	0.074		0.012		0.066	
Likelihood ratio test	40.126***		6.235* (0.013)		35.560***	

Model	D (df=3)		E (df=2)		F (df=2)	
Measure	B (SE)	OR [95% CI]	B (SE)	OR [95% CI]	B (SE)	OR [95% CI]
Constant	2.597*** (0.204)	13.428	1.953*** (0.141)	7.050	**2.504*** (0.173)**	**12.228**
Person measure self-concept	0.120* (0.055)	1.128 [1.013, 1.256]	0.170*** (0.053)	1.185 [1.068, 1.315]	**0.142** (0.049)**	**1.152 [1.047, 1.268]**
Self-concept type (1)	-0.194 (0.219)	0.824 [0.537, 1.265]	-0.164 (0.213)	0.849 [0.559, 1.290]	-	-
Sex (female)	-1.080*** (0.211)	0.340 [0.225, 0.513]	-	-	**-1.075*** (0.211)**	**0.341 [0.226, 0.516]**
Correctly classified cases	87.4%		87.4%		**87.4%**	
Nagelkerke's R^2	0.083		0.031		**0.081**	
Likelihood ratio test	44.947***		16.528***		**44.162*** **	

cont.

Model	G (df=1)	
Measure	B (SE)	OR [95% CI]
Constant	1.876*** (0.096)	6.528
Self-concept person measure	0.187*** (0.048)	1.206 [1.098, 1.324]
Self-concept type (1)	-	-
Sex (female)	-	-
Correctly classified cases	87.4%	
Nagelkerke's R^2	0.030	
Likelihood ratio test	15.940***	

student's physics-related self-concept (continuous Rasch measures and categorical most probable group assignment obtained from the mixed Rasch analysis) and sex (categorical). In total, I calculated 7 regression models regarding each content area. Each regression model comprised a different combination of independent variables. To compare the models, I used the Likelihood ratio test, the Nagelkerke's R^2, and the percentage of correctly classified cases.

Table 32 lists the results of these analyses for interest in mechanics. The simplest best fitting model (F) in comparison to the baseline model included the students' Rasch person measure indicating their physics-related self-concept (continuous) and their sex (categorical) as significant independent variables for interest type assignment in mechanics (Chi-squared=44, p=<0.001, df=2). This model fit the data better than model D, although model D had a slightly higher Chi-squared value, because model D included an additional independent variable (categorical self-concept) which did not significantly contribute to the model. Hence, model F was the simplest best fitting model including only variables that were significantly contributing to the model. Model F could correctly assign 87% of the cases to their mechanics interest type. The odds ratio that a female student is in type 2_M was OR=0.34. Since sex is a categorical variable, I could calculate the relative risk corresponding to this odds ratio[33]. This relative risk was RR=0.39 which means that female students were 61% less likely to be in type 2_M. The odds ratio that a student is in type 2_M for a student with higher self-concept values was OR=1.2; that is, for a one-unit increase in self-concept, the odds of being in type 2_M increase by the factor 1.2. This means that students with higher self-concept values were more likely to be in type 2_M. Since this is the continuous self-concept variable, I could not calculate the relative risk corresponding to this odds ratio.

Table 33 lists the results of these analyses for interest in particle physics. The best fitting model (F) in comparison to the baseline model (Chi-squared=100, p<0.001, df=2) included the students'

---

[33] RR=0.341-(0.341-1)*(37/540)=0.386

Table 33. Regression models analysis of the Rasch person measures for interest in particle physics (2 types, reference is type 1_{PP}). ***p < 0.001, **p < 0.01, *p < 0.05 | *B* is the correlation coefficient, *SE* is the standard error, *OR* is the odds ratio, *CI* is confidence interval, and *df* is degrees of freedom.

Model	A (df=2)		B (df=1)		C (df=1)	
Measure	B (SE)	OR [95% CI]	B (SE)	OR [95% CI]	B (SE)	OR [95% CI]
Constant	2.358*** (0.159)	10.570	1.692*** (0.112)	5.432	2.079*** (0.137)	8.000
Self-concept person measure	-	-	-	-	-	-
Self-concept type (1)	-0.663*** (0.164)	0.515 [0.373, 0.711]	-0.707*** (0.159)	0.493 [0.361, 0.674]	-	-
Sex (female)	-1.236*** (0.172)	0.291 [0.208, 0.407]	-	-	-1.260*** (0.170)	0.284 [0.203, 0.396]
Correctly classified cases	79.9%		79.9%		79.9%	
Nagelkerke's R^2	0.115		0.031		0.091	
Likelihood ratio test	75.872***		19.662***		59.564***	
Model	D (df=3)		E (df=2)		F (df=2)	
Measure	B (SE)	OR [95% CI]	B (SE)	OR [95% CI]	B (SE)	OR [95% CI]
Constant	2.039*** (0.167)	7.679	1.400*** (0.118)	4.056	**1.924*** (0.139)**	**6.848**
Self-concept person measure	0.241*** (0.048)	1.272 [1.158, 1.398]	0.285*** (0.046)	1.330 [1.214, 1.456]	**0.267*** (0.043)**	**1.306 [1.199, 1.422]**
Self-concept type (1)	-0.237 (0.184)	0.789 [0.551, 1.131]	-0.204 (0.178)	0.815 [0.575, 1.156]	**-**	**-**
Sex (female)	-1.106*** (0.175)	0.331 [0.235, 0.466]	-	-	**-1.100*** (0.174)**	**0.333 [0.236, 0.468]**
Correctly classified cases	80.7%		80.0%		**81.2%**	
Nagelkerke's R^2	0.153		0.091		**0.151**	
Likelihood ratio test	102.094***		59.479***		**100.435*** **	

cont.

Model	G (df=3)	
Measure	B (SE)	OR [95% CI]
Constant	1.304*** (0.081)	3.683
Self-concept person measure	0.306*** (0.042)	1.359 [1.251, 1.476]
Self-concept type (1)	-	-
Sex (female)	-	-
Correctly classified cases	80.1%	
Nagelkerke's R^2	0.089	
Likelihood ratio test	58.167***	

physics-related self-concept (continuous Rasch person measure) and their sex (categorical) as significant independent variables for interest group assignment in particle physics. This model could correctly assign 81% of the cases to their particle physics interest type. The odds ratio that a female student is in type 2_{PP} was OR=0.33. Again, since sex is a categorical variable, I could calculate the relative risk corresponding to this odds ratio[34]; that is, RR=0.26. This means that female students are 74% less likely to be in type 2_{PP}. The odds ratio that a student is in type 2_{PP} for a student with lower self-concept vales was OR=1.3. That is, for a one-unit increase in self-concept, the odds of being in type 2_M increase by factor 1.3. This means that students with higher self-concept values were more likely to be in type 2_{PP}. Since this is the continuous self-concept variable, I could not calculate the relative risk corresponding to this odds ratio.

Next, I performed an independent's sample t-test to check whether there is a significant difference in mean self-concept between female and male students. First, I conducted this analysis for the full sample. The Levene's test did not show a significant difference; that is, equal variances for female and male students are assumed. I found that female students had a significantly lower physics related self-concept (-1.00 ± 2.04 logits) compared to male students (-0.19 ± 2.04 logits), t(999)=6.27, p=<0.001. This is a small effect (d_C=0.4, 95% *CI* [0.27, 0.52]).

Second, I conducted this analysis separately for the two self-concept types, the type SC_1 consisting of the groups 1_{SC} and 3_{SC} and the type SC_2 consisting of the group 2_{SC}. For both types, the Levene's test did not show a significant difference; that is, equal variances for female and male students are assumed. For the first type SC_1, I found that female students had significantly lower physics related self-concept (-1.70 ± 1.98 logits) compared to male students (-1.02 ± 1.96 logits), t(609)=4.18, p=<0.001. This is a small effect (d_C=0.3, 95% *CI* [0.18, 0.50]). For the second type SC_2, I also found that female students had statistically significant, lower physics related self-concept (0.24 ± 1.50 logits) compared

---

[34] RR=0.333 -(0.333 -1)*(60/540)=0.259

to male students (0.94 ± 1.54 logits), t(388)=4.53, p=<0.001. This is a medium-sized effect (d_C=0.5, 95% *CI* [0.26, 0.66]).

Third, I conducted this analysis separately for the two mechanics interest types, the type M_1 consisting of the groups 1_M, 2_M, and 3_M and the type M_2 consisting of the group 4_M. For both types, the Levene's test did not show a significant difference; that is, equal variances for female and male students are assumed. For the type M_1, I found that female students had significantly lower physics related self-concept (-1.02 ± 2.03 logits) compared to male students (-0.32 ± 1.95 logits), t(873)=5.10, p=<0.001. This difference translates into a small effect (d_C=0.3, 95% *CI* [0.21, 0.48]). For the type M_2, I also found that female students have a significantly lower physics-related self-concept (-0.71 ± 2.17 logits) compared to the male students' physics-related self-concept (0.38 ± 2.31 logits), t(124)=2.45, p=0.016. This is a medium-sized effect (d_C=0.5, 95% *CI* [0.09, 0.87]).

Finally, I conducted this analysis separately for the two particle physics interest types, type PP_1 consisting of the groups 1_{PP} and 2_{PP} and the type PP_2 consisting of the group 3_{PP}. For the type PP_1, the Levene's test did not show a significant difference; that is, equal variances for female and male students are assumed. I found that female students had significantly lower physics related self-concept (-1.05 ± 2.00 logits) compared to male students (-0.61 ± 1.92 logits), t(798)=3.05, p=0.002. This is a small effect (d_C=0.2, 95% *CI* [0.08, 0.36]). For the type PP_2, the Levene's test showed a significant difference (p=0.045); that is, equal variances for female and male students are not assumed and the degrees of freedom are adjusted. I found a statistically significant difference between the female students' physics-related self-concept (-0.63 ± 2.35 logits) compared to the male students' physics-related self-concept (0.77 ± 1.97 logits), t(96.206)=4.07, p=<0.001. This is a medium-sized effect (d_C=0.7, 95% *CI* [0.36, 0.98]).

Interestingly, for all the above analyses, the difference in physics-related self-concept between female and male students within the types with higher mean interest translate into a bigger effect than in the other type.

In addition, I performed a linear regression/correlation analysis to examine the association between the students' self-concept and their interest in mechanics or particle physics, respectively, for the full sample. The correlations between self-concept and interest in mechanics and particle physics, respectively, were similar ($r_{M,SC}$=0.38, 95% CI [0.32,0.43], p=<0.001; $r_{PP,SC}$=0.47, 95% CI [0.42,0.52], p=<0.001). This similarity can also be seen in the scatterplot showing the students' Rasch person measures indicating their degrees of interest in mechanics versus their degrees of physics-related self-concept (**Figure 23**) and in the scatterplot showing the students' Rasch person measures indicating their degrees of interest in particle physics versus their degrees of physics-related self-concept (**Figure 24**).

Then, I performed a linear regression/correlation analysis to check whether the association between the students' self-concept and their interest in mechanics or particle physics, respectively, differs between female and male students within the different interest types. First, I found that for the mechanics type M_1, comprising the groups 1_M, 2_M, and 3_M, the correlation is similar for both female and male students ($r_{M,SC,1f}$=0.39, 95% CI [0.31,0.46], p=<0.001; $r_{M,SC,1m}$=0.41, 95% CI [0.32,0.49], p=<0.001).

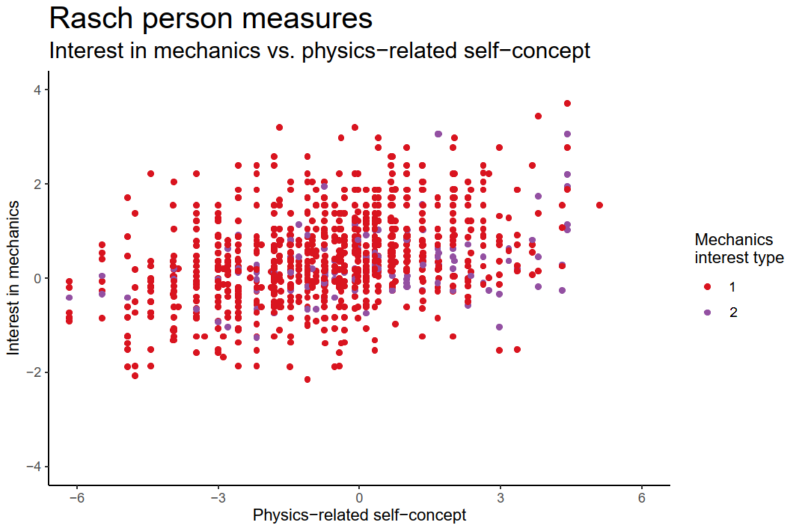


Figure 23. The students' Rasch person measures in logit units indicating their degrees of interest in mechanics versus their degrees of physics-related self-concept. The colour coding indicates to which mechanics interest type a person was assigned.

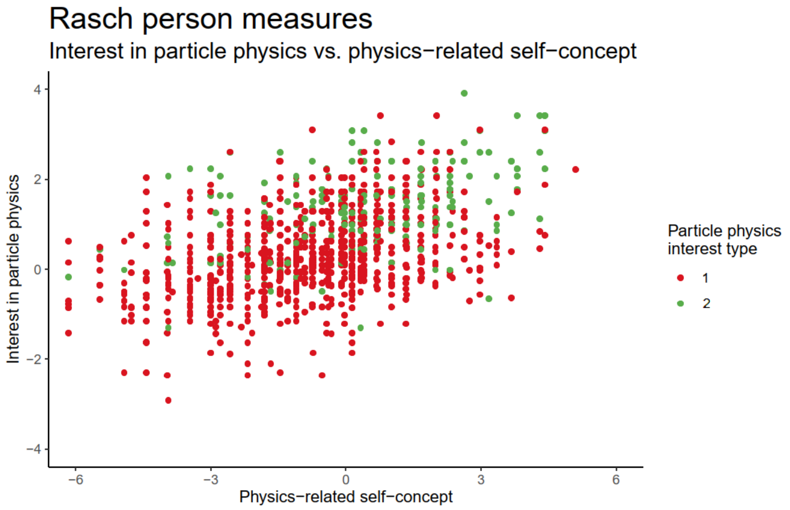


Figure 24. The students' Rasch person measures in logit units indicating their degrees of interest in particle physics versus their degrees of physics-related self-concept. The colour coding indicates to which particle physics interest type a person was assigned.

The same is true for mechanics type M_2, comprising the group 4_M ($r_{M,SC,2f}$ =0.35, 95% CI [0.02,0.60], p=0.036; $r_{M,SC,2m}$=0.34, 95% CI [0.14,0.51], p=0.001). Second, I found that for the particle physics type PP_1, that is the groups 1_{PP} and 2_{PP}, the correlation is similar for both female and male students ($r_{PP,SC,1f}$=0.42, 95% CI [0.35,0.50], p=<0.001; $r_{PP,SC,1m}$=0.38, 95% CI [0.28,0.47], p=<0.001). The same is true for particle physics type PP_2, that is the group 3_{PP} ($r_{PP,SC,2f}$=0.44, 95% CI [0.21,0.62], p=<0.001; $r_{PP,SC,2m}$=0.46, 95% CI [0.32,0.58], p=<0.001).

In addition, I performed an independent samples t-test to check whether there is a significant difference in mean interest in mechanics and particle physics, respectively, between female and male students. First, I conducted this analysis for the full sample. For mechanics, the Levene's test did not show a significant difference; that is, equal variances for female and male students are assumed. I did not find a statistically significant difference between the female students' mean interest in mechanics (0.42 ± 0.93 logits) and the male students' mean interest in mechanics (0.46 ± 0.92 logits), p=0.568. For particle physics, the Levene's test showed a significant difference (p=0.040); that is, equal variances for female and male students are not assumed and the degrees of freedom are adjusted. I found a statistically significant difference between the female students' mean interest in particle physics (0.38 ± 0.98 logits) and the male students' mean interest in particle physics (0.67 ± 1.09 logits), t(934.550)=4.36, p=<0.001. This is a small effect (d_C=0.3, 95% CI [0.27, 0.52]).

Second, I conducted this analysis separately for the two mechanics interest types M_1 and M_2. For both types, the Levene's test did not show a significant difference; that is, equal variances for female and male students are assumed. There was no statistically significant difference between the female students' mean interest in mechanics (type M_1: 0.44 ± 0.95 logits, type M_2: 0.23 ± 0.65 logits) compared to the male students' mean interest in mechanics (type M_1: 0.48 ± 0.95 logits, type M_2: 0.34 ± 0.79 logits; p=0.465 and p=0.457, respectively).

Finally, I conducted this analysis separately for the two particle physics interest types. For both types, the Levene's test did not show a significant difference; that is, equal variances for female and male students are assumed. For the first type PP_1, there was no statistically significant difference between the female students' mean interest in particle physics (0.29 ± 0.95 logits) compared to the male students' mean interest in particle physics (0.31 ± 0.94 logits), p=0.798. For the second type PP_2, I found a statistically significant difference between the female students' mean interest in particle physics (1.10 ± 0.92 logits) compared to the male students' mean interest in particle physics (1.49 ± 0.93 logits), t(199)=2.67, p=0.008. This is a small effect (d_C=0.4, 95% CI [0.11, 0.72]).

Discussion

First, a direct logistic regression analysis was performed. The dependent variable was the students' most probable type assignment concerning their interest in mechanics and in particle physics, that is, the joint description of the groups obtained from the mixed Rasch model as interest types (as presented in **Section 6.3.2.**). The independent variables were the student's physics-related self-concept (continuous: Rasch measures, categorical: most probable type assignment) and sex (categorical).

The best fitting logistic regression models in comparison to the respective baseline model for mechanics and particle physics, respectively, to distinguish between the two mechanics interest types M_1 and M_2 and between the two particle physics interest types PP_1 and PP_2, respectively, comprised the students' physics-related self-concept as a continuous Rasch person measure and their sex (categorical) as significant independent variables for type assignment (**Table 32** and **Table 33**).

For mechanics, this model correctly assigned 87% of the cases. However, all models assigned the same percentage of students correctly because they all assigned all cases to type M_1 ('Physics? Only in certain contexts!'). This indicates that, in general, a logistic regression model may not be the best choice for describing the data. The *Nagelkerke* R^2 of 0.08 provides further evidence that the model only provided poor improvement in fit over the baseline model. The rule-of-thumb guideline says that models with R^2 values from 0 to 0.1 only provide a poor improvement of the fit in comparison to the baseline model (Muijs, 2004). However, the models showed that female students are 61% less likely to be in the second type of interest (i.e. the type M_2, 'students interested in physics relating to the motion of cars') compared to male students. Moreover, the odds to be in the type M_2 ('students interested in physics relating to the motion of cars') of a person with a one-unit increase in physics-related self-concept increase by a factor 1.2. Since (to my knowledge) it is not possible to calculate the probability of a one-unit increase on a continuous scale, I could not calculate the relative risk from the odds ratio. Since the odds ratio does not provide sound information about the strength of this positive relationship, I can only conclude that as the students' physics-related self-concept increased, their odds of being assigned to type M_2 in mechanics increased.

For particle physics, this model correctly assigns 81% of the cases. This is the best percentage of correctly assigned students compared to the other models. When all cases are assigned to type PP_1 ('Physics? Only in certain contexts!'), a percentage of 80% is correctly assigned; that is, the improvement of correctly assigned students is 1%. This indicates that, in general, a logistic regression model may not be the best choice for describing the data. The *Nagelkerke* R^2 of 0.15 provides further evidence that the model only provided modest improvement in fit over the baseline model. The rule-of-thumb guideline says that models with R^2 values from 0.1 to 0.3 only provide a modest improvement of the fit in comparison to the baseline model (Muijs, 2004). However, the models showed that female students are 74% less likely to be in the second type of interest (i.e. the type PP_2, 'particle physics lovers') compared to male students. Moreover, the odds to be in the type PP_2 ('particle physics lovers') of a person with a one-unit increase in physics-related self-concept increase by factor 1.3. Again, I could not calculate the relative risk from the odds ratio, and I can only say that as the students' physics-related self-concept increased, their odds of being assigned to type PP_2 in particle physics increased.

In sum, for both content areas, a logistic regression model comprising the physics-related self-concept as a continuous variable in addition to sex described the data better than a model comprising sex alone, each compared to the corresponding baseline model. Considering the large differences in amount of female and male students in the mechanics type M_2 ('students interested in physics relating to the motion of cars') and the particle physics type PP_2 ('particle physics lovers'), it is remarkable that the physics-related self-concept significantly contributed to the respective models, that is, the model was significantly better than if the variable 'physics-related self-concept' was 0. Hence, I argue that logistic regression models provided evidence that the physics-related self-concept is an important independent variable for describing students' interest. However, the interpretability of odds ratios in logistic regression models is limited to the direction of the effect as the effect size cannot be interpreted in a sound manner. Moreover, to my knowledge, it is not possible to calculate the probabilities of a one-unit increase on a continuous variable, and hence to calculate the relative risk. Due to these limitations, it was not possible to compare the effect sizes between the categorical variable sex and the continuous variable physics-related self-concept using logistic regression models in my study. It would have been possible to draw conclusion about which variable is better than the other using a logistic regression models, if, for example, the model that only comprises the physics-related self-concept was significantly better than the one only comprising sex, each compared to the baseline model.

I argue that sex was a significant variable because there is an uneven distribution of sexes for the interest type M_2 ('students interested in physics relating to the motion of cars') in mechanics and PP_2 ('particle physics lovers') in particle physics. There were more male than female students belonging to this type and the students' mean physics-related self-concept was higher. Since students with higher physics-related self-concept were rather male than female, it is not surprising that both independent variables ‘physics-related self-concept’ and ‘sex’ were significant when the number of male and female students is very unequal.

The students’ characteristics for the different physics-related self-concept types were described in **Table 31**. Self-concept type SC_1 students had on average higher physics-related self-concept and both aspects ‘self-perception of ability’ and ‘beliefs about the perception as a physics person by others’ contributed equally positive (or negative) to their physics-related self-concept. Self-concept type SC_2 students had on average lower physics-related self-concept and their self-perception of ability contributed more positively to their physics-related self-concept than their beliefs about the perception as a physics person by others. Moreover, I investigated the assignment of the different interest types to the different self-concept types (**Table 19** and **Table 23**). I found that the students of interest type M_1 in mechanics and PP_1 in particle physics (‘Physics? Only in certain contexts!’) rather belonged to the type SC_1 in self-concept. The students of interest type M_2 (‘students interested in physics relating to the motion of cars’) in mechanics and PP_2 (‘particle physics lovers’) in particle physics equally belonged to the self-concept types SC_1 and SC_2. Hence, I argue that the categorical self-concept type was not statistically significant in the logistic regression models because both self-concept types were rather equally assigned to both interest types concerning both content areas.

In sum, I argue that the logistic regression model did not provide evidence that the students’ physics-related self-concept is a better independent variable for distinguishing between the two interest types concerning both content areas, that is, between the first type M_1 and PP_1 (‘Physics? Only in certain contexts!’) and the second type M_2 in mechanics (‘students interested in physics relating to the motion of cars’) and PP_2 in particle physics (‘particle physics lovers’), respectively. Both variables, sex and continuous Rasch person measures indicating the students’ degrees of self-concept, were significantly improving the model (compared to a model where these variables are 0). This contrasts with my hypothesis that physics-related self-concept is a better independent variable than sex. To learn more about the relationship between interest, self-concept, and sex, I additionally calculated different t-tests, correlation coefficients, and linear regression models.

First, I analysed whether female and male students significantly differ in their mean physics-related self-concept using t-tests. In general, female students had a lower mean physics-related self-concept than male students. Overall, this difference in mean self-concept translated into a small effect. However, separate analysis for the self-concept types showed that the effect was small for type SC_1 (both aspects ‘self-perception of ability’ and ‘beliefs about the perception as a physics person by others’ are equivalent) but medium-sized for type SC_2 (‘self-perception of ability’ is more positive than ‘beliefs about the perception as a physics person by others’). Similarly, in the separate analysis for the mechanics and particle physics interest types, the effect was small for the types M_1 and PP_1 (‘Physics? Only in certain contexts!’) but medium-sized for the type M_2 in mechanics (‘students interested in physics relating to the motion of cars’) and PP_2 in particle physics (‘particle physics lovers’), respectively. It is remarkable that the differences between the sexes were more pronounced in the types with higher mean physics-related self-concept (SC_2, M_2, PP_2). This provides evidence that the differences are especially pronounced when focusing on the students with high levels of physics-related self-concept, whereas for most students these differences are comparatively less pronounced. This finding is somewhat surprising and suggests that in the types of students with higher mean self-concept, the sex differences are even more pronounced. Similarly, Buccheri et al. (2011) analysed the PISA 2006 data of four different countries and found that ‘female high-performers have a significantly lower [ability] self-concept in sciences than their male colleagues’.

Second, I investigated whether the association between self-concept and interest in mechanics and particle physics, respectively, is different between female and male students within the different interest types using a linear regression/correlation analysis. However, for both sexes in all types the correlation coefficients between self-concept and interest were similar and in the range of the commonly reported coefficient of $r_{I,SC} = 0.4$ (e.g. Woithe, 2020). This finding shows that although there were significant differences in the mean self-concept values of female and male students, the association between self-concept and interest was similarly strong for both sexes; that is, for both sexes, the self-concept was equally associated with interest. Accordingly, it is not surprising that when female students have a lower mean self-concept, they also have a lower mean interest compared to male students. Hence, I conclude that the key to fostering female students’ interest is helping them develop a more positive physics-related self-concept; that is, programmes aiming at fostering interest should foster self-concept at the same time.

Third, I investigated whether there is a significant difference in mean interest in mechanics and particle physics, respectively, between female and male students using t-tests. For mechanics, I did not find a statistically significant difference, neither for the full sample nor for the different mechanics interest types. This is the same for particle physics, except for the type PP_2 (‘particle physics lovers’). For this type, I found a statistically significant difference between the female and male students’ mean interest in particle physics translating into a small effect. This is also a remarkable finding because it shows that for most students there was no significant difference in mean interest associated with their sex. Yet, concerning the content area ‘particle physis’ and the type of students referred to as ‘particle physics lovers’ (PP_2), there is a statistically significant lower interest of female students compared to male students. This provides evidence that there are no differences in mean interest between the sexes of all students concerning mechanics and of the vast majority of students (i.e. type PP_1, ‘Physics? Only in certain contexts!’, 79% of students) concerning particle physics. Only when focusing on the students with higher mean interest in particle physics (i.e. type PP_2, ‘particle physics lovers’, 21% of students), male students had a significantly higher interest than female students. This finding contrasts with our hypothesis that particle physics is equally and highly interesting for all students. Yet, it also provides evidence that (apart from this exception) there were no significant differences in mean interest between the sexes overall.

In sum, using t-tests I found that the difference in mean interest between the sexes is only significant for the particle physics lover type of students concerning the content area ‘particle physics’. The association between self-concept and interest was found to be similar for all students of all types in both content areas using linear regression/correlation analyses. Yet, I found that the mean self-concept significantly differs between the sexes, with girls having lower mean self-concept. This effect was relatively larger for the types (concerning self-concept, mechanics interest, and particle physics interest, respectively), which have the higher mean physics-related self-concept (i.e. types SC_2, M_2, PP_2). These results combined provide evidence that the students differ in their degrees of interest because of their differences in physics-related self-concept. Yet, the physics-related self-concept significantly differs for female and male students. Hence, I argue that the direct association between sex and interest is less descriptive for the students’ interest than the association between sex and physics-related self-concept because self-concept itself is similarly associated with interest for all students. The results of these analyses provide evidence that the association of sex and interest is mediated via self-concept. Similarly, Kalender et al. (2019b) have shown using structural equation modelling that gender has a significant effect on perceived recognition, which in turn has an effect on interest. Moreover, they showed that (1) self-efficacy and belonging and (2) physics identity and interest are strongly aligned and do not factor out separately for either gender (Kalender et al., 2019a). There was one exception, namely, for female students the perceived recognition (i.e. external identity) item 'My TA or instructor see me as a physics person' loads strongly with the factor 'self-efficacy or belonging' instead of 'interest or identity' (Kalender et al., 2019a). That is, female students' perceived recognition by their TA or instructor is closely linked to their self-efficacy and sense of belonging in a physics classroom (Kalender et al., 2019a). They also showed that perceived recognition has an effect on competency beliefs (Kalender et al., 2019b). Moreover, Godwin et al. (2016) showed that competency beliefs have a very large direct effect on interest and on perceived recognition. However, in my study perceived recognition (i.e. beliefs about the perception as a physics person by others) and competency beliefs (i.e. self-perception of ability) were combined as self-concept. Nevertheless, it is remarkable that also in the study by Kalender et al. (2019b), the effect of gender was mediated by perceived recognition. Hence, I argue that a focus on the students' sex for describing their interest is not sufficient, instead both variables 'sex (or gender)' and 'self-concept' should jointly be discussed when describing the students' interest in physics.

6.3.7. Previous experience with the content areas in school – does it matter?

In this section the results related to my research question about the students' previous experience with the content areas in school are discussed. RQ10 asked *'To what extent does the students' previous experience with the content areas in school affect their expressed interest in these content areas?'*. The hypothesis for this research question was that the students' expressed interest in the content areas changes depending on their previous experiences with it in school. I investigated to what extent the students' previous experience with the content areas affects their expressed interest in these content areas using an analysis of the mean individual experience differences and one-sample t-tests, as well as linear regression analyses and independent-samples t-tests.

Results

First, I calculated a Rasch model for students' (1) previous experience with mechanics in school and (2) previous experience with particle physics in school. **Table 34** lists the items sorted by their item measure, that is, how much experience the students reported to each content, for both item sets. One can see that regarding mechanics the students reported relatively more previous experience in the contents 'force', 'velocity', and 'energy of motion', and less in the contents 'pulley', 'pump', and 'lifting platform'. Regarding particle physics they reported relatively more previous experience in the contents 'electron', 'nucleus space', and 'speed of light', and less in the contents 'particle accelerator', 'the big bang' and 'particle detector'.

Table 34. Items, item measures, and corresponding standard errors (SE) for students' previous experience in school regarding mechanics and particle physics, respectively.

Mechanics	Measure	SE	Particle Physics	Measure	SE
lifting platform	1.03	0.04	particle detector	0.9	0.04
pump	0.95	0.04	the big bang	0.57	0.04
pulley	0.67	0.03	particle accelerator	0.48	0.04
braking distance	0.1	0.03	elementary particles	0.1	0.03
energy of motion	-0.74	0.03	speed of light	-0.42	0.03
velocity	-0.95	0.04	nucleus space	-0.59	0.03
force	-1.07	0.04	electron	-1.04	0.03

Note: The Rasch item measures and standard errors are listed in logit units. Lower and higher item measures represent items regarding which students report less and more experience, respectively. The mean item measure of each group is set to zero in mixed Rasch analysis. The items are ordered according to their measures. The colour scale represents the degree of reported experience; dark blue is the item regarding which the most experience was reported and dark red is the item regarding which the least experience was reported.

When conducting a Rasch analysis, an evaluation of the obtained item and person measures is crucial as it provides further evidence for the quality of the chosen model. First, I evaluated the Rasch item measures regarding previous experience with mechanics in school. I found that all items fit well (*ZQ* values ranging from -0.31 to 1.24) except for the item 'lifting platform' (*ZQ* = -2.13). I argue that this is because 'lifting platform' is the item with the highest item measure; that is, the students reported the least previous experience with the content 'lifting platform'. Hence, it may not seem to fit well to the one-dimensional trait 'previous experience with mechanics'. However, mechanics is a broad content area comprising many different contents (cf. Linacre (2020) about mathematics as a broad domain). Second, I evaluated the Rasch item measures in previous experience regarding particle physics in school. I found that some items fit well (*ZQ* values ranging from -1.82 to 1.87), whereas other items do not fit well, namely 'particle detector' (*ZQ* = -3.56), 'particle accelerator' (*ZQ* = -2.00), and 'electron' (*ZQ* = 2.44). I argue that this is because 'electron' is the item with the lowest item measure, that is, the students reported the highest previous experience with the content 'electron'. Similarly, the contents 'particle accelerator' and 'particle detector' both have relatively high item measure; that is, the students reported rather low previous experience with these contents. Hence, these three contents may not seem to fit well to the one-dimensional trait 'previous experience with particle physics'. However, particle physics is a broad content area comprising many different contents (cf. Linacre (2020) about mathematics as a broad domain). In sum, despite the reported item misfits, the data supported one trait, previous experience with the content area in school, for all groups, that is, the measurement instruments about students' (1) previous experience with the mechanics in school and (2) previous experience with particle physics in school are unidimensional. Then, I evaluated the Rasch person measures for the students' previous experience. I investigated the Rasch person fit and excluded those students from further analysis, whose newfit 1 and/or 2 value was $\ll -3$ or $\gg 3$ or for who no newfit values were calculated. Concerning previous experience with mechanics, I excluded 20 students from further analyses, and concerning previous experience with particle physics, I excluded 40 students. Since there was some overlap (i.e. some students were excluded because of both fit evaluations), I excluded in total 56 students (i.e. the sample size for the analysis presented in this chapter was *N* = 945 students).

First, I analysed in which content area mechanics or particle physics the students have overall more previous experience in school by investigating the individual differences in Rasch person measures. I analysed the means of the individual differences in Rasch person measures (experience in mechanics minus experience in particle physics). To check whether the means of the individual differences in Rasch person measures are significantly different from 0, I conducted a one-sample t-test.

For the full sample, the mean difference (M=0.3555, 95% CI = [0.2974, 0.4135]) was significantly higher than 0, t(944)=12.018, p=<0.001. This difference translates into is a small effect (d_C=0.39, 95% CI [0.325, 0.457]).

I also conducted this analysis separately for the mechanics interest types. For type M_1, the mean difference (M=0.3641, 95% CI = [0.3020, 0.4263]) was significantly different from 0, t(826)=11.498, p=<0.001. This difference translates into a small effect (d_C=0.40, 95% CI [0.329, 0.471]). For type M_2, the mean difference (M=0.2949, 95% CI = [0.1307, 0.4591]) was significantly higher than 0, t(117)=3.557, p=<0.001. This is a small effect (d_C=0.33, 95% CI [0.142, 0.512]).

Then, I conducted this analysis separately for the particle physics interest types. For type PP_1, the mean difference (M=-0.3433, 95% CI = [0.2796, 0.4070]) was significantly higher than 0, t(756)=-10.584, p=<0.001. This difference translates into a small effect (d_C=0.39, 95% CI [0.311, 0.458]). For type PP_2, the mean difference (M=0.4046, 95% CI = [0.2642, 0.5449]) was significantly higher than 0, t(187)=5.688, p=<0.001. This difference translates into a small effect (d_C=0.42, 95% CI [0.265, 0.563]).

In addition, I performed an independent-samples t-test to check whether there is a significant difference in previous experience in school concerning the content areas between the students assigned to the two interest types in mechanics and particle physics, respectively. For mechanics, the Levene's test did not show a significant difference (p=0.094); that is, equal variances for students in type M_1 and M_2 are assumed. I found that the students' mean experience in school with mechanics for type M_1 students (-0.31 ± 0.73 logits) was significantly lower than for type M_2 students (-0.14 ± 0.83 logits), t(943)=2.298, p=0.022. This difference translates into a small effect (d_C=-0.23, 95% CI [-0.419, -0.033]). For particle physics, the Levene's test did not show a significant difference (p=0.093); that is, equal variances for students in type PP_1 and PP_2 are assumed. I did not find a significant difference between the students' mean experience with particle physics of type PP_1 students (-0.65 ± 0.82 logits) compared to type PP_2 students (-0.58 ± 0.90 logits), t(943)=1.107, p=0.269.

I also performed a linear regression/correlation analysis to investigate the association between the students' interest in mechanics and particle physics, respectively, and their previous experience with these content areas in school.

First, I conducted this analysis for the full sample. I found that the correlation between the students' interest in and previous experience with mechanics is $r_{M,E}$=0.15 (95% CI [0.09,0.21], p=<0.001). The correlation between the students' interest in and previous experience with particle physics is $r_{PP,E}$=0.16 (95% CI [0.09,0.22], p=<0.001)

Second, I investigated whether this association is different for the students assigned to different interest types. I found that the correlation for the mechanics interest type M_1 is $r_{M,E,1}$=0.16 (95% CI [0.10,0.23], p=<0.001) and for mechanics interest type M_2 is $r_{M,E,2}$=0.13, 95% CI [-0.05,0.30], p=0.166). Second, I found that the correlation for the particle physics interest type PP_1 is $r_{PP,E,1}=0.19$ (95% CI [0.12,0.26], p=<0.001) and for the particle physics interest type PP_2 is $r_{PP,E,2}=0.03$ (95% CI [-0.11,0.18], p=0.655).

Discussion

First, I analysed regarding which content area (mechanics or particle physics) the students report overall more previous experience in school. I investigated the individual differences in Rasch person measures indicating the students' degrees of previous experience. Using one-sample t-tests, I analysed whether the means of the individual differences in Rasch person measures significantly differ from 0. I found that, on average, the students reported significantly more previous experience in mechanics than in particle physics. This is also the case when analysing the different interest groups individually.

Then, I calculated independent samples t-tests separately for both content areas to check whether the students' assigned to the different interest types differ in their previous experience in school. It is remarkable that the students assigned to type M_1 ('Physics? Only in certain contexts!') reported significantly lower previous experience in mechanics than the students assigned to type M_2 ('students interested in physics relating to the motion of cars'). I argue that the students aged 14 to 16 years may have already been forced to decide on or against physics during their educational path. Thus, part of the students indeed has more previous experience in school with mechanics than others because of the differing choices that they have made. The students may have decided on physics based on their interest; that is, the students who are interested in physics decide on physics. Similarly, Maltese et al. (2014) point towards a 'selection bias where those who are already interested seek out more coursework'. I argue that students who are interested in technical objects, such as cars, are more likely to decide on physics, as physics is often presented in school in relation to technical objects. Most of the students that participated in the study were from Austrian secondary school and in grade 9. Typically, in Austria, students who have decided on physics have physics lessons in grade 9, which is not the case for students who have decided against physics. In Austria, in the curriculum of grade 9 one term solely focuses on mechanics. Thus, it seems plausible that students who were interested in the context 'car' also reported more previous experience with mechanics in school.

For particle physics, the two interest types PP_1 ('Physics? Only in certain contexts!') and PP_2 ('particle physics lovers') did not differ significantly in their previous experience with particle physics in school. This finding is not surprising as particle physics is not yet established in the curricula for students aged 14 to 16 years. This means that because of this lack of presence in school curricula the differences in previous experiences with particle physics are little. Hence, the previous experience with particle physics in school is equally little for all students. This also is supported by the findings of the comparison of individual experience differences (one-sample t-test described above), which showed that all students have, on average, more previous experience regarding mechanics than particle physics in school.

Moreover, I performed a linear regression/correlation analysis to investigate the association between the students' interest in mechanics and particle physics, respectively, and their previous experience with these content areas in school. I found a small positive correlation between interest and previous experience for both content areas. When analysing the interest types separately, the correlation was significant for most students, that is, for the type of students that is only interested in physics set in certain contexts (type M_1 concerning mechanics and PP_1 concerning particle physics).

Interestingly, for the type M_2 ('students interested in physics relating to the motion of cars') and the type PP_2 ('particle physics lovers') there was no significant correlation between the students' interest and their previous experience with the content area. Here, three explanations seem plausible. First, it may be that the correlation is not significant because the sample sizes were smaller for the types M_2 and PP_2 in comparison to M_1 and PP_1 ('Physics? Only in certain contexts!'). Second, when students have a well-developed dispositional interest in a certain aspect, such as the content area 'particle physics' or the context 'car', the previous experience with the content area in school is less associated with the interest expressed in an item referring to these aspects. Third, for students with a well-developed dispositional interest, their interest may rather be associated with previous out-of-school experience.

However, the observed significant correlations for these associations were rather small ($r_{M,E}$=0.15 and $r_{PP,E}$=0.16), considering that the Pearson correlation coefficient ranges from -1 to 1. It is remarkable, though, that the correlation values were similar for both content areas, which may provide evidence that there is a non-negligible positive association between interest and previous experience in school. First, as already outlined above the differences in previous experience may have resulted from the students' already made decision on or against physics during their educational path. That is, more interested students may decide on physics, and hence have more previous experience in school because of their curriculum. This explanation applies grade 9 students concerning the content area 'mechanics', but it does not apply to the content area 'particle physics', with which all students should have equally little previous experience based on the curriculum. Nevertheless, considering that teachers focus on different physics content more (or less) exhaustively and may also tackle extra-curricular content, it is likely that students differ in their previous experiences with the content areas in school regardless of the choices in their educational path. Hence, this positive association may also provide evidence that the students' previous experience and interest foster each other; that is, more interested students making more previous experiences and more experienced students being more interested.

The theoretical and empirical literature mostly focuses on knowledge concerning a content area instead of previous experience with it. I argue that previous experience with a content area in school also provides a measure for knowledge as it can be regarded as an inevitable mediator for knowledge. Yet, I did not assess or make an association of interest with the students' knowledge concerning the content areas. Neither did I assess the students' out-of-school previous experiences with the content area. Both would also be relevant for making a strong link to the theoretical and empirical background on knowledge. However, a positive association between previous experience with or knowledge in a content area and interest in it is in line with the theoretical background, such as the 'Four-phase-model of interest development' (Hidi & Renninger, 2006), the 'Expectancy-value theory of achievement motivation' (Eccles, 2009; Wigfield & Eccles, 2000) and the 'Social cognitive career theory' (Sheu et al., 2010; Lent et al., 1994). An empirical longitudinal study by Höft et al. (2019) showed that interest in cognitively activating tasks or in tasks that involve the communication of knowledge (e.g. solving theoretical problems, debating with or explaining something to classmates) is increasingly associated with knowledge. Since my instrument also measures interest in specific content-context-task combinations, I argue that my finding of a positive correlation between interest in a content area and previous experience with it aligns with the findings by Höft et al. (2019).

6.4. Strengths, limitations, and directions for future research

This study investigated the students' types of interest in mechanics and in particle physics and their relation to the students' characteristics sex, physics-related self-concept, and previous experience with the content areas in school. Below I will describe the strengths and limitations of my study, as well as directions for future research sorted thematically.

Selection and definition of the assessed constructs

One strength of my study is the careful selection and thorough definition of the assessed constructs. First, based on an extensive literature review, I decided on, discussed, and provided a precise definition of interest considering the four aspects 'content', 'context', 'task', and 'learning environment'. I decided on measuring interest in a content area using a set of items that present different corresponding contents set in different contexts. This contrasts with other studies in which the items are based on rather broad definitions of interest (i.e. relating to physics as a domain or subject). Although the physics contents were additionally presented in a range of different tasks, the tasks were not varied in an equally systematic manner as the contexts. That is, there were seven different contexts represented by one or two items each, whereas the four different tasks were represented by one, two, or seven items each. Moreover, I did not specify nor vary the learning environment in the items. These are limitations of my study. Yet, it is a strength of my study that by using Rasch models I was perfectly aligning my analysis methods with my theoretical background about the construct 'interest'. In the 'Person-object theory of interest' both aspects, (a) the students and their interests and (b) the objects and their interestingness, play an equal role. Only by applying a Rasch model to the collected data, I could draw conclusions about both aspects. The students' person measures reflected their degrees of interest and the items' item measures reflected their degrees of interestingness. Using mixed Rasch models I could show how interesting different contents and contexts were relative to each other for different groups of students and each content area. However, I chose to apply the Rasch model separately to the data about both content areas (instead of applying it to all data combined). Here, I followed the approach used in the IPN study to ensure comparability of the results. They had chosen this approach because they had eight content areas and the mixed Rasch model cannot be applied to large item sets (88 items in their case). In future studies, it would be very interesting to investigate items that present different contents from additional content areas set in different contexts based on the same context categories and see whether the students' types of interest found in my study can also be replicated. Overall, for future studies I recommend considering interest in specific content areas and corresponding contents instead of physics in general.

Moreover, I decided on, discussed, and provided a broad definition of physics-related self-concept comprising two aspects, the 'self-perception of physics ability' and the 'beliefs about the perception as a physics person by others' (perceived recognition). This approach was based on recent study results on the importance of perceived recognition by Kalender et al. (2019a, 2019b). Indeed, applying a mixed Rasch model, the item set formed a one-dimensional scale. In addition, I could show how the two assessed aspects contribute to the physics-related self-concept of different types of students. For example, the type with overall lower physics-related self-concept had less positive beliefs about the perception as a physics person by others (perceived recognition) compared to the relatively more positive self-perception of ability (i.e. the two aspects contribute differently to the physics-related self-concept). This allows unique insights into the structure of the physics-related self-concept. Hence, the instrument to measure physics-related self-concept (i.e. ‘ability self-perception’ and ‘beliefs about the perception as a physics person by others’) and data analysis method (i.e. mixed Rasch models) used in my study could also be used as diagnostic tools by researchers interested in fostering students’ physics-related self-concept. For example, it would be an interesting direction for future research to investigate whether the types of self-concept differ for different age groups. In general, I recommend for future studies – in line with Kalender et al.'s recommendation to examine the ‘structure of multiple motivational factors together’ (2019a) – to combine two or more motivational constructs and analyse the corresponding data together using a mixed Rasch model. When combining the data regarding two (or more) aspects one can still draw conclusions about the individual aspects and their relationship with one another for different student groups by applying mixed Rasch models. Yet, it is a limitation of my study that I did not investigate the self-perception as a physics person as suggested by the physics identity studies, which can be easily assessed with one item (‘I see myself as a physics person’). It would be an interesting direction for future research to do the same type of mixed Rasch analysis but additionally including the self-perception as a physics person (i.e. to combine the aspects ‘self-perception of ability’, ‘beliefs about the perception by others as a physics person’, and ‘self-perception as a physics person’). Here, one could investigate how self-perception of ability, self-perception as a physics person, and beliefs about the perception as a physics person by others contribute relative to one another to the physics-related self-concept of different student groups. Moreover, I highlighted the differences and transitions of physics-related ability self-concept to the similar construct ‘self-efficacy’, which in contrast to physics-related ability self-concept refers to specific contents and tasks. Hence, I recommend for future studies to choose self-efficacy as a construct, if they want to assess the students’ abilities in a rather specific manner (i.e. referring to specific contents and tasks within different physics content areas), whereas they should choose ability self-concept, if they want to assess ability in a rather broad manner (i.e. referring to physics as a domain). Here, it also becomes clear that there is not a sharp distinction between the two constructs but rather a transition from one to the other. Since the interest items used in my study are rather specific, it could also be an interesting direction for future research to construct very specific self-efficacy items for every interest item and to investigate the students’ interest and its association with the corresponding specific self-efficacy.

Moreover, I decided on, discussed, and provided a rather uncommon approach to investigating the students’ previous experience. In particular, I focused on the students’ quantitative previous experience with the content area in school in contrast to other studies who either investigated the students’ knowledge or qualitative experiences (e.g. with parents, with teachers in the classroom, in out-of-school learning environments). This is a strength of my study as it shifted the focus from knowledge to previous experience. This approach provided the opportunity to present a broader picture of the association between interest and knowledge. For example, it may be that at the time when their interest and knowledge are assessed the students do not know anymore the physics content but still have a feeling for how often they have experienced it in school. One limitation related to asking the students about their previous experience with the content areas in school may be that it is a subjective self-report. For example, students who are not interested as much might remember lessons on the (for them not interesting) content area less well or perceive their experience as lower for other reasons. In comparison, assessing the students’ knowledge would be more objective. However, one strength of my study is that the students were asked about the previous experience with several specific contents from each content area, which is more concrete than asking them for their general experience with mechanics and particle physics, respectively. In future studies, one may investigate the association between interest in and previous experience with a content area in school by aligning the wording of both item sets even more than done in my study. Another limitation, especially concerning the content area 'particle physics', is the focus on the previous experience in school (i.e. excluding out-of-school previous experience). Particle physics is usually not covered in school curricula at the investigated age group. Yet, interested students might have engaged a lot with particle physics in their leisure time (e.g. by watching videos). For example, a previous study by Woithe (2020) indicates that about one third of students (on average aged 17 years) gained their previous experience with particle physics out of school via online videos and websites. Focusing on previous experience in school only, I could not investigate the association between interest and previous quantitative out-of-school experience.

One limitation of my study is that I assessed sex instead of gender. Future studies could assess gender. Moreover, quantitative measurement instruments to assess queer identities could be used. For example, Dockendorff and Geist (2022) propose an instrument to assess feminine/masculine/androgynous self-perception and beliefs about the corresponding perception by other people. It would be interesting to investigate the relationship between gender/queer self-concept, physics-related self-concept, and interest.

Mixed Rasch models

One strength of my study is the use of (mixed) Rasch models to analyse the collected data on students' interest and physics-related self-concept. By conducting a mixed Rasch rating scale analysis for distinguishing different groups of interest as well as different groups of physics-related self-concept, I also investigated differential item functioning (DIF; Quandt, 2012). For example, I found that the items do function in the same manner for female and male students since the groups comprise a rather balanced ratio of female and male students. The advantage of using mixed Rasch analysis over standard DIF testing is that standard DIF testing is based on pre-defined categorisations, such as sex or age, which are routinely assessed (Quandt, 2012). Yet, the absence of DIF (e.g. regarding the students' sex) in standard testing does not mean that there is no DIF at all (Quandt, 2012). Other types of DIF (e.g. for the students' response style) might be overlooked (Quandt, 2012). I furthermore argue that DIF on pre-existing groups might reinforce stereotypes. In my study, I could show that there is DIF for response style using mixed Rasch models. Hence, the students were categorised into more than one interest and self-concept group, respectively, that could actually be described in terms of one single type of interest and self-concept, respectively. In sum, I suggest that testing for DIF using mixed Rasch models may be beneficial for other researchers in future studies. One limitation of my study is that I did not double check the DIF for response style by applying constrained mixed Rasch models as suggested by Wetzel et al. (2013). Hence, I also recommend this as a direction for future research.

Moreover, I recommend using mixed Rasch models in future studies to avoid generalising conclusions that only apply for a small group of students. Here, another strength of my study was that I could point towards such commonly reported generalisations (e.g. modern physics content areas being more interesting than classical ones) that may only be an artefact resulting from the interests of a subset of students (e.g. the 'particle physics lovers'). Such subsets that qualitatively differ in their interest can only be detected using a mixed Rasch analysis because other methods, such as multilevel analyses, are based on pre-defined categorisations (e.g. sex, class, high vs. low interest). Hence, mixed Rasch models allow unique insights into students' interest and physics-related self-concept. For example, using a mixed Rasch model one can show how interesting different contexts are relative to one another and whether there are different groups of students, of which the qualitative interests are similar within these groups but different across these groups.

Measurement instrument

One limitation of my study is that the rating scale format changes for the items assessing different constructs. Interest in a content area is assessed using five categories and physics-related self-concept is assessed using four categories. For both variables, I used the original rating scale format. Thus, I cannot fully compare the response styles of the students found separately concerning the interest and self-concept items with one another. Nevertheless, I found different groups of students that rather resemble an ERS or a NERS as well as groups with a rather 'normal' response style, that is, groups that use the rating scale in the intended manner, concerning both constructs. It would be a very interesting direction for future research to investigate whether students apply a certain response style, such as ERS, consistently across items used to assess different constructs, when using the same rating scale format.

One limitation of my study is that I did not assess the students' interest while they were conducting an actual task as, for example, done in the PISA 2006 study (OECD, 2007a) or by Rösler et al. (2014, 2018) in a biology education research project. This has two drawbacks. First, the students may not like physics as a school subject and respond to the presented item according to this aversion; that is, although they would be interested in the content, context, or task presented in the item, when conducting an actual task, they express relatively less interest in the non-embedded item because of their physics-aversion. Moreover, in the PISA 2006 study embedded items lead to strikingly different results than non-embedded ones (Stern et al., 2009). Especially the often-reported lower interest of girls evened out in embedded items (Stern et al., 2009). I claim that the difference evens out because when engaging with tasks, students' may not recognise which tasks belong to which domain, and thus gender stereotypical answers disappear. Second, in the classroom the students encounter different physics contents, contexts, and tasks during actual activities, which cannot be resembled by non-embedded items. Hence, I recommend for future interest studies to use embedded items for assessing interest to not trigger an eventually existing aversion against physics as a school subject and to resemble actual classroom situations. However, it may also be considered as an advantage of my study that I am assessing the students' interest in a non-embedded way. I argue that it enables me to draw conclusions about the contexts that may trigger the students' operating interest, which may either be caused by the already existing individual interest of the students or it may be a situational interest triggered by the content, contexts, and tasks presented in the items acting as 'catch factors' (Mitchell, 1993). Hence, the recommendations given based on my results may be especially relevant and useful for engaging students at the beginning of a lesson. One limitation of my study is that I am measuring operating interest. I do not assess whether their interest is caused by the students' disposition or the interestingness of the situation; that is, I do not know whether the expressed operating interest reflects the students' actualised individual interest or their triggered situational interest.

Sample

Another limitation of my study is that the sample might be biased towards a higher amount of 'particle physics lovers' than existing. The teachers were contacted to participate in the study, and as a thank you they were offered a CERN science show related to particle physics or an overview talk about CERN. Hence, it might be that the teachers that participated in my study have personal interest in particle physics which they might also (un)intentionally foster in their students. Thus, I argue that the type PP_2 ('particle physics lovers') might comprise less than the observed 21% of students. Moreover, that the teachers participated with their students in a physics education research project might indicate that they are more dedicated than other teachers, which might also influence their students' interest in physics. One strength related to my sample is the age group (i.e. students aged 14 to 16 years, mostly in grade 9). Previous research (e.g. Osborne et al., 2003) suggests that between the ages 10 and 14 interest in science is diminished. Hence, I could investigate the students' interest after this crucial phase. Moreover, the age group of my sample make my results comparable to the ones of large-scale studies on students' interest, such as the PISA and ROSE studies, which also focused on students aged 15 years. One limitation related to the age group is that the school curricula typically mention particle physics, if at all, rather in higher grades. Thus, one may argue that it would have been better to compare interest in particle physics and in mechanics using an older sample. However, one aim of my study was explicitly to compare students' interest in two content areas, with which they have different degrees of previous experience in school. One limitation related to my sample is that it was only German-speaking students. Here, one strength is nevertheless that it comprised students from three different German-speaking countries (Austria, Germany, and Switzerland). One strength of my study is the large sample size ($N > 1200$).

Hierarchy of levels of interest in physics (HOLIP)

One strength related to the proposed conceptualisation of students' interest as a hierarchy of levels of interest in physics (HOLIP) is that I could interpret the levels using the three components of interest (i.e. the emotional, value-related, and cognitive epistemic components) aligned with the 'Four-phase model of interest development' (Hidi & Renninger, 2006). First, I characterised the level of focused interest in physics as being interested in physics when it is set in contexts related to one's own body, socio-scientific issues, or existential questions of humankind, that is, contexts that arouse emotions, which aligns with the first phase in interest development. Second, students at the level of open interest were additionally interested in physics set in the context 'everyday life', that is, the students recognise personal value, which aligns with the second phase in interest development. Third, students at the level of broad interest were interested in physics, even when set in a purely scientific or technical context (i.e. addressing the cognitive-epistemic component of interest), which aligns with the third and fourth phases in interest development. Future studies could investigate students' interest in physics, especially focusing on the three components of interest, in addition to varying different contents and contexts. This would provide additional evidence on the conceptualisation of students' interest as a hierarchy of levels of interest in physics (HOLIP) aligned with the hierarchy of interest components in the 'Four-phase model of interest development' (Hidi & Renninger, 2006). Moreover, it would be a very interesting direction for future studies to conduct a longitudinal study to investigate whether the hierarchy of interest levels indeed reflect a development process. One strength related to my conceptualisation of students' interest as a hierarchy of levels of interest in physics (HOLIP) is that it provides a concise overview of which context are more (or less) interesting relative to each another and can be used by educators as a tool to develop interesting learning activities.

One limitation of my study is that tasks were included in the interest measurement instruments but not addressed in the hierarchy of the students' levels of interest in physics. I decided against this as the variation of tasks is not sufficiently large to analyse the influence of tasks on students' interest (as outlined above). However, in line with past empirical findings I could observe that the students generally find items that describe the tasks ‘conducting an experiment’ or ‘planning an experiment’ relatively more interesting. Moreover, the focus on contexts is also supported by the finding of Blankenburg et al.’s study (2016) that students who indicate higher interest in a certain context are more interested in all activities in that context. Future studies could nevertheless combine a variation of contexts and tasks to extend the hierarchy of level of interest to include tasks.

Combining replicative and innovative aspects

One strength of my study is that it is both (a) a partial replication study of the IPN study concerning the classical physics content area ‘mechanics’ and (b) an innovative extension to the modern physics content area ‘particle physics’. I could show that the IPN findings concerning the students’ interest types in mechanics can only partially be replicated and that the IPN findings can only partially be applied to particle physics. It would be an interesting direction for future research to investigate whether the IPN findings can be replicated or applied concerning further content areas.

6.5. Conclusions and implications for practice

Based on previous findings, I hypothesised that there are three different types of interest in physics that are valid for modern and classical physics content areas. Moreover, I hypothesised that different contexts are more (or less) interesting relative to each other within the students’ different types of interest. I conducted a cross-cohort study with German-speaking students aged 14 to 16 years (N = 1219). I assessed students’ interest in mechanics and particle physics, their physics-related self-concept, previous experience with both content areas in school, and sex. Applying the mixed Rasch rating scale model, I found that most students could be categorised into one single type of interest regarding both content areas (i.e. 86% of students regarding mechanics, 79% of students regarding particle physics), the type of students that were only interested in physics set in certain contexts referred to as ‘Physics? Only in certain contexts!’. Few students (i.e. 14 % of students) that were interested in mechanics set in the context ‘car’ formed a second type of interest in mechanics, the ‘students interested in physics relating to the motion of cars’; and few students (i.e. 21% of students) that were highly interested in particle physics as a scientific endeavour formed a second type of interest in particle physics, the ‘particle physics lovers’. This finding is important for particle physics outreach practices. Within the research community one may assume that everyone else is as interested in particle physics as a scientific endeavour as the community is. Yet, this result indicates that outreach efforts should focus on the contexts in which particle physics can be set and that are interesting for most students. Calculating logistic regression models, I found that whether the students belonged to one or the other interest type was significantly better described with a model that included both their sex and their degrees of physics-related self-concept compared to the baseline model. To describe the relative interestingness of different contexts for the type of students that was only interested in physics set in certain contexts, that is, the vast majority students, I could apply my conceptualisation of students’ interest as a hierarchy of levels of interest in physics (HOLIP). It can be used by educators as a tool to develop learning activities that are interesting for most students, regardless of their sex or physics-related self-concept. Although my conceptualisation of interest had originally been developed for particle physics, I found that it suited to describe most students’ interest in particle physics and in mechanics. Hence, I conclude that educators can apply the HOLIP to all physics content areas if they can be set in the suggested contexts. Moreover, I could also apply the mixed Rasch rating scale model to the students’ physics-related self-concept, and thereby show how the assessed aspects of physics-related self-concept (i.e. ‘self-perception of ability’ and ‘beliefs about the perception as a physics person by others’) contribute differently to the physics-related self-concept of the different types of physics-related self-concept. Finally, I also investigated the association between interest in a content area and different student characteristics, such as physics-related self-concept, sex, and previous experience with the content area in school.

Comparing my results to those of the IPN study, I could show that the definition of interest types by the IPN was based on the individual groups resulting from a mixed Rasch analysis. I argue that not all groups were actual interest types but rather occurred due to the different use of the rating scale of the students. Hence, two or more groups, which had a similar interest profile but differed in their rating scale use, could be combined to one type of interest. In this sense, also in the re-analysis of the IPN study conducted by Sievers (1999) only two types of interest were actually found, one type comprising students who are only interested in physics set in certain contexts and a second type of students who are rather equally interested in all items. In my study, using the original IPN instrument to measure students' interest in mechanics, I could only partially replicate their findings. However, I could show that also the interestingness of a modern physics content area can be analysed in terms of interest types.

I also conclude that my study provided further evidence that physics-related self-concept is a broad construct comprising several aspects, such as the assessed aspects 'self-perception of ability' and 'beliefs about the perception as a physics person by others'. Students in type SC_1 had on average lower physics-related self-concept values and agreed less to items on their beliefs about the perception as a physics person by others than to items on their self-perception of ability. For these students there was a difference between (1) what they think about their own abilities and (2) whether they think that others perceive them as a physics person. Students in type SC_2 had on average higher physics-related self-concept values and agreed to a similar extent to perceived recognition items and to ability self-concept items. For these students (1) what they think about their own abilities matches with (2) whether they think that others perceive them as a physics person.

For education research practice, I imply that researchers should de-emphasise the role of the students' sex or gender and instead focus on the students' physics-related self-concept as a reference variable. This has already been suggested in past empirical studies, for example, by Höft and Bernholt (2019). I acknowledge that it requires more items to assess and more effort to conduct analyses, such as regression models, using physics-related self-concept instead of sex. Even though previous studies found that other variables, such as self-efficacy, are also associated with interest, the focus of analysis usually is on gender issues, for example, how to make physics more interesting for girls. In different studies, topics were presented as interesting either for boys or for girls using phrasings, such as 'girls are interested in' and 'boys are interested in'. Yet, the results of data analyses of different studies indicated that in most cases the allocated girls' topic was equally interesting for boys as the boys' topic, but not the other way round. This has, for example, been found by Häußler and Hoffmann (1998; IPN study). Moreover, Jenkins and Nelson (2005; ROSE study) emphasise that a high interest in a certain topic by one sex or gender does not necessarily mean that the same topic is not interesting for the other. In sum, by focusing on sex or gender as analysis variables one may reemphasise differences that may not even exist to the commonly assumed extent. In addition, one may reemphasise that girls are 'problematic' in physics and need particular support and that one needs to adjust teaching practices to the girls' interests. Yet, I argue that it is not the girls but the students who have lower physics-related self-concept and the reasons why some students have a lower self-concept that one should focus on. I found that the students with on average lower physics-related self-concept agreed less to perceived recognition items than to ability self-concept items; that is, they rather struggled with whether they think that others see them as a physics person than with their self-perception of ability. Hence, focusing on girls (and thereby unwillingly reemphasising that girls are 'problematic' in physics and need particular support) may not help them to feel more perceived as a physics person (i.e. as a person apt to do physics). Similarly, previous studies indicate that students may differ in the way they create self-concept with male students (i.e. students who commonly report higher degrees of self-concept) using mastery experience as source whereas it was 'social persuasion and vicarious experiences' for female students (i.e. students who commonly report lower degrees of self-concept; Zeldin et al., 2008). I conclude that education research needs a sex- or gender-neutral clustering variable to truly reflect that it is not the sex but the physics-related self-concept that matters. I also acknowledge that it may appear better to focus on gender issues than on self-concept as it enables researchers to provide rather straightforward recommendations for educational practice as it is easier for educators to assign students to gender than to self-concept. However, focussing on the similarities, that is, the contents and contexts that are equally and highly interesting for most students regardless of their sex or gender and self-concept, even better reflects the classrooms that teachers commonly encounter.

Current physics education mainly presents physics in the contexts 'science and technology'. Yet, already the science-technology-society initiative in the 1970s and 80s aimed to increase the students' interest by brining society as a context in the physics classroom (Zoller & Watson, 1974; Aikenhead, 1980). My study indeed showed that the contexts 'science' and 'technology' are only interesting for few students. Instead, I found that most students can be described with one single type of interest and that I can describe their interest with my conceptualisation of interest as a hierarchy of levels of interest in physics (HOLIP). Although the type of students interested in physics relating to the motion of cars (M_2) regarding mechanics and the particle physics lover type of students (PP_2) regarding particle physics could only to some extent be described with this conceptualisation, I recommend that educators match their learning activities with the HOLIP conceptualisation. The type of students interested in physics relating to the motion of cars and the particle physics lover type of students are most likely still interested in mechanics and particle physics, respectively, if set in contexts that are the most interesting for most students, although for them other contexts may be relatively more interesting. I conclude that by matching the design of learning activities to the type of students that is only interested in physics set in certain contexts, the type of students interested in physics relating to the motion of cars and the particle physics lover type of students will not lose their interest in mechanics and particle physics, respectively. For example, for the particle physics lover type of students, particle physics in the context 'scientific endeavour' (e.g. quantitative physics in the context of particle collisions) is more interesting than in everyday life contexts (e.g. digital camera as particle detector). Yet, they are also highly interested in physics when set in an everyday life context. In contrast, particle physics in the context 'scientific endeavour' will only be interesting for the particle physics lover type students, and not for most students. Hence, I imply for educational practice that educators can apply the HOLIP conceptualisation to all physics content areas if they can be set in the suggested contexts, and thereby match the design of their learning activities to the interests of most students regardless of their sex or gender and physics-related self-concept.

7. SUMMARY OF THE MAIN FINDINGS

The summary presented in this chapter provides a compact overview of the main finding of my doctoral research project. The first study aimed to develop the instrument to measure particle physics interest (IPPI) and to propose a hierarchy of levels of interest in particle physics. Second, the main study sought to categorise students into different types of interest in mechanics and particle physics, to investigate the applicability of the proposed hierarchy of levels of interest interest to the different types of interest in mechanics and in particle physics, and the association of interest with the assessed student characteristics ‘physics-related self-concept’ and ‘previous experience with the content areas in school’.

Introducing a novel approach in selecting items from an initial item pool

The IPPI was created based on the 11 item categories introduced in the IPN study (Häußler, Lehrke, et al., 1998). Initially, I created at least three items for each category. Applying the Rasch partial credit model, I ultimately selected one item per category following clear, stepwise, and reproducible criteria based on the category probability curves, item fit indices, and the sign of an even distribution of items on the Wright Map. The Rasch analysis provided evidence supporting content, construct, statistical, and fit validity of the IPPI. The results demonstrate that I have successfully developed a valid and reliable instrument to measure interest in particle physics. This novel approach in selecting items from an initial item pool may be useful for other researchers developing instruments.

Using different introductory texts is associated with small qualitative differences in interest

In the main study, I used two different versions of the mechanics introductory text in combination with the same items. Overall, there was no statistically significant, quantitative difference in the students’ interest associated with the use of different versions of the introductory text. Instead, the results provide some evidence that there were qualitative differences in interest; that is, certain items seem to have been relatively more interesting for some students who were presented with the original IPN introductory text. I discussed several possible reasons. For example, I argued that the original IPN introductory text mostly focused on the context ‘car’, and hence the interest of some students in all contents set in the context ‘car’ might have been triggered. In contrast, the newly developed introductory text presented a variety of contexts, and hence the students might have been more ‘sensitive’ to the different contexts presented in the items. Indeed, when combining the datasets for the mixed Rasch analysis, these small qualitative differences also resulted in four interest groups instead of three as found in the separate analyses.

Redefinition of students’ types of interest in physics

Overall, I could distinguish three types of interest in physics in the main study. For both content areas, I could replicate the IPN interest type C (Langeheine et al., 2001; Rost et al.,1999; Sievers, 1999). This type comprised students that are only interested in physics set in certain contexts. It was referred to as type M_1 for mechanics and PP_1 for particle physics and comprised 84% and 79% of the students, respectively. However, I could not replicate the IPN interest type NG (Langeheine et al., 2001; Rost et al.,1999; Sievers, 1999) using my approach, since students with differing response style were also included in the interest type M_1 and PP_1, respectively. Second, I could partially replicate the IPN interest type A (Langeheine et al., 2001; Rost et al.,1999; Sievers, 1999) for both content areas. The type M_1 of interest in mechanics (14% of the students) comprised students that are relatively more interested in mechanics contents set in the context 'car' and I referred to this type as the 'students interested in physics relating to the motion of cars'. However, their quantitative degree of interest in mechanics was not significantly different than of the other type, which does not align with the IPN interest type A. Yet, all items were rather equally interesting for the type of students interested in physics relating to the motion of cars, which may resemble the IPN interest type A. The type PP_2 of interest in particle physics (21% of the students) comprised students who were interested in items presenting the specifics of particle physics as a content area and as a scientific endeavour. They had the highest mean interest in particle physics. Hence, I referred to this type as the 'particle physics lovers'. Yet, for the particle physics lover type of students not all items were equally and highly interesting for the students which does not align with the IPN interest type A. For example, the item PP05 (masses of particles) presenting quantitative particle physics without being set in a context was rated relatively less interesting. Despite the qualitative differences in interest of the observed interest types, the Rasch model could be applied for all students and both content areas, and hence my instruments to measure interest were found to be unidimensional.

The hierarchy of levels of interest in physics (HOLIP) applies to most students

I introduced the conceptualisation of students' interest in particle physics comprising three levels of interest to describe the item hierarchy in my first study (Zoechling et al., 2022; see **Chapter 5**). In my main study about the students' interest in particle physics and mechanics, this conceptualisation could also be applied to the item hierarchy of the type of students that is only interested in physics set in certain contexts, that is, most students in mechanics (type M_1, 84% of the students) and particle physics (type PP_1, 79% of the students; see **Chapter 6**). Hence, the conceptualisation was suggested as a guideline for physics in general and renamed 'hierarchy of students' levels of interest in physics' (HOLIP, see **Figure 18**). The HOLIP provides a concise overview of how interesting different contexts (i.e. storylines), in which physics content may be set, are relative to each other. That is, different levels of interest among students were mainly determined by the context in which the physics content was set. I also interpreted the hierarchy of students' levels of interest based on the results of past empirical research and the 'Four-phase model of interest development' (Hidi & Renninger, 2006). I found that: (1) Students at a focused level of interest are interested in physics when set in a context, which arouses emotions. That is, the most interesting contexts are related to one's own body (e.g. 'artificial joints in medicine', 'medical diagnostics', 'driver'), socio-scientific issues (e.g. 'car accident', 'smuggled arms'), and existential questions of humankind (e.g. '*Where do we come from?*'). (2) Students at an open level of interest were additionally interested in physics when it was set in an everyday life context (e.g. 'mobile phone'). (3) Students at a level of broad interest were even interested in physics when it was set in scientific and technical contexts (e.g. 'electronics industry', 'car repair shop'). When interpreting the hierarchy of students' levels of interest, it is crucial to consider that these levels are cumulative. This means that the level of broad interest includes the level of open interest, which further includes the level of focused interest. In contrast, I found a second type of interest in both content areas which cannot be described with the conceptualisation. The students assigned to the second type of interest in mechanics were mostly interested in the context 'car' (type M_2, 16% of the students). The students assigned to the second type of interest in particle physics (type PP_2, 21% of the students) were mostly interested in the special aspects of particle physics as a content area and as a scientific endeavour.

Mechanics and particle physics are equally interesting

My main study provided evidence that most students (type M_1 and PP_1) were equally interested in both content areas. Yet, particle physics was found to be more interesting than mechanics for the type of students interested in physics relating to the motion of cars regarding mechanics (M_2) and the particle physics lover type of students regarding particle physics (PP_2). This finding contrasts with past empirical studies which suggested that modern physics contents, such as the history of the universe, may be more interesting than classical ones (e.g. OECD, 2016; Sjøberg & Schreiner, 2012; Häußler, Lehrke, et al., 1998). The findings of my main study indicate that such generalised observations may only result from the differences in interest of a small fraction of students and cannot be applied to most students. Most importantly, these findings emphasise the importance of analysing data concerning students' interest using methods, such as mixed Rasch analyses, that do not rely on pre-defined categorisations (e.g. sex) but instead distinguish groups based on their responses to the used interest measurement instrument.

Definition of students' types of physics-related self-concept

In my main study, I used a broad definition of self-concept comprising two aspects, namely self-perception of ability and beliefs about the perception as a physics person by others. The results of my main study provide evidence that there are two different types of physics-related self-concept. For 60% of the students, referred to as the type SC_1, the two assessed aspects of physics-related self-concept contributed differently to their self-concept; that is, the students had a more positive self-perception of ability than beliefs about the perception as a physics person by others. The remaining 40% of the students, referred to as the type SC_2, had equivalent self-perception of ability and beliefs about the perception as a physics person by others. Despite the different contributions of the two aspects to the students' self-concept for the two types, the Rasch model could be applied for all students and my instrument to measure self-concept was found to be unidimensional. This supports a broad definition of physics-related self-concept and further contributes to the empirical background for assessing it in future studies.

Physics-related self-concept as a mediator for sex differences in interest

The main study showed that there were overall no significant differences in mean interest between the sexes. There was a small difference for the particle physics lover type of students concerning the content area 'particle physics' with male students showing slightly higher mean interest than female students. The association between physics-related self-concept and interest was found to be similar for both sexes and both interest types concerning both content areas using linear regression/correlation analyses. Yet, the mean physics-related self-concept significantly differed between the sexes, with girls having lower mean physics-related self-concept. This effect was relatively larger for the type of self-concept, mechanics interest, and particle physics interest, respectively, which had the higher mean physics-related self-concept (i.e. SC_2, M_2, and PP_2). These results combined provided evidence that the students differ in their degrees of interest because of their differences in physics-related self-concept. Yet, the physics-related self-concept significantly differed for female and male students. Hence, my main study provided evidence that the direct association between sex and interest is less descriptive for the students' interest than the association between sex and interest with physics-related self-concept as a mediator. This is because the sexes significantly differed in their mean self-concept but the correlation between self-concept and interest was similar for both sexes. These findings provided evidence that the association of sex and interest is mediated via self-concept.

Small positive correlation between previous experience in school and interest

The main study showed that students have significantly different previous experience with both content areas in school. They reported having more previous experience with mechanics compared to particle physics in school. This was expected based on the higher representation of mechanics in school curricula in contrast to the little representation of particle physics. The mechanics interest types differed significantly in their previous experience with mechanics; that is, the type of students interested in physics relating to the motion of cars (M_2) had more experience than the type of students who were only interested in physics in certain context (M_1). The particle physics interest types did not differ significantly in their previous experience with particle physics, which also was expected based on the almost non-existing representation of particle physics in school curricula for students aged 14 to 16 years. My main study provided evidence that for most students (type M_1 and PP_1) there is a very small positive significant correlation between previous experience in school with and interest in mechanics and particle physics, respectively ($r_{M,E,1}=0.16$; $r_{PP,E,1}=0.19$). Yet, for the students interested in physics relating to the motion of cars (M_2) and the particle physics lovers (PP_2), there was no significant correlation. Past empirical studies differed in their observed correlations between the students' interest and their knowledge; correlations from 0 to strong are reported (Schiefele et al., 1992). Although knowledge and previous experience are different variables, they are associated with one another, and hence a reference to this field of research seems beneficial. The results of my main study provided evidence that such generalised, non-agreeing observations may result from the differences in correlations between most students (types M_1 and PP_1) and a small fraction of the students (types M_2 and PP_2).

8. REFERENCES

Adams, W. K., Perkins, K. K., Podolefsky, N. S., Dubson, M., Finkelstein, N. D., & Wieman, C. E. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Physical Review Special Topics - Physics Education Research, 2*(1), 010101. https://doi.org/10.1103/PhysRevSTPER.2.010101

Adams, R. J., Wu, M. L., & Wilson, M. (2012). The Rasch Rating Model and the Disordered Threshold Controversy. *Educational and Psychological Measurement, 72*(4), 547–573. https://doi.org/10.1177/0013164411432166

Aikenhead, G. S. (1980). *Science in Social Issues: Implications for Teaching*. Publications Office, Science Council of Canada.

Ainley, M., & Ainley, J. (2011). A cultural perspective on the structure of student interest in science. *International Journal of Science Education*, *33*(1), 51–71. https://doi.org/10.1080/09500693.2010.518640

Ainley, M., & Hidi, S. (2014). Interest and enjoyment. In R. Pekrun & L. Linnenbrink-Garcia (Eds.), *International handbook of emotions in education* (pp. 205–227). Routledge/Taylor & Francis Group.

Andrich, D. (1978). A rating formulation for ordered response categories. *Psychometrika, 43*, 561-573.

Andrich, D. (2005). The Rasch model explained. In S. Alagumalai, D. D. Durtis, & N. Hungi (Eds.), *Applied Rasch measurement: A book of exemplars* (pp. 308-328). Berlin, Germany: Springer-Kluwer.

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2010). “Doing” science versus “being” a scientist: Examining 10/11-year-old schoolchildren's constructions of science through the lens of identity. *Science Education, 94*(4), 617–639. https://doi.org/10.1002/sce.20399

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2013). ‘Not girly, not sexy, not glamorous’: primary school girls’ and parents’ constructions of science aspirations. *Pedagogy, Culture & Society, 21*(1), 171–194. https://doi.org/10.1080/14681366.2012.748676

Aschbacher, P. R., Li, E., & Roth, E. J. (2010). Is science me? High school students' identities, participation and aspirations in science, engineering, and medicine. *Journal of Research in Science Teaching, 47*(5), 564-582. https://doi.org/10.1002/tea.20353

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Baumeister, R. (1998). The self. In D. Gilbert, S. Fiske, & G. Lindzey (Eds.), *Handbook of social psychology* (4th ed., pp. 680–740). New York: McGraw-Hill.

Bennett, J., Lubben, F., & Hogarth, S. (2007). Bringing science to life: A synthesis of the research evidence on the effects of context-based and STS approaches to science teaching. *Science Education*, *91*(3), 347–370. https://doi.org/10.1002/sce.20186

Blankenburg, J. S., Höffler, T. N., Peters, H., & Parchmann, I. (2016). The effectiveness of a project day to introduce sixth grade students to science competitions. *Research in Science and Technological Education*, *34*(3), 342–358. https://doi.org/10.1080/02635143.2016.1222361

Blankenburg, J. S., & Scheersoi, A. (2018). Interesse und Interessenentwicklung [Interest and development of interest]. In D. Krüger, I. Parchmann & H. Schecker (Eds.), *Theorien in der naturwissenschaftsdidaktischen Forschung*(pp. 245–259). Springer. https://doi.org/10.1007/978-3-662-56320-5_15

Bleeker, M. M., & Jacobs, J. E. (2004). Achievement in Math and Science: Do Mothers' Beliefs Matter 12 Years Later? *Journal of Educational Psychology, 96*(1), 97–109. https://doi.org/10.1037/0022-0663.96.1.97

Bond, T. G., Yan, Z., & Heene, M. (2020). *Applying the Rasch model: Fundamental measurement in the human sciences* (fourth edn.). Routledge. https://doi.org/10.4324/9780429030499

Bong, M., & Clark, R. E. (1999). Comparison between self-concept and self-efficacy in academic motivation research. *Educational Psychologist, 34*(3), 139–153. https://doi.org/10.1207/s15326985ep3403_1

Boone, W. J., & Staver, J. R. (2020). *Advances in Rasch analyses in the human sciences*. Springer. https://doi.org/10.1007/978-3-030-43420-5

Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch analysis in the human sciences*. Springer. https://doi.org/10.1007/978-94-007-6857-4

Bøe, M., Henriksen, E., Lyons, T. & Schreiner, C. (2011). Participation in science and technology: young people’s achievement-related choices in late-modern societies, *Studies in Science Education, 47*(1), 37–72, https://doi.org/10.1080/03057267.2011.549621

Buccheri, G., Gürber, N. A., & Brühwiler, C. (2011). The Impact of Gender on Interest in Science Topics and the Choice of Scientific and Technical Vocations. *International Journal of Science Education, 33*(1), 159–178. https://doi.org/10.1080/09500693.2010.518643

Bundesrecht konsolidiert [Austrian federal law consolidated] (2022). Gesamte Rechtsvorschrift für Lehrpläne – allgemeinbildende höhere Schulen [Complete act for curricula – general secondary schools] (Publication no. BGBl 88/1985). https://www.ris.bka.gv.at/GeltendeFassung.wxe?Abfrage=Bundesnormen&Gesetzesnummer=10008568

Bybee, R. & McCrae, B. (2011). Scientific Literacy and Student Attitudes: Perspectives from PISA 2006 science. *International Journal of Science Education, 33*(1), 7–26. https://doi.org/10.1080/09500693.2010.518644

Carlone, H. B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching, 44*(8), 1187-1218. https://doi.org/10.1002/tea.20237

Carlone, H., Huffling, L., Tomasek, T., Hegedus, T., Matthews, C., Allen, M. & Ash, M. (2015). ‘Unthinkable’ Selves: Identity boundary work in a summer field ecology enrichment program for diverse youth, *International Journal of Science Education, 37*(10), 1524-1546, https://doi.org/10.1080/09500693.2015.1033776

Cheung, D. (2018). The key factors affecting students’ individual interest in school science lessons. *International Journal of Science Education*, *40*(1), 1–23. https://doi.org/10.1080/09500693.2017.1362711

Chung, J., Cannady, M. A., Schunn, C., Dorph, R., & Bathgate, M., (2016) *Measures Technical Brief: Fascination in Science.* Available online: http://activationlab.org/wp-content/uploads/2018/03/Fascination-Report-3.2-20160331.pdf (accessed on 27 February 2023).

Cleaves, A. (2005). The formation of science choices in secondary school. *International Journal of Science Education, 27*(4), 471-486. https://doi.org/10.1080/0950069042000323746

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Routledge. https://doi.org/10.4324/9780203771587

Cooley, C. H. (1902). *Human Nature and the Social Order*. New York: Scribner's Sons

Cross, S. and Bagilhole, B. (2002). Girls' Jobs for the Boys? Men, Masculinity and Non-Traditional Occupations. *Gender, Work and Organization, 9*(2), 204–226. https://doi.org/10.1111/1468-0432.00156

Dasgupta, N., & Asgari, S. (2004). Seeing is believing: Exposure to counterstereotypic women leaders and its effect on the malleability of automatic gender stereotyping. *Journal of Experimental Social Psychology 40*(5), 642–658. https://doi.org/10.1016/j.jesp.2004.02.003

de Barba, P. G., Kennedy, G. E., & Ainley, M. D. (2016). The role of students' motivation and participation in predicting performance in a MOOC: Motivation and participation in MOOCs. *Journal of Computer Assisted Learning*, *32*(3), 218–231. https://doi.org/10.1111/jcal.12130

Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977), Maximum Likelihood from Incomplete Data Via the *EM* Algorithm. *Journal of the Royal Statistical Society: Series B (Methodological), 39*(1), 1-22. https://doi.org/10.1111/j.2517-6161.1977.tb01600.x

DeWitt, J., Osborne, J., Archer, L., Dillon, J., Willis, B., & Wong, B. (2013). Young Children's Aspirations in Science: The unequivocal, the uncertain and the unthinkable. *International Journal of Science Education, 35*(6), 1037-1063. https://doi.org/10.1080/09500693.2011.608197

DeWitt, J., & Archer, L. (2015). Who Aspires to a Science Career? A comparison of survey responses from primary and secondary school students. *International Journal of Science Education, 37*(13), 2170-2192. https://doi.org/10.1080/09500693.2015.1071899

Dierks, P. O., Höffler, T. N., & Parchmann, I. (2014) Profiling interest of students in science: Learning in school and beyond, *Research in Science & Technological Education, 32*(2), 97–114, https://doi.org/10.1080/02635143.2014.895712

Dierks, P. O., Höffler, T. N., Blankenburg, J. S., Peters, H., & Parchmann, I. (2016). Interest in science: a RIASEC-based analysis of students' interests. *International Journal of Science Education, 38*(2), 238–258. https://doi.org/10.1080/09500693.2016.1138337

Dockendorff, K. J., & Geist, C. (2022). Beyond the Binary: Gender Image and Experiences of Marginalization on Campus, *Journal of Critical Scholarship on Higher Education and Student Affairs*, 6 (1). Available online: https://ecommons.luc.edu/jcshesa/vol6/iss1/3 (accessed on 8 March 2023).

Drechsel, B., Carstensen, C., & Prenzel, M. (2011). The Role of Content and Context in PISA Interest Scales: A study of the embedded interest items in the PISA 2006 science assessment. *International Journal of Science Education*, *33*(1), 73–95. https://doi.org/10.1080/09500693.2010.518646

Duranti, A., & Goodwin, C. (1992). Rethinking Context: An introduction. In A. Duranti, & C. Goodwin (Eds.), *Rethinking context: Language as an interactive phenomenon* (pp. 1–42). Cambridge University Press.

Eccles, J. S. (2009). Who am I and what am I going to do with my life? Personal and collective identities as motivators of action. Educational Psychologist, 44(2), 78–89. https://doi.org/10.1080/00461520902832368

Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data* (revised edn.). Bradford Books/The MIT Press.

Erikson, E. H. (1968). *Identity youth and crisis*. New York: Norton.

Eurofound and European Commission Joint Research Centre (2021). *European Jobs Monitor 2021: Gender gaps and the employment structure*. European Jobs Monitor series, Publications Office of the European Union. https://doi.org/10.2806/16416

European Commission (2000). *Women and Men in Paid Employment in the European Union*. Luxembourg: Office for the Publications of the European Communities.

Fechner, S. (2009). *Effects of context oriented learning on student interest and achievement in chemistry education*. Berlin: Logos.

Fechner, S., Van Vorst, H., Kölbach, E., & Sumfleth, E. (2015). It's the Situation That Matters: Affective Involvement in Context-Oriented Learning Tasks. In *Affective Dimensions in Chemistry Education* (pp. 159-176). Springer.

Field, A. (2009). *Discovery Statistics Using SPSS: and sex, and drugs and rock'n'roll*. Los Angeles, CA: Sage.

Finkelstein, N. (2005). Learning Physics in Context: A study of student learning about electricity and magnetism. *International Journal of Science Education, 27*(10), 1187–1209. https://doi.org/10.1080/09500690500069491

Fouad, N. A., & Smith, P. L. (1996). A test of a social cognitive model for middle school students: Math and science. *Journal of Counseling Psychology, 43*(3), 338–346. https://doi.org/10.1037/0022-0167.43.3.338

Fouad, N. A., Smith, P. L., & Zao, K. E. (2002). Across academic domains: Extensions of the social-cognitive career model. *Journal of Counseling Psychology, 49*(2), 164–171. https://doi.org/10.1037/0022-0167.49.2.164

Frey, A., Taskinen, P., & Schütte, K. (2009). *PISA 2006 Skalenhandbuch. Dokumentation der Erhebungsinstrumente [PISA 2006 scale book. Documentation of the measurement instruments].* Waxmann.

Galton, M. (2009). Moving to secondary school: Initial encounters and their effects. *Perspectives on Education*, *2*, 5–21.

Gardner, P. L., & Tamir, P. (1989). Interest in biology. Part I: A multidimensional construct. *Journal of Research in Science Teaching, 26*(5), 409–423. https://doi.org/10.1002/tea.3660260506

Gayles, J. G., & Ampaw, F. D. (2011). Gender matters: An examination of differential effects of the college experience on degree attainment in STEM. *New Directions for Institutional Research, 2011*(152), 19-25. https://doi.org/10.1002/ir.405

Gilbert, J. K. (2006). On the nature of "context" in chemical education. *International Journal of Science Education*, *28*(9), 957–976. https://doi.org/10.1080/09500690600702470

Godwin, A., Potvin, G., Hazari, Z., & Lock, R. (2016). Identity, Critical Agency, and Engineering: An Affective Model for Predicting Engineering as a Career Choice. *Journal of Engineering Education, 105*(2), 312–340. https://doi.org/10.1002/jee.20118

Grice, J. W., & Barrett, P. T. (2014). A Note on Cohen's Overlapping Proportions of Normal Distributions. *Psychological Reports, 115*(3), 741-747. https://doi.org/10.2466/03.PR0.115c29z4

Habig, S., Blankenburg, J., van Vorst, H., Fechner, S., Parchmann, I., & Sumfleth, E. (2018). Context characteristics and their effects on students' situational interest in chemistry, *International Journal of Science Education, 40*(10), 1154–1175, https://doi.org/10.1080/09500693.2018.1470349

Härtig, H., Nordine, J. C., & Neumann, K. (2020). Contextualization in the assessment of students' learning about science. In I. Sánchez Tapia (Ed.), *International perspectives on the contextualization of science education* (pp. 113–144). Springer. https://doi.org/10.1007/978-3-030-27982-0_6

Häußler, P. (1987). Measuring students' interest in physics - Design and results of a cross-sectional study in the Federal Republic of Germany, *International Journal of Science Education, 9*(1), 79-92, https://doi.org/10.1080/0950069870090109

Häußler, P., Bünder, W., Duit, R., Gräber, W., & Mayer, J. (1998). *Naturwissenschaftsdidaktische Forschung: Perspektiven für die Unterrichtspraxis*. Institut für die Pädagogik der Naturwissenschaften.

Häußler, P., Hoffman, L., Langeheine, R., Rost, J., & Sievers, K. (1996). Qualitative Unterschiede im Interesse an Physik und Konsequenzen für den Unterricht [Qualitative differences in interest in physics and consequences for teaching]. *Zeitschrift für Didaktik der Naturwissenschaften, 2*(3), 57–69.

Häußler, P., Hoffman, L., Langeheine, R., Rost, J., & Sievers, K. (1998). A typology of students' interest in physics and the distribution of gender and age within each type. *International Journal of Science Education, 20*(2), 223–238, https://doi.org/10.1080/0950069980200207

Häußler, P., Lehrke, M., & Hoffmann, L. (1998). *Die IPN-Interessenstudie Physik [The IPN study about interest in physics]*. IPN Kiel.

Häußler, P. and Hoffmann, L. (2000). A curricular frame for physics education: Development, comparison with students' interests, and impact on students' achievement and self-concept. *Science Education, 84*(6), 689–705. https://doi.org/10.1002/1098-237X(200011)84:6<689::AID-SCE1>3.0.CO;2-L

Hazari, Z., Sonnert, G., Sadler, P., & Shanahan, M. (2010) Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study, *Journal of Research in Science Teaching, 47*(8), 978–1003. https://doi.org/10.1002/tea.20363

Hidi, S. (2006). Interest: A unique motivational variable. *Educational Research Review, 1*(2), 69–82. https://doi.org/10.1016/j.edurev.2006.09.001

Hidi, S., & Renninger, K. A. (2006). The four-phase model of interest development. *Educational Psychologist*, *41*(2), 111–127. https://doi.org/10.1207/s15326985ep4102_4

Hoffmann, L., Krapp, A., Renninger, K. A., & Baumert, J. (1998) *Interest and Learning, Proceedings of the Seeon Conference on Interest and Gender*. Kiel.

Höft, L., Bernholt, S., Blankenburg, J. S., & Winberg, M. (2019). Knowing more about things you care less about: Cross-sectional analysis of the opposing trend and interplay between conceptual understanding and interest in secondary school chemistry. *Journal of Research in Science Teaching, 56*(2), 184-210. https://doi.org/10.1002/tea.21475

Höft, L., & Bernholt, S. (2019). Longitudinal couplings between interest and conceptual understanding in secondary school chemistry: an activity-based perspective. *International Journal of Science Education, 41*(5), 607-627. https://doi.org/10.1080/09500693.2019.1571650

Holland, J. L. (1997). *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments, 3rd ed*. Edessa, FL: Psychological Assessment Resources.

Holstermann, N., & Bögeholz, S. (2007). Interesse von Jungen und Mädchen an naturwissenschaftlichen Themen am Ende der Sekundarstufe I [Gender-specific interests of adolescent learners in science topics]. *Zeitschrift für Didaktik der Naturwissenschaften, 13*, 71–86.

Hox, J. J., Moerbeek, M., & Van de Schoot, R. (2017). *Multilevel analysis: Techniques and Applications*. Routledge.

Hyde, J. S. (2005). The gender similarities hypothesis. *American Psychologist, 60*(6), 581–592. https://doi.org/10.1037/0003-066X.60.6.581

International Baccalaureate Organization (2014). *Diploma Programme Physics Guide: First Assessment 2016*. Available online: http://www.holyheart.ca/wp-content/uploads/2016/10/IB-Physics-Guide-2016.pdf (accessed on 26 February 2023).

Jacobs, J. E., & Eccles, J. S. (1992). The impact of mothers' gender-role stereotypic beliefs on mothers' and children's ability perceptions. *Journal of Personality and Social Psychology, 63*(6), 932–944. https://doi.org/10.1037/0022-3514.63.6.932

Jacobs, J. E., Finken, L. L., Griffin, N. L., & Wright, J. D. (1998). The Career Plans of Science-Talented Rural Adolescent Girls. *American Educational Research Journal, 35*(4), 681-704. https://doi.org/10.3102/00028312035004681

Jacobs, J. E., & Eccles, J. S. (2000). Chapter 14 - Parents, task values, and Real-Life achievement-related choices. In C. Sansone & J. M. Harackiewicz (Eds.), *Intrinsic and Extrinsic Motivation* (pp. 405-439). San Diego: Academic Press. https://doi.org/10.1016/B978-012619070-0/50036-2

James, W. (1890). *The principles of psychology.* New York: Henry Holt and Company.

Jirout, J., & Klahr, D. (2012). Children's scientific curiosity: In search of an operational definition of an elusive concept. *Developmental Review, 32*(2), 125–160. https://doi.org/10.1016/j.dr.2012.04.002

Jenkins, E. W., & Nelson, N. W. (2005). Important but not for me: Students' attitudes towards secondary school science in England. *Research in Science & Technological Education*, *23*(1), 41–57. https://doi.org/10.1080/02635140500068435

Kalender, Z. Y., Marshman, E., Schunn, C. D., Nokes-Malach, T. J., & Singh, C. (2019a). Gendered patterns in the construction of physics identity from motivational factors. *Physical Review Physics Education Research*, *15*(2). https://doi.org/10.1103/PhysRevPhysEducRes.15.020119

Kalender, Z. Y., Marshman, E., Schunn, C., Nokes-Malach, T., & Singh, C. (2019b). Why female science, technology, engineering, and mathematics majors do not identify with physics: They do not think others see them that way. *Physical Review Physics Education Research, 15*(2), 020148. https://doi.org/10.1103/PhysRevPhysEducRes.15.020148

Kalender, Z. Y., Marshman, E., Schunn, C. D., Nokes-Malach, T. J., & Singh, C. (2020). Damage caused by women's lower self-efficacy on physics learning. *Physical Review Physics Education Research, 16*(1). https://doi.org/10.1103/PhysRevPhysEducRes.16.010118

Kashdan, T. B., & Silvia, P. J. (2009). Curiosity and interest: The benefits of thriving on novelty and challenge. In C. R. Snyder & S. J. Lopez (Eds.), *Oxford handbook of positive psychology*, *2nd edition* (pp. 367-374). Oxford University Press.

Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Association, 90*(430), 773-795. https://doi.org/10.1080/01621459.1995.10476572

Kauertz, A., & Fischer, H. E. (2006). Assessing students' level of knowledge and analysing the reasons for learning difficulties in physics by Rasch analysis. In X. Liu & W. J. Boone (Eds.), *Applications of Rasch measurement in science education* (pp. 212–246). JAM Press.

Kirschner, S., Borowski, A., Fischer, H. E., Gess-Newsome, J., & von Aufschnaiter, C. (2016). Developing and evaluating a paper-and-pencil test to assess components of physics teachers' pedagogical content knowledge. *International Journal of Science Education*, *38*(8), 1343–1372. https://doi.org/10.1080/09500693.2016.1190479

Kjærnsli, M. & Lie, S. (2011). Students' Preference for Science Careers: International comparisons based on PISA 2006, *International Journal of Science Education, 33*(1), 121-144, https://doi.org/10.1080/09500693.2010.518642

Kölbach, E. (2011). *Kontexteinflüsse beim Lernen mit Lösungsbeispielen [Influence of contexts on learning with example solutions].* Berlin: Logos.

Krapp, A. (2002a). Structural and dynamic aspects of interest development: Theoretical considerations from an ontogenetic perspective. *Learning and Instruction*, *12*(4), 383–409. https://doi.org/10.1016/S0959-4752(01)00011-1

Krapp, A. (2002b). An educational-psychological theory of interest and its relation to SDT. In E. L. Deci & R. M. Ryan (Eds.), *Handbook of self-determination research* (pp. 405–427). University of Rochester Press.

Krapp, A. (2005). Basic needs and the development of interest and intrinsic motivational orientations. *Learning and Instruction, 12*, 383–409.

Krapp, A., & Prenzel, M. (2011). Research on Interest in Science: Theories, methods, and findings. *International Journal of Science Education*, *33*(1), 27–50. https://doi.org/10.1080/09500693.2010.518645

Kuhn, J. (2010). *Authentische Aufgaben im theoretischen Rahmen von Instruktions- und Lehr-Lern-Forschung. Optimierung von Ankermedien für eine neue Aufgabenkultur im Physikunterricht*. Wiesbaden: Vieweg+Teubner Verlag.

Kuhn, J., Müller, A., Müller, W., & Vogt, P. (2010). Kontextorientierter Physikunterricht: Konzeptionen, Theorien und Forschung zu Motivation und Lernen. *Praxis der Naturwissenschaften–Physik in der Schule, 5*(59), 13–25.

Langeheine, R., Häußler, P., Hoffmann, L., Rost, J., & Sievers, K. (2001). Strukturelle Veränderungen des Physikinteresses von der 7. zur 9. Jahrgangsstufe. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie, 33*(1), 20–29. https://doi.org/10.1026//0049-8637.33.1.20

Lavonen, J., Ávalos, B., Upadyaya, K., Araneda, S., Juuti, K., Cumsille, P., Inkinen, J., & Salmela-Aro, K. (2021). Upper secondary students' situational interest in physics learning in Finland and Chile. *International Journal of Science Education, 43*(16), 2577–2596, https://doi.org/10.1080/09500693.2021.1978011

Lazarsfeld, P. F., & Henry, N. W. (1968). *Latent structure analysis.* Boston: Houghton Mifflin Co.

Lent, R. W., Brown, S. D., & Hackett, G. (1994). Toward a unifying social cognitive theory of career and academic interest, choice, and performance. *Journal of Vocational Behavior*, *45*(1), 79–122. https://doi.org/10.1006/jvbe.1994.1027

Lent, R.W., Hackett, G. and Brown, S.D. (1999). A Social Cognitive View of School-to-Work Transition. *The Career Development Quarterly, 47*: 297–311. https://doi.org/10.1002/j.2161-0045.1999.tb00739.x

Levene, H. (1961). *Robust Tests for Equality of Variances*. In: I. Olkin et al. (Eds.*), Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*. Stanford University Press (pp. 278-292).

Levrini, O., De Ambrosis, A., Hemmer, S., Laherto, A., Malgieri, M., Pantano, O., & Tasquier, G. (2017). Understanding first-year students' curiosity and interest about physics—Lessons learned from the HOPE project. *European Journal of Physics*, *38*(2), 025701. https://doi.org/10.1088/1361-6404/38/2/025701

Li, C. Y., Romero, S., Bonilha, H. S., Simpson, K. N., Simpson, A. N., Hong, I., & Velozo, C. A. (2018). Linking existing instruments to develop an activity of daily living item bank. *Evaluation and the Health Professions*, *41*(1), 25–43. https://doi.org/10.1177/0163278716676873

Linacre, J. M. (1991). Step disordering and Thurstone thresholds. *Rasch Measurement Transactions, 5* (3), 171.

Linacre, J. (2021). *A User's guide to WINSTEPS, MINISTEP, Rasch model computer programs.* Program Manual 5.1.7. https://www.winsteps.com/winman/copyright.htm Linacre, J. M. (2002). Optimizing rating scale category effectiveness. *Journal of Applied Measurement*, *3*(1), 85–106.

Lindqvist, A., Gustafsson Sendén, M. & Renström, E. (2021). What is gender, anyway: a review of the options for operationalising gender, *Psychology & Sexuality, 12*(4), 332-344, https://doi.org/10.1080/19419899.2020.1729844

Litman, J. (2005). Curiosity and the pleasures of learning: Wanting and liking new information. *Cognition and Emotion, 19*(6), 793-814, https://doi.org/10.1080/02699930541000101

Liu, X. (2010). *Using and developing measurement instruments in science education: A Rasch modeling approach*. Information Age Publishing.

Lohaus, A. & Vierhaus, M. (2019) *Entwicklungspsychologie des Kinder- und Jugendalters für Bachelor.* Berlin, Heidelberg: Springer. https://doi.org/10.1007/978-3-662-59192-5

Luce, M. R. & Hsi, S. (2015), Science-Relevant Curiosity Expression and Interest in Science: An Exploratory Study. *Science Education, 99*(1), 70-97. https://doi.org/10.1002/sce.21144

Luo, T., Wang, J., Liu, X., & Zhou, J. (2019). Development and application of a scale to measure students' STEM continuing motivation. *International Journal of Science Education*, *41*(14), 1885–1904. https://doi.org/10.1080/09500693.2019.1647472

Mahalanobis, P. C. (1936). On the generalised distance in statistics. In *Proceedings of the national Institute of Science of India* (pp. 49-55).

Malespina, A., & Singh, C. (2022). Gender differences in test anxiety and self-efficacy: why instructors should emphasize low-stakes formative assessments in physics courses. *European Journal of Physics*, *43*(3), 035701. https://doi.org/10.1088/1361-6404/ac51b1

Maltese, A. V., & Tai, R. H. (2011). Pipeline persistence: Examining the association of educational experiences with earned degrees in STEM among US students. *Science Education*, *95*(5), 877–907. https://doi.org/10.1002/sce.20441

Maltese, A. V., Melki, C. S., & Wiebke, H. L. (2014). The Nature of Experiences Responsible for the Generation and Maintenance of Interest in STEM. *Science Education, 98*(6), 937-962. https://doi.org/10.1002/sce.21132

Mang, J., Ustjanzew, N., Leßke, I., Schiepe-Tiska, A., & Reiss, K. (2019). *PISA 2015 Skalenhandbuch. Dokumentation der Erhebungsinstrumente [PISA 2015 scale book. Documentation of the measurement instruments].* Waxmann.

Marsh, H. W. (1990). The structure of academic self-concept: The Marsh/Shavelson model. *Journal of Educational Psychology, 82*(4), 623–636. https://doi.org/10.1037/0022-0663.82.4.623

Marshman, E. M., Kalender, Z. Y., Nokes-Malach, T., Schunn, C., & Singh, C. (2018a). Female students with A's have similar physics self-efficacy as male students with C's in introductory courses: A cause for alarm? *Physical Review Physics Education Research, 14*(2). https://doi.org/10.1103/PhysRevPhysEducRes.14.020123

Marshman, E. M., Kalender, Z. Y., Schunn, C., Nokes-Malach, T., & Singh, C. (2018b). A longitudinal analysis of students' motivational characteristics in introductory physics courses: Gender differences. *Canadian Journal of Physics, 96*(4), 391-405. https://doi.org/10.1139/cjp-2017-0185

Martino, W. (2008). Male teachers as role models: Addressing issues of masculinity, pedagogy and the re-masculinization of schooling. *Curriculum inquiry*, *38*(2), 189-223. https://doi.org/10.1111/j.1467-873X.2007.00405.x

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*(2), 149–174. https://doi.org/10.1007/BF02296272

Mesic, V., & Muratovic, H. (2011). Identifying predictors of physics item difficulty: A linear regression approach. *Physical Review Special Topics – Physics Education Research*, *7*(1). https://doi.org/10.1103/PhysRevSTPER.7.010110

Mestre, J. P. (2002). Probing adults' conceptual understanding and transfer of learning via problem posing. *Journal of Applied Developmental Psychology*, *23*(1), 9–50.

Mitchell, M. (1993). Situational interest: Its multifaceted structure in the secondary school mathematics classroom. *Journal of Educational Psychology*, *85*(3), 424–436. https://doi.org/10.1037/0022-0663.85.3.424

Muckenfuß (1995). *Lernen im sinnstiftenden Kontext: Entwurf einer zeitgemäßen Didaktik des Physikunterrichts [Learning in meaningful contexts: proposal of a modern physics education]*. Berlin: Cornelsen.

Muijs, D. (2004). *Doing quantitative research in education with SPSS*. Sage.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics.* Boston College, TIMSS & PIRLS International Study Center. https://timssandpirls.bc.edu/timss2015/international-results/advanced/

Nakamura, J., & Csikszentmihalyi, M. (2009). Flow theory and research. In S. J. Lopez & C. R. Snyder (Eds.), *Oxford handbook of positive psychology* (pp. 195–206). Oxford University Press.

National Research Council (2013). *Next generation science standards: For states, by states*. The National Academies Press. https://doi.org/10.17226/18290

Neumann, I., Neumann, K., & Nehm, R. (2011). Evaluating instrument quality in science education: Rasch-based analyses of a Nature of Science test. *International Journal of Science Education*, *33*(10), 1373–1405. https://doi.org/10.1080/09500693.2010.511297

Neumann, K., Viering, T., Boone, W. J., & Fischer, H. E. (2013). Towards a learning progression of energy. *Journal of Research in Science Teaching*, *50*(2), 162–188. https://doi.org/10.1002/tea.21061

Niu, L. (2020). A review of the application of logistic regression in educational research: common issues, implications, and suggestions. *Educational Review,* *72*(1), 41-67. https://doi.org/10.1080/00131911.2018.1483892

Nuutila, K., Tapola, A., Tuominen, H., Kupiainen, S., Pásztor, A., & Niemivirta, M. (2020). Reciprocal predictions between interest, self-efficacy, and performance during a task. *Frontiers in Education*, *5*(36). https://doi.org/10.3389/feduc.2020.00036

OECD (2005). *Student Questionnaire for PISA 2006*. Available online: https://www.oecd.org/pisa/pisaproducts/PISA06_Student_questionnaire.pdf (accessed on 27 February 2023).

OECD (2006). *Assessing scientific, reading and mathematical literacy: A framework for PISA 2006*. PISA, OECD Publishing.

OECD (2007a). *PISA 2006 Science competencies for tomorrow's world. Volume 1. Analysis*. PISA, OECD Publishing.

OECD (2007b). *PISA 2006 Science competencies for tomorrow's world. Volume 2. Data*. PISA, OECD Publishing.

OECD (2008). *Encouraging student interest in science and technology studies*. Paris: Global Science Forum.

OECD (2009). *PISA 2006 Technical Report*. PISA, OECD Publishing.

OECD (2016). *PISA 2015 results (Volume I): Excellence and equity in education*. PISA, OECD Publishing. https://doi.org/10.1787/9789264266490-en

OECD (2017). *PISA 2015 Assessment and analytical framework: Science, reading, mathematic, financial literacy and collaborative problem solving*. PISA, OECD Publishing. https://doi.org/10.1787/9789264281820-en

Olsen, R. V., & Lie, S. (2011). Profiles of Students' Interest in Science Issues around the World: Analysis of data from PISA 2006. *International Journal of Science Education,* *33*(1), 97–120. https://doi.org/10.1080/09500693.2010.518638

Orion, N., & Hofstein, A. (1994). Factors that influence learning during a scientific field trip in a natural environment. *Journal of Research in Science Teaching,* *31*(10), 1097. https://doi.org/10.1002/tea.3660311005

Osborne, J. & Collins, S. (2001). Pupils' views of the role and value of the science curriculum: a focus-group study. *International Journal of Science Education*, *23*(5), 441–467. https://doi.org/10.1080/09500690320000032199

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: A review of the literature and its implications. *International Journal of Science Education*, *25*(9), 1049–1079. https://doi.org/10.1080/0950069032000032199

Oyserman, D., Elmore, K., & Smith, G. (2012). Self, self-concept, and identity. In J. Tangney and M. Leary (Eds). *The Handbook of Self and Identity* (2nd ed., pp 69-104), New York: Guilford Press.

Parchmann, I., Gräsel, C., Baer, A., Nentwig, P., Demuth, R., & Ralle, B. (2006). "Chemie im Kontext": A symbiotic implementation of a context-based teaching and learning approach. *International Journal of Science Education, 28*(9), 1041–1062. https://doi.org/10.1080/09500690600702512

Parchmann, I., & Kuhn, J. (2018). Lernen im Kontext. In D. Krüger, I. Parchmann, & H. Schecker (Eds.), *Theorien in der naturwissenschaftsdidaktischen Forschung* (pp. 193–207). Springer Spektrum. https://doi.org/10.1007/978-3-662-56320-5_12

Planinic, M., Boone, W. J., Susac, A., & Ivanjek, L. R. (2019). Rasch analysis in physics education research: Why measurement matters. *Physical Review Physics Education Research*, *15*(2), 020111. https://doi.org/10.1103/PhysRevPhysEducRes.15.020111

Preinerstorfer, D., & Formann, A. K. (2012). Measure recovery and model selection in mixed Rasch models. *British Journal of Mathematical and Statistical Psychology, 65*(2), 251-262. https://doi.org/10.1111/j.2044-8317.2011.02020.x

Prenzel, M., Schütte, K., & Walter, O. (2007). Interesse an den Naturwissenschaften [Interest in natural sciences]. In PISA Konsortium Deutschland (Ed.), *Pisa 2006. Die Ergebnisse der Internationalen Vergleichsstudie [PISA 2006. The results of the third international comparison study]* (pp. 107-124). Waxmann.

Quandt, M. (2012). Using the Mixed Rasch Model in the Comparative Analysis of Attitudes. In E. Davidov, P. Schmidt, & J. Billiet (Eds.). *Cross-Cultural Analysis: Methods and Applications* (1st ed., pp. 455-482). Routledge. https://doi.org/10.4324/9780203882924

Radišić, J., Selleri, P., Carugati, F., & Baucal, A. (2021). Are students in Italy really disinterested in science? A person-centred approach using the PISA 2015 data. *Science Education, 105*(2), 438–468. https://doi.org/10.1002/sce.21611

Rask, K. (2010). Attrition in STEM fields at a liberal arts college: The importance of grades and pre-collegiate preferences. *Economics of Education Review, 29*(6), 892–900. https://doi.org/10.1016/j.econedurev.2010.06.013

Renninger, K. A., & Hidi, S. (2011). Revisiting the conceptualization, measurement, and generation of interest. *Educational Psychologist, 46*(3), 168–184. https://doi.org/10.1080/00461520.2011.587723

Rippon, G. (2019). *The Gendered Brain: The New Neuroscience that Shatters the Myth of the Female Brain.* London: The Bodley Head. Print.

Rösler, M., Wellnitz, N., & Mayer, J. (2014). Motivationale Einflüsse auf schriftliche Testleistungen im Fach Biologie [Motivational influences on achievements in written tests in biology]. In D. Krüger, P. Schmiemann, A. Dittmer, & A. Möller (Eds.), *Erkenntnisweg Biologiedidaktik 13* (pp. 179-195). Fachsektion Didaktik der Biologie.

Rösler, M., Wellnitz, N., & Mayer, J. (2018). The role of interesting and motivating contexts in the assessment of content knowledge and decision-making. In N. Gericke & M. Grace (Eds.), *Challenges in biology education research* (pp. 135–150). Karlstad University Printing Office.

Rost, J., & von Davier, M. (1994). A Conditional Item-Fit Index for Rasch Models. *Applied Psychological Measurement, 18*(2), 171–182. https://doi.org/10.1177/014662169401800206

Rost, J., & von Davier, M. (1995). Mixture Distribution Rasch Models. In G. H. Fischer & I. W. Molenaar (Eds.), *Rasch Models: Foundations, Recent Developments, and Applications* (pp. 257-268). New York, NY: Springer.

Rost, J., Sievers, K., Häußler, P., Hoffmann, L., & Langeheine, R. (1999). Struktur und Veränderung des Interesses an Physik bei Schülern der 6. bis 10. Klassenstufe [Structure and change of interest in physics among grade 6 to grade 10 students]. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *31*(1), 18–31. https://doi.org/10.1026//0049-8637.31.1.18

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist, 55*(1), 68.

Salchegger, S., Wallner-Paschon, C., Schmich, J., & Höller, I. (2016). Kompetenzentwicklung im Kontext individueller, schulischer und familiärer Faktoren [Development of competences in the context of individual, school and family factors]. In B. Suchań, & S. Breit (Eds.), *PISA 2015. Grundkompetenzen am Ende der Pflichtschulzeit im internationalen Vergleich [PISA 2015. International comparison of the basic competences at the end of obligatory schooling].* Graz: Leykam

Sander, D. (2013). Models of Emotion. In J. Armony & P. Vuilleumier (Eds.), *The Cambridge Handbook of Human Affective Neuroscience* (pp. 5-54). Cambridge: Cambridge University Press.

Sandmann, A. (2014). Lautes Denken–die Analyse von Denk-, Lern-und Problemlöseprozessen [Thinking aloud—The analysis of thinking, learning, and problem-solving processes]. In D. Krüger, I. Parchmann & H. Schecker (Eds.), *Methoden in der naturwissenschaftsdidaktischen Forschung* (pp. 179–188). Springer Spektrum. https://doi.org/10.1007/978-3-642-37827-0_15

Schiefele, U., Krapp, A., Wild, K.-P., & Winteler, A. (1993). Der „Fragebogen zum Studieninteresse“ (FSI) [The study interest questionnaire (SIQ)]. *Diagnostica, 39*(4), 335–351.

Schreiner, C., & Sjøberg, S. (2004). *Sowing the seeds of ROSE: Background, rationale, questionnaire development and data collection for ROSE (The Relevance of Science Education): A comparative study of students' views of science and science education*. Oslo: ILS og forfatterne.

Schriebl, D., Müller, A. & Robin, N. (2022). Modelling Authenticity in Science Education. *Science & Education*. https://doi.org/10.1007/s11191-022-00355-x

Schwarz, N. (1999). Self-Reports: How the Questions Shape the Answers. *American Psychologist, 54*(2), 93-105.

Sen, S. (2018). Spurious Latent Class Problem in the Mixed Rasch Model: A Comparison of Three Maximum Likelihood Estimation Methods under Different Ability Distributions. *International Journal of Testing, 18*(1), 71-100. https://doi.org/10.1080/15305058.2017.1312408

Seymour, E., & Hewitt, N. (1997). *Talking about leaving: Why undergraduates leave the sciences*. Boulder, CO: Westview.

Shapiro, C. A., & Sax, L. J. (2011). Major selection and persistence for women in STEM. *New Directions for Institutional Research, 2011*(152), 5-18. https://doi.org/10.1002/ir.404

Shavelson, R. J., Hubner, J. J., & Stanton, G. C. (1976). Self-Concept: Validation of Construct Interpretations. *Review of Educational Research, 46*(3), 407-441. https://doi.org/10.3102/00346543046003407

Sheu, H.-B., Lent, R. W., Brown, S. D., Miller, M. J., Hennessy, K. D., & Duffy, R. D. (2010). Testing the choice model of social cognitive career theory across Holland themes: A meta-analytic path analysis. *Journal of Vocational Behavior, 76*(2), 252-264. https://doi.org/10.1016/j.jvb.2009.10.015

Sievers, K. (1999). *Struktur und Veränderung von Physikinteressen bei Jugendlichen [Structure and change of physics interest among adolescents]*. Universität Kiel.

Silvia, P. J. (2001). Interest and interests: The psychology of constructive capriciousness. *Review of General Psychology, 5*(3), 270–290. https://doi.org/10.1037/1089-2680.5.3.270

Silvia, P. J. (2008). Interest—The Curious Emotion. *Current Directions in Psychological Science, 17*(1), 57-60. https://doi.org/10.1111/j.1467-8721.2008.00548.x

Sjøberg, S., & Schreiner, C. (2010). *The ROSE project: An overview and key findings*. Available online: http://roseproject.no/network/countries/norway/eng/nor-Sjoberg-Schreiner-overview-2010.pdf (accessed on 26 February 2023).

Sjøberg, S., & Schreiner, C. (2012). Results and perspectives from the ROSE Project: Attitudinal aspects of young people and science in a comparative perspective. In D. Jorde & J. Dillon (Eds.), *Science education research and practice in Europe. Cultural perpectives in science education, 5* (pp. 203–236). Sense Publishers. https://doi.org/10.1007/978-94-6091-900-8_9

Speering, W., & Rennie, L. (1996). Students' perceptions about science: The impact of transition from primary to secondary school. *Research in Science Education, 26*(3), 283-298. https://doi.org/10.1007/BF02356940

Stern, T., Jelemenská, P., & Radits, F. (2009). Das Interesse an Naturwissenschaften: Eine Analyse der österreichischen PISA-2006-Ergebnisse. In C. Schreiner & U. Schwantner (Eds.), *PISA 2006: Österreichischer Expertenbericht zum Naturwissenschafts-Schwerpunkt*. Graz: Leykam.

Swarat, S., Ortony, A., & Revelle, W. (2012). Activity matters: Understanding student interest in school science. *Journal of Research in Science Teaching, 49*(4), 515-537.

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.

Todt, E. (1978). *Das Interesse Empirische Untersuchungen zu einem Motivationskonzept.* Huber.

Turner, S. L., Steward, J. C., & Lapan, R. T. (2004). Family Factors Associated With Sixth-Grade Adolescents' Math and Science Career Interests. *The Career Development Quarterly, 53*(1), 41-52. https://doi.org/10.1002/j.2161-0045.2004.tb00654.x

Tyson, W., Lee, R., Borman, K. M., & Hanson, M. A. (2007). Science, technology, engineering, and mathematics (STEM) pathways: High school science and math coursework and postsecondary degree attainment. *Journal of Education for Students Placed at Risk, 12*(3), 243–270. https://doi.org/10.1080/10824660701601266

van Vorst, H., Dorschu, A., Fechner, S., Kauertz, A., Krabbe, H., & Sumfleth, E. (2014). Charakterisierung und Strukturierung von Kontexten im naturwissenschaftlichen Unterricht – Vorschlag einer theoretischen Modellierung. *Zeitschrift für Didaktik der Naturwissenschaften, 21*(1), 29–39. https://doi.org/10.1007/s40573-014-0021-5

van Vorst, H., Fechner, S., & Sumfleth, E. (2018). Unterscheidung von Kontexten für den Chemieunterricht. *Zeitschrift für Didaktik der Naturwissenschaften, 24*(1), 167-181. https://doi.org/10.1007/s40573-018-0081-z

Vogt, H. (2007). Theorie des Interesses und des Nicht-Interesses [Theory of interest and non-interest]. In D. Krüger & H. Vogt (Eds.), *Theorien in der biologiedidaktischen Forschung: Ein Handbuch für Lehramtsstudenten und Doktoranden [Theories in biology education research: a handbooks for teacher and doctoral students]* (pp. 9–20). Berlin, Heidelberg: Springer Berlin Heidelberg.

Vogt, P. (2010). *"Werbeaufgaben" in Physik: Motivations- und Lernwirksamkeit authentischer Texte, untersucht am Beispiel von Werbeanzeigen*. Wiesbaden: Vieweg+Teubner Verlag.

von Davier, M. (2001). *WINMIRA 2001 user manual*. Available online: http://208.76.80.46/~svfklumu/wmira/winmiramanual.pdf (accessed on 3 March 2023).

Vorholzer, A., von Aufschnaiter, C., & Kirschner, S. (2016). Entwicklung und Erprobung eines Tests zur Erfassung des Verständnisses experimenteller Denk- und Arbeitsweisen [Development and evaluation of an instrument to measure the understanding of experimental ways of thinking and working]. *Zeitschrift für Didaktik der Naturwissenschaften*, *22*(1), 25–41. https://doi.org/10.1007/s40573-015-0039-3

Voss, R., Rudolf, P., Saunders, F., & Lee, D. (2019). *The Importance of Physics to the Economies of Europe*.

Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika, 54*(3), 427-450. https://doi.org/10.1007/BF02294627

West, C., & Zimmerman, D. H. (1987). Doing Gender. *Gender & Society, 1*(2), 125-151. https://doi.org/10.1177/0891243287001002002

Wetzel, E., Böhnke, J. R., Carstensen, C. H., Ziegler, M., & Ostendorf, F. (2013). Do individual response styles matter? Assessing differential item functioning for men and women in the NEO-PI-R. *Journal of Individual Differences, 34*(2), 69–81. https://doi.org/10.1027/1614-0001/a000102

Wilkinson, L., & American Psychological Association Task Force on Statistical Inference (1999). Statistical methods in psychology journals: Guidelines and explanations. *American Psychologist, 54*(8), 594–604. https://doi.org/10.1037/0003-066X.54.8.594

Wigfield, A., & Eccles, J. S. (2000). Expectancy-value theory of achievement motivation. *Contemporary Educational Psychology, 25*(1), 68–81. https://doi.org/10.1006/ceps.1999.1015

Woithe, J. (2020). *Designing, measuring and modelling the impact of the hands-on particle physics learning laboratory S'Cool LAB at CERN Effects of student and laboratory characteristics on high-school students' cognitive and affective outcomes* (Doctoral dissertation, Technical University of Kaiserslautern). Available online: https://cds.cern.ch/record/2727453 (accessed on 28 February 2023).

Woolnough, B. E., Guo, Y., Leite, M. S., Almeida, M. J. d., Ryu, T., Wang, Z., & Young, D. (1997). Factors Affecting Student Choice of Career in Science and Engineering: parallel studies in Australia, Canada, China, England, Japan and Portugal. *Research in Science & Technological Education, 15*(1), 105-121. https://doi.org/10.1080/0263514970150108

Wright, B., & Stone, M. (1979). *Best test design.* MESA Press.

Wyer, M. (2003). Intending to stay: Images of scientists, attitudes toward women, and gender as influences on persistence among science and engineering majors. *Journal of Women and Minorities in Science and Engineering, 9*(1), 1 – 16. https://doi.org/10.1615/JWomenMinorScienEng.v9.i1.10

Yao, J.-X., Guo, Y.-Y., & Neumann, K. (2017). Refining a learning progression of energy. *International Journal of Science Education*, *39*(17), 2361–2381. https://doi.org/10.1080/09500693.2017.1381356

Young, D. M., Rudman, L. A., Buettner, H. M., & McLean, M. C. (2013). The Influence of Female Role Models on Women's Implicit Science Cognitions. *Psychology of Women Quarterly, 37*(3), 283–292. https://doi.org/10.1177/0361684313482109

Zeldin, A. L., & Pajares, F. (2000). Against the Odds: Self-Efficacy Beliefs of Women in Mathematical, Scientific, and Technological Careers. *American Educational Research Journal, 37*(1), 215-246. https://doi.org/10.3102/00028312037001215

Zeldin, A. L., Britner, S. L., & Pajares, F. (2008). A comparative study of the self-efficacy beliefs of successful men and women in mathematics, science, and technology careers. *Journal of Research in Science Teaching, 45*(9), 1036. https://doi.org/10.1002/tea.20195

Zoechling, S., Hopf, M., Woithe, J., & Schmeling, S. (2022). Students' interest in particle physics: conceptualisation, instrument development, and evaluation using Rasch theory and analysis.

International Journal of Science Education, 44(15), 2353-2380, https://doi.org/10.1080/09500693.2022.2122897

Zoller, U., & Watson, F. G. (1974). Teacher Training for the "Second Generation" of Science Curricula: The Curriculum-Proof Teacher. *Science Education, 58*(1), 93-103.

9. DATA AVAILABILITY STATEMENT

The qualitative data (i.e. transcripts of think-aloud interviews) that support the findings of these studies are available on request. The quantitative data that support the findings of the studies presented in this thesis are openly available in the Austrian Social Science Data Archive at https://doi.org/10.11587/OUDFJK (Developing the IPPI, **Chapter 5**) and at https://doi.org/10.11587/PHPAUN (Main study, **Chapter 6**).

10. APPENDICES

10.1. Appendix 1: Developping the IPPI

10.1.1. Introductory text on particle physics

a) Original German Text:

TEILCHENPHYSIK oder *Woraus wir eigentlich bestehen*

Alles, was man zumindest theoretisch berühren kann, wird als Materie bezeichnet. Dazu zählen nicht nur wir Menschen, sondern auch Sterne und Planeten. Teilchenphysiker*innen erforschen, woraus alle Materie besteht und was ihre Bestandteile zusammenhält. Ein menschliches Haar zum Beispiel ist aus Atomen aufgebaut, und ein Atom aus einem Atomkern-Bereich und Elektronen, die diesen umgeben. Das Elektron ist ein sogenanntes *Elementarteilchen*. Diese sind unteilbar.

Erkenntnisse über den Aufbau der Materie gewinnt man mithilfe von Experimenten. Zum Beispiel beschleunigt man Teilchen auf sehr hohe Geschwindigkeiten, um sie dann zusammenstoßen zu lassen. Bei diesem Zusammenstoß entstehen neue Teilchen, die von Detektoren aufgezeichnet werden. Diese sind mehrere Stockwerke hohe Geräte, die bis zu 40 Millionen Aufzeichnungen pro Sekunde machen können. Teilchenphysiker*innen werten diese Aufzeichnungen aus. So können sie zum Beispiel die Prozesse erforschen, die sehr kurz nach dem Urknall stattgefunden haben, um besser zu verstehen, wie unser Universum entstanden ist.

Im Grunde können wir alle physikalischen Phänomene auf die Wechselwirkungen zwischen Teilchen zurückführen, wenn wir ganz genau hinschauen. Zum Beispiel wechselwirken Elektronen miteinander, weil sie eine elektrische Ladung haben: Diese Wechselwirkung verhindert, dass du durch den Stuhl fällst, auf dem du vermutlich gerade sitzt. Denn die Elektronen deines Körpers und die Elektronen des Stuhls stoßen sich gegenseitig ab und können nicht beliebig nah zusammengebracht werden.

Außerdem hat Forschung in der Teilchenphysik viele Anwendungen, zum Beispiel bei der Diagnose und Behandlung von Krankheiten oder bei der Feststellung der Echtheit eines Kunstwerks.

Wie gerne würdest du im Zusammenhang mit diesem Thema das Folgende tun? Bitte klicke „Weiter“!

b) English Translation of the Text:

PARTICLE PHYSICS or *What we are actually made of*

Everything that can be touched, at least in theory, is called matter. This includes not only us humans, but also the stars and the planets. Particle physicists are investigating what all matter is made of and what holds its components together. A human hair, for example, is made of atoms, and an atom is made of a nucleus space and electrons surrounding it. The electron is a so-called *elementary particle*. These are indivisible.

We gain knowledge about the structure of matter through experiments. For example, particles are accelerated to very high speeds and then forced to collide. This collision creates new particles that are recorded by detectors. These are devices that are several storeys high and can make up to 40 million recordings per second. Particle physicists analyse these recordings. This enables them, for example, to investigate the processes that took place very shortly after the Big Bang to better understand how our universe came into being.

We can explain all physical phenomena in terms of the interactions between particles if we look very closely. For example, electrons interact with each other because they have an electric charge: This interaction prevents you from falling through the chair on which you are probably sitting on right now. This is because the electrons of your body and the electrons of the chair repel each other and cannot be brought infinitely close together.

In addition, research in particle physics has many applications, for example in the diagnosis and treatment of diseases or in determining the authenticity of a piece of art.

In relation to this topic, how interested are you in doing the following? Please click "Next"!

10.1.2. Items to measure particle physics interest

Table 35. Original German wordings and English paraphrases of the items to measure particle physics interest.

Item ID(s)	Original German wording	English paraphrase
I011	Mehr darüber erfahren, wie ein Teilchenbeschleuniger funktioniert	Particle accelerator
I012 (PP01)	Mehr darüber erfahren, wie Geräte funktionieren, die Teilchen detektieren (z.B. Digitalkamera)	Devices that detect particles (e.g. digital camera)
I013	Mehr darüber erfahren, wie Geräte funktionieren, die Teilchen beschleunigen (z.B. Elektronenmikroskop)	Devices that accelerate particles (e.g. electron microscope)
I021 (PP02)	Mehr darüber erfahren, wie Teilchenphysik zum Verständnis des Urknalls beiträgt	Particle physics and the big bang
I022	Mehr darüber erfahren, wie Teilchenphysik zum Verständnis von Polarlichtern beiträgt	Particle physics and the northern lights
I023	Mehr darüber erfahren, welche Elementarteilchen aus dem Kosmos bis zur Erdoberfläche gelangen	Cosmic particles
I031	Mehr darüber erfahren, wie ein Teilchenbeschleuniger zur friedlichen Zusammenarbeit verschiedener Nationen beiträgt	A particle accelerator and the peaceful collaboration of diverse nations
I032 (PP03)	Mehr darüber erfahren, wie mithilfe von Teilchendetektoren geschmuggelte Waffen in einem Container entdeckt werden können	Particle detectors and smuggled arms
I033	Mehr darüber erfahren, wie mithilfe der Teilchenphysik festgestellt werden kann, ob ein Kunstwerk echt ist	Particle physics and art authentication
I041	Mehr darüber erfahren, welche Elementarteilchen und Wechselwirkungen es gibt	Particles and interactions
I042 (PP04)	Mehr darüber erfahren, welche Elementarteilchen man im Atomkern-Bereich findet	Particles in the nucleus space of an atom
I043	Mehr darüber erfahren, welche Wechselwirkung die Elementarteilchen im Atomkern-Bereich zusammenhält	The interaction binding together the nucleus space of an atom
I051	Mehr darüber erfahren, aus wie vielen Elementarteilchen ein Gegenstand, z.B. ein Stift, besteht	Particles of objects (e.g. pen) (quantitative)
I052	Mehr darüber erfahren, warum man Teilchen zu Forschungszwecken auf sehr hohe Geschwindigkeiten beschleunigen muss	Acceleration of particles (quantitative)
I053 (PP05)	Mehr darüber erfahren, wie groß die Massen der Elementarteilchen im Vergleich zueinander sind	Masses of particles (quantitative)

cont.

I061	Die Vielfalt der verschiedenen Berufsgruppen, die in der Teilchenphysik mitarbeiten, kennenlernen	Occupational groups contributing to particle physics
I062	Mehr Einblick erhalten, in welchen Bereichen - abgesehen von Forschung - Teilchenphysiker*innen arbeiten	Jobs outside science for particle physicists
I063 (PP06)	Mehr Einblick erhalten, wie in der Elektronik-Industrie mit Teilchenbeschleunigern gearbeitet wird	Particle accelerators in the electronics industry
I071 (PP07)	Mehr Einblick erhalten, wie in einem medizinischen Diagnose-Zentrum gearbeitet wird	Medical diagnostics
I072	Mehr Einblick erhalten, wie Krankheiten mithilfe von Teilchenbeschleunigern behandelt werden	Particle accelerators to cure diseases
I073	Mehr Einblick erhalten, wie man das Innere von Vulkanen oder Pyramiden mithilfe von Teilchendetektoren erkennen kann	Particle accelerators for studying volcanoes or pyramids
I081	Einen Teilchendetektor aus Alltagsgegenständen selbst bauen und ausprobieren	Building a particle detector out of daily life objects (hands-on)
I082	Einen Elektromagneten bauen und damit die Bewegungsrichtung eines Teilchens verändern	Building an electromagnet to influence the direction of a particle (hands-on)
I083 (PP08)	Ein Handy in einen Teilchendetektor umbauen und ausprobieren	Transforming a mobile phone into a particle detector (hands-on)
I091 (PP09)	Ein Experiment planen, um zu zeigen, wie Teilchen beschleunigt werden	Planning an experiment on particle acceleration (minds-on)
I092	Sich ein Experiment ausdenken, um zu zeigen, wie man den Aufbau eines Atoms erforschen kann	Planning an experiment to study the structure of an atom (minds-on)
I093	Ein Experiment planen, um zu zeigen, wie man die Bewegungsrichtung eines Teilchens verändern kann	Planning an experiment to influence the direction of a particle (minds-on)
I101 (PP10)	Berechnen, wie groß die Energie beim Zusammenstoß zweier Teilchen ist, die sich mit nahezu Lichtgeschwindigkeit bewegen	Calculating the energy of a particle collision at the speed of light (minds-on)
I102	Berechnen, aus wie vielen Elementarteilchen ein menschliches Haar besteht	Calculating number of particles in human hair (minds-on)
I103	Die Massen verschiedener Elementarteilchen berechnen, weil man sie nicht einfach abwiegen kann	Calculating the mass of particles (minds-on)

cont.

I111 (PP11)	Darüber diskutieren, wie Erkenntnisse im Bereich der Teilchenphysik unser Alltagsleben verändert haben	Particle physics has changed our daily life (discussion)
I112	Darüber diskutieren, warum Forschung in der Teilchenphysik für unsere Gesellschaft wichtig ist	The societal relevance of particle physics (discussion)
I113	Darüber diskutieren, warum die EU in den letzten fünf Jahren 10 Mio. € in die Entwicklung von Teilchendetektoren investiert hat	EU investments in particle detectors (discussion)

10.2. Appendix 2: Main study

10.2.1. Original IPN introductory text on mechanics

a) Original German Text:

VON DEN BEWEGUNGEN und *Wie man Kraft sparen kann*

Von schwerer körperlicher Anstrengung befreit zu sein, ist ein uralter Menschheitstraum, und schon die alten Griechen waren Meister darin, Kraft zu sparen. So kannten sie schon den Flaschenzug, eine sinnvolle Konstruktion aus Seilen und Rollen, mit der sie zum Beispiel schwere Tempelsäulen aufrichten konnten.

Heute gibt es viele Geräte, die uns Anstrengungen abnehmen: Eine hydraulische Hebebühne hebt zum Beispiel Autos in Autowerkstätten, eine Pumpe holt Erdöl aus großer Tiefe an die Erdoberfläche, und ein leichtes Tippen auf das Gaspedal beschleunigt den Sportwagen auf 180 Sachen (Kilometer pro Stunde)!

Aber macht man sich immer klar, welche Kräfte man da entfesselt hat? Wer Physik gelernt hat, der kann leicht ausrechnen, dass ein frontaler Aufprall eines Autos auf eine Mauer bei einer Geschwindigkeit von nur 50 Kilometer pro Stunde genauso gefährlich ist wie ein Sturz aus 10 Meter Höhe!

Weil die meisten Unfälle bei verhältnismäßig geringen Geschwindigkeiten und im Stadtverkehr passieren, sehen viele einen Ausweg darin, nach physikalischen Erkenntnissen konstruierte Sicherheitsautos zu bauen. Dazu gehören zum Beispiel Sicherheitsgurte, die verhindern, dass man bei einem Aufprall durch die Frontscheibe fliegt, oder sogenannte Knautschzonen im vorderen und hinteren Bereich des Wagens, die die Bewegungsenergie durch Verformung von Blechteilen aufzehren sollen und so die Wucht des Aufpralls bei einem Unfall mindern.

Wie gerne würdest du im Zusammenhang mit diesem Thema das Folgende tun? Bitte klicke auf "Weiter"!

b) English Translation of the Text:

ABOUT MOVEMENTS and ***How to save effort***

To be freed from heavy physical effort is an ancient dream of humankind, and even the ancient Greeks were masters at saving effort. They already knew the pulley, a useful construction made of ropes and pulleys, with which they could lift heavy temple columns, for example.

Today, there are many devices that save us effort: A hydraulic lift, for example, lifts cars in garages, a pump brings oil from great depths to the earth's surface, and a light tap on the throttle accelerates a sports car to 180 kilometres per hour!

But do we always realise what forces we have unleashed? Anyone who has studied physics can easily calculate that a head-on collision of a car with a wall at a speed of only 50 kilometres per hour is as dangerous as a fall from a height of 10 metres!

Because most accidents happen at relatively low speeds and in city traffic, many consider constructing safety cars based on physics to be a way out. These include, for example, seat belts that prevent you from flying through the windscreen in the case of a collision, or so-called crumple zones in the front and rear of the car, which are designed to absorb kinetic energy by deforming sheet metal parts and thus reduce the force of the collision in the event of an accident.

In relation to this topic, how interested are you in doing the following? Please click on "Next"!

10.2.2. Newly developed introductory text on mechanics

a) Original German Text:

MECHANIK oder *Was Kräfte bewirken*

Nicht alles, was wir im Alltag als Kraft bezeichnen, ist auch im physikalischen Sinne eine Kraft. So hast du bestmmt schon mal gehört, dass etwas «Kraft hat». Doch Kraft kann man nicht haben, man kann sie nur ausüben. Oft erkennen wir eine Kraft an ihrer Wirkung, zum Beispiel etwas in Bewegung zu versetzen oder zu verformen.

Manchmal sind wir Menschen nicht stark genug, um eine genügend große Kraft auszuüben. Manchmal wollen wir aber einfach von schwerer körperlicher Anstrengung befreit sein. Zum Glück wusste man schon in der Antike, wie man Kräfte verstärken kann. So erfand man schon vor 3000 Jahren den Flaschenzug, eine nützliche Konstruktion aus Seilen und Rollen, mit der man zum Beispiel schwere Tempelsäulen aufrichten konnte. Heute gibt es viele Geräte, die uns Anstrengungen abnehmen: Pumpen holen Grundwasser aus großer Tiefe an die Erdoberfläche. Hebebühnen heben Elefanten in Tierkliniken, Patient*innen in Krankenhäusern oder Autos in Autowerkstätten. Außerdem genügt ein leichtes Tippen auf das Gaspedal, um ein Fahrzeug auf 130 km/h zu beschleunigen!

Aber macht man sich immer klar, welche Energien man dabei erreicht? Ein frontaler Aufprall eines Autos auf eine Mauer bei einer Geschwindigkeit von nur 50 km/h ist zum Beispiel genauso gefährlich wie ein Sturz aus 10 Metern Höhe!

Um Fahrzeuge sicherer zu machen, werden Erkenntnisse der Mechanik genutzt. Zum Beispiel verhindert der Sicherheitsgurt, dass du bei einem Aufprall nach vorne fliegst. Außerdem gibt es im vorderen und hinteren Bereich jedes Fahrzeugs sogenannte Knautschzonen, in denen die Bewegungsenergie durch Verformung von Blechteilen aufgenommen wird. So wird die Wucht des Aufpralls bei einem Unfall verringert.

Wie gerne würdest du im Zusammenhang mit diesem Thema das Folgende tun? Bitte klicke auf "Weiter"!

b) English Translation of the Text:

MECHANICS or *What forces do*

Not everything that we call a force in everyday life is actually a force in the physical sense. You have probably heard that someone "is a force". But you can't actually be a force, you can only exert forces. We often recognise a force by its effect, for example, to set something in motion or to deform it.

Sometimes we humans are not strong enough to exert sufficient force. Sometimes, however, we simply want to be free from heavy physical effort. Fortunately, people already knew how to amplify strength in ancient times. As early as 3000 years ago, people invented the pulley, a useful construction made of ropes and pulleys that could be used, for example, to put up heavy temple columns. Today there are many devices that relieve us of our efforts: Pumps bring groundwater from great depths to the earth's surface. Lifting platforms lift elephants in veterinary clinics, patients in hospitals or cars in garages. Moreover, a light tap on the throttle is enough to accelerate a vehicle to 130 km/h!

But do we always realise what energies we are reaching? For example, a head-on collision of a car with a wall at a speed of only 50 km/h is as dangerous as a fall from a height of 10 metres!

To make vehicles safer, knowledge of mechanics is used. For example, the seat belt prevents you from falling forward in case of a collision. In addition, there are so-called crumple zones at the front and rear of every vehicle, in which the kinetic energy is absorbed by deforming sheet metal parts. This reduces the force of the collision in an accident.

In relation to this topic, how interested are you in doing the following? Please click on "Next"!

10.2.3. Items to measure mechanics interest

Table 36. Original German wordings and English paraphrases of the item to measure mechanics interest.

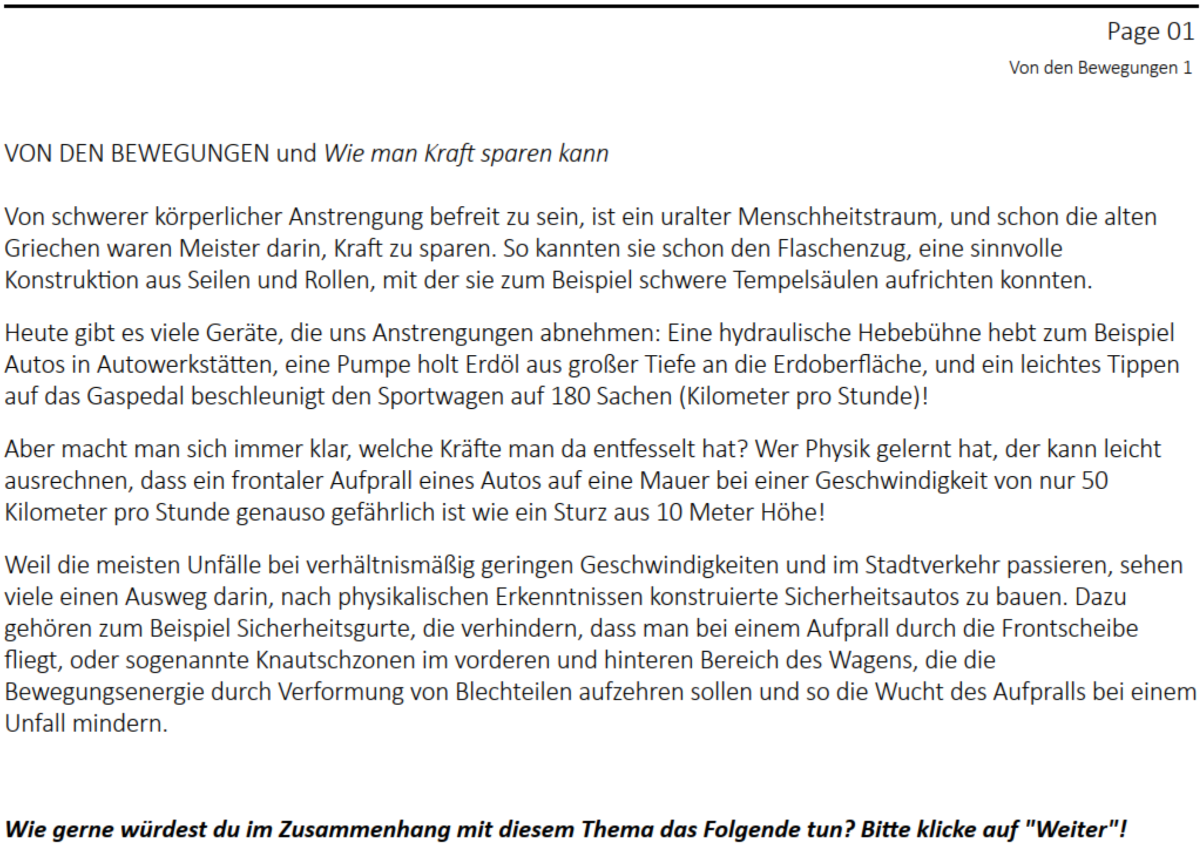
Item ID	Original German wording	English paraphrase
M01	Mehr darüber erfahren, wie Geräte funktionieren, die Kräfte verstärken (z.B. Flaschenzug, Hebebühne)	Devices that amplify forces (e.g. pulley, lifting platform)
M02	Mehr darüber erfahren, wie man Erdöl aus sehr großen Tiefen (z.B. aus 3000 Metern) heraufpumpen kann	Pumping oil from great depths
M03	Mehr darüber erfahren, wie die Wahrscheinlichkeit eines Autounfalls und die Schwere der Unfallfolgen mit zunehmender Geschwindigkeit wachsen	Increase of probability and consequences of a car accident with speed
M04	Mehr darüber erfahren, wie die Bewegungsenergie (Wucht) eines Fahrzeugs in andere Energieformen umgewandelt werden kann (z.B. in den Bremsen oder in der Knautschzone)	Transformation of energy of motion into other forms of energy
M05	Mehr Einblick erhalten, wie die Bewegungsenergie eines Fahrzeugs aus seiner Geschwindigkeit berechnet werden kann	Calculating the energy of motion based on the speed of a car (quantitative)
M06	Mehr Einblick erhalten, welche kraftsparenden Geräte in einer Autowerkstatt verwendet werden	Force-saving devices in a car repair shop
M07	Mehr Einblick erhalten, welche künstlichen Organe (z.B. Herz als Blutpumpe) und Gelenke heute in der Medizin zur Verfügung stehen	Artificial organs and joints in medicine
M08	Mit Rollen und Seilen verschiedene Flaschenzüge bauen und ausprobieren	Constructing different pulleys and trying them out (hands-on)
M09	Sich ein Sicherheitsfahrzeug ausdenken, in dem auch bei schweren Unfällen Fahrer und Beifahrer wenig oder nichts passiert	Designing a safety car for severe accidents (minds-on)
M10	Darüber nachdenken, wie man aus dem Bremsweg eines Autos seine Geschwindigkeit vor dem Abbremsen berechnen kann	Calculating the braking path based on the speed of a car
M11	Sich mit Unfallstatistiken beschäftigen und über den Sinn von Geschwindigkeitsbegrenzungen diskutieren	Accident statistics and speed limits (discussion)

10.2.4. Items to measure physics-related self-concept

Table 37. Original German and English wordings of the items to measure physics-related self-concept.

Item ID	Original German wording	English wording	Aspect
SC01	Ich glaube, dass ich anspruchsvollen Stoff im Physikunterricht leicht lernen kann.	Learning advanced physics topics would be easy for me	*Physics ability self-concept*
SC02	Normalerweise kann ich Prüfungsfragen im Physikunterricht gut beantworten.	I can usually give good answers to test questions on physics topics	*Physics ability self-concept*
SC03	Mein*e Lehrer*in sieht mich als Physik-Person.	My TA or Instructor see me as physics person	*Physics perceived recognition*
SC04	Ich lerne neuen Stoff im Physikunterricht schnell.	I can learn physics topics quickly	*Physics ability self-concept*
SC05	Meine Eltern sehen mich als Physik-Person.	My parents see me as physics person	*Physics perceived recognition*
SC06	Den Stoff im Physikunterricht finde ich einfach.	Physics topics are easy for me	*Physics ability self-concept*
SC07	Meine Mitschüler*innen sehen mich als Physik-Person.	My classmates see me as physics person	*Physics perceived recognition*
SC08	Wenn ich in Physik unterrichtet werde, verstehe ich neue Begriffe leicht.	When I am being taught physics, I can understand the concepts very well	*Physics ability self-concept*
SC09	Meine Freund*innen sehen mich als Physik-Person.	My friends see me as physics person	*Physics perceived recognition*
SC10	Es fällt mir leicht, neue Ideen im Physikunterricht zu verstehen.	I can easily understand new ideas in physics	*Physics ability self-concept*

10.2.5. Version 1 of the questionnaire with original IPN mechanics text



Page 01

Von den Bewegungen 1

VON DEN BEWEGUNGEN und *Wie man Kraft sparen kann*

Von schwerer körperlicher Anstrengung befreit zu sein, ist ein uralter Menschheitstraum, und schon die alten Griechen waren Meister darin, Kraft zu sparen. So kannten sie schon den Flaschenzug, eine sinnvolle Konstruktion aus Seilen und Rollen, mit der sie zum Beispiel schwere Tempelsäulen aufrichten konnten.

Heute gibt es viele Geräte, die uns Anstrengungen abnehmen: Eine hydraulische Hebebühne hebt zum Beispiel Autos in Autowerkstätten, eine Pumpe holt Erdöl aus großer Tiefe an die Erdoberfläche, und ein leichtes Tippen auf das Gaspedal beschleunigt den Sportwagen auf 180 Sachen (Kilometer pro Stunde)!

Aber macht man sich immer klar, welche Kräfte man da entfesselt hat? Wer Physik gelernt hat, der kann leicht ausrechnen, dass ein frontaler Aufprall eines Autos auf eine Mauer bei einer Geschwindigkeit von nur 50 Kilometer pro Stunde genauso gefährlich ist wie ein Sturz aus 10 Meter Höhe!

Weil die meisten Unfälle bei verhältnismäßig geringen Geschwindigkeiten und im Stadtverkehr passieren, sehen viele einen Ausweg darin, nach physikalischen Erkenntnissen konstruierte Sicherheitsautos zu bauen. Dazu gehören zum Beispiel Sicherheitsgurte, die verhindern, dass man bei einem Aufprall durch die Frontscheibe fliegt, oder sogenannte Knautschzonen im vorderen und hinteren Bereich des Wagens, die die Bewegungsenergie durch Verformung von Blechteilen aufzehren sollen und so die Wucht des Aufpralls bei einem Unfall mindern.

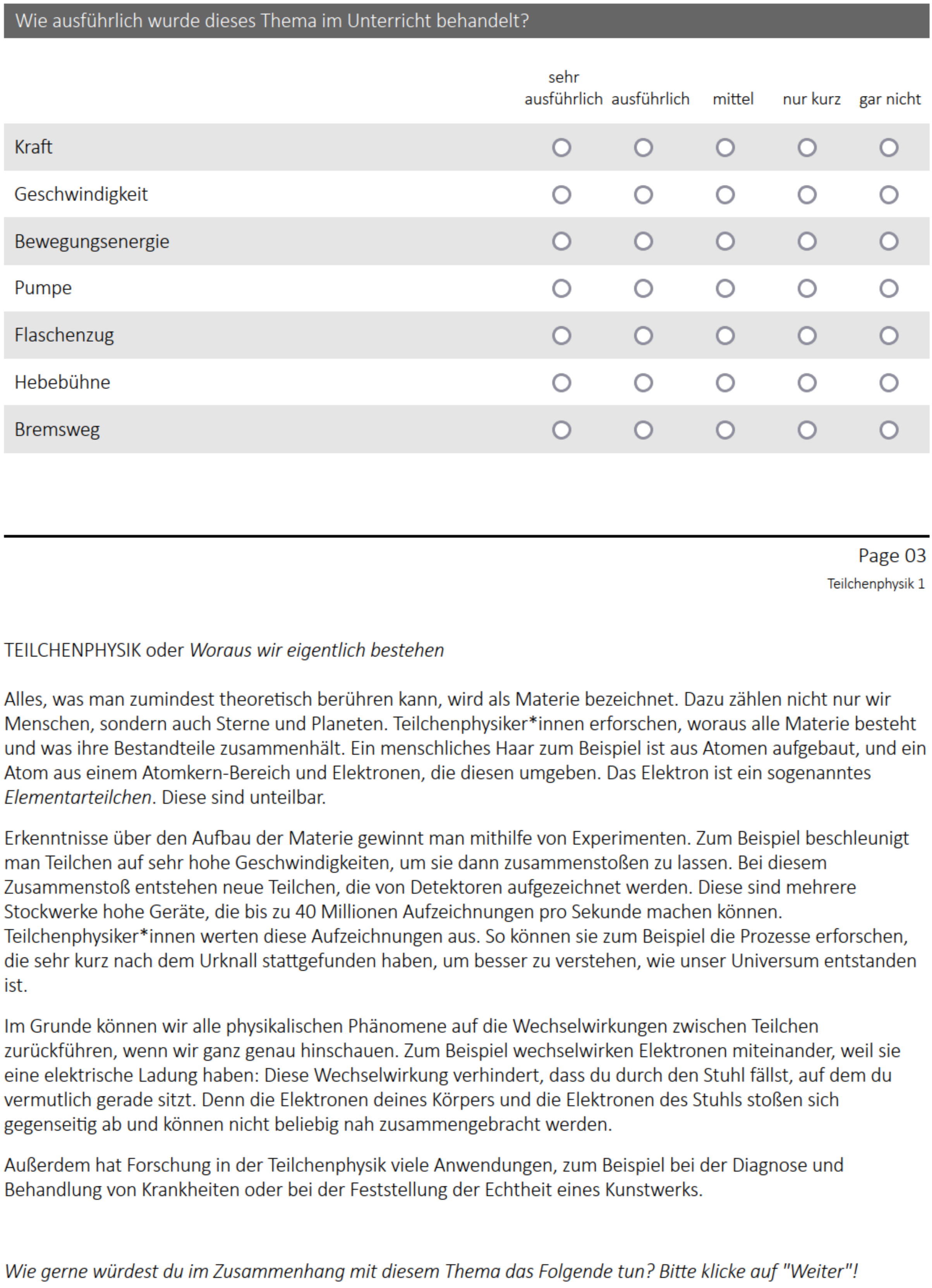
Wie gerne würdest du im Zusammenhang mit diesem Thema das Folgende tun? Bitte klicke auf "Weiter"!

Page 02

Von den Bewegungen 2

Mein Interesse daran ist

	sehr groß	groß	mittel	gering	sehr gering
Mehr darüber erfahren, wie Geräte funktionieren, die Kräfte verstärken (z.B. Flaschenzug, Hebebühne)	○	○	○	○	○
Mehr darüber erfahren, wie man Erdöl aus sehr großen Tiefen (z.B. aus 3000 Metern) heraufpumpen kann	○	○	○	○	○
Mehr darüber erfahren, wie die Wahrscheinlichkeit eines Autounfalls und die Schwere der Unfallfolgen mit zunehmender Geschwindigkeit wachsen	○	○	○	○	○
Mehr darüber erfahren, wie die Bewegungsenergie (Wucht) eines Fahrzeugs in andere Energieformen umgewandelt werden kann (z.B. in den Bremsen oder in der Knautschzone)	○	○	○	○	○
Mehr Einblick erhalten, wie die Bewegungsenergie eines Fahrzeugs aus seiner Geschwindigkeit berechnet werden kann	○	○	○	○	○
Mehr Einblick erhalten, welche kraftsparenden Geräte in einer Autowerkstatt verwendet werden	○	○	○	○	○
Mehr Einblick erhalten, welche künstlichen Organe (z.B. Herz als Blutpumpe) und Gelenke heute in der Medizin zur Verfügung stehen	○	○	○	○	○
Mit Rollen und Seilen verschiedene Flaschenzüge bauen und ausprobieren	○	○	○	○	○
Sich ein Sicherheitsfahrzeug ausdenken, in dem auch bei schweren Unfällen Fahrer und Beifahrer wenig oder nichts passiert	○	○	○	○	○
Darüber nachdenken, wie man aus dem Bremsweg eines Autos seine Geschwindigkeit vor dem Abbremsen berechnen kann	○	○	○	○	○
Sich mit Unfallstatistiken beschäftigen und über den Sinn von Geschwindigkeitsbegrenzungen diskutieren	○	○	○	○	○



Wie ausführlich wurde dieses Thema im Unterricht behandelt?

	sehr ausführlich	ausführlich	mittel	nur kurz	gar nicht
Kraft	○	○	○	○	○
Geschwindigkeit	○	○	○	○	○
Bewegungsenergie	○	○	○	○	○
Pumpe	○	○	○	○	○
Flaschenzug	○	○	○	○	○
Hebebühne	○	○	○	○	○
Bremsweg	○	○	○	○	○

Page 03

Teilchenphysik 1

TEILCHENPHYSIK oder *Woraus wir eigentlich bestehen*

Alles, was man zumindest theoretisch berühren kann, wird als Materie bezeichnet. Dazu zählen nicht nur wir Menschen, sondern auch Sterne und Planeten. Teilchenphysiker*innen erforschen, woraus alle Materie besteht und was ihre Bestandteile zusammenhält. Ein menschliches Haar zum Beispiel ist aus Atomen aufgebaut, und ein Atom aus einem Atomkern-Bereich und Elektronen, die diesen umgeben. Das Elektron ist ein sogenanntes *Elementarteilchen*. Diese sind unteilbar.

Erkenntnisse über den Aufbau der Materie gewinnt man mithilfe von Experimenten. Zum Beispiel beschleunigt man Teilchen auf sehr hohe Geschwindigkeiten, um sie dann zusammenstoßen zu lassen. Bei diesem Zusammenstoß entstehen neue Teilchen, die von Detektoren aufgezeichnet werden. Diese sind mehrere Stockwerke hohe Geräte, die bis zu 40 Millionen Aufzeichnungen pro Sekunde machen können. Teilchenphysiker*innen werten diese Aufzeichnungen aus. So können sie zum Beispiel die Prozesse erforschen, die sehr kurz nach dem Urknall stattgefunden haben, um besser zu verstehen, wie unser Universum entstanden ist.

Im Grunde können wir alle physikalischen Phänomene auf die Wechselwirkungen zwischen Teilchen zurückführen, wenn wir ganz genau hinschauen. Zum Beispiel wechselwirken Elektronen miteinander, weil sie eine elektrische Ladung haben: Diese Wechselwirkung verhindert, dass du durch den Stuhl fällst, auf dem du vermutlich gerade sitzt. Denn die Elektronen deines Körpers und die Elektronen des Stuhls stoßen sich gegenseitig ab und können nicht beliebig nah zusammengebracht werden.

Außerdem hat Forschung in der Teilchenphysik viele Anwendungen, zum Beispiel bei der Diagnose und Behandlung von Krankheiten oder bei der Feststellung der Echtheit eines Kunstwerks.

Wie gerne würdest du im Zusammenhang mit diesem Thema das Folgende tun? Bitte klicke auf "Weiter"!

Page 04

Teilchenphysik 2

Mein Interesse daran ist

	sehr groß	groß	mittel	gering	sehr gering
Mehr darüber erfahren, wie Geräte funktionieren, die Teilchen detektieren (z.B. Digitalkamera)	○	○	○	○	○
Mehr darüber erfahren, wie Teilchenphysik zum Verständnis des Urknalls beiträgt	○	○	○	○	○
Mehr darüber erfahren, wie mithilfe von Teilchendetektoren geschmuggelte Waffen in einem Container entdeckt werden können	○	○	○	○	○
Mehr darüber erfahren, welche Elementarteilchen man im Atomkern-Bereich findet	○	○	○	○	○
Mehr darüber erfahren, wie groß die Massen der Elementarteilchen im Vergleich zueinander sind	○	○	○	○	○
Mehr Einblick erhalten, wie in der Elektronik-Industrie mit Teilchenbeschleunigern gearbeitet wird	○	○	○	○	○
Mehr Einblick erhalten, wie in einem medizinischen Diagnose-Zentrum gearbeitet wird	○	○	○	○	○
Ein Handy in einen Teilchendetektor umbauen und ausprobieren	○	○	○	○	○
Ein Experiment planen, um zu zeigen, wie Teilchen beschleunigt werden	○	○	○	○	○
Berechnen, wie groß die Energie beim Zusammenstoß zweier Teilchen ist, die sich mit nahezu Lichtgeschwindigkeit bewegen	○	○	○	○	○
Darüber diskutieren, wie Erkenntnisse in der Teilchenphysik unser Alltagsleben verändert haben	○	○	○	○	○

Wie ausführlich wurde dieses Thema im Unterricht behandelt?

	sehr ausführlich	ausführlich	mittel	nur kurz	gar nicht
Elementarteilchen	○	○	○	○	○
Teilchendetektor	○	○	○	○	○
Teilchenbeschleuniger	○	○	○	○	○
Atomkern-Bereich	○	○	○	○	○
Elektron	○	○	○	○	○
Urknall	○	○	○	○	○
Lichtgeschwindigkeit	○	○	○	○	○

Page 05

Vergleich

Mein Interesse daran ist

	sehr groß	groß	mittel	gering	sehr gering
Teilchenphysik	○	○	○	○	○
Von den Bewegungen	○	○	○	○	○

Page 06

Wie sehr stimmst du mit den folgenden Aussagen überein?

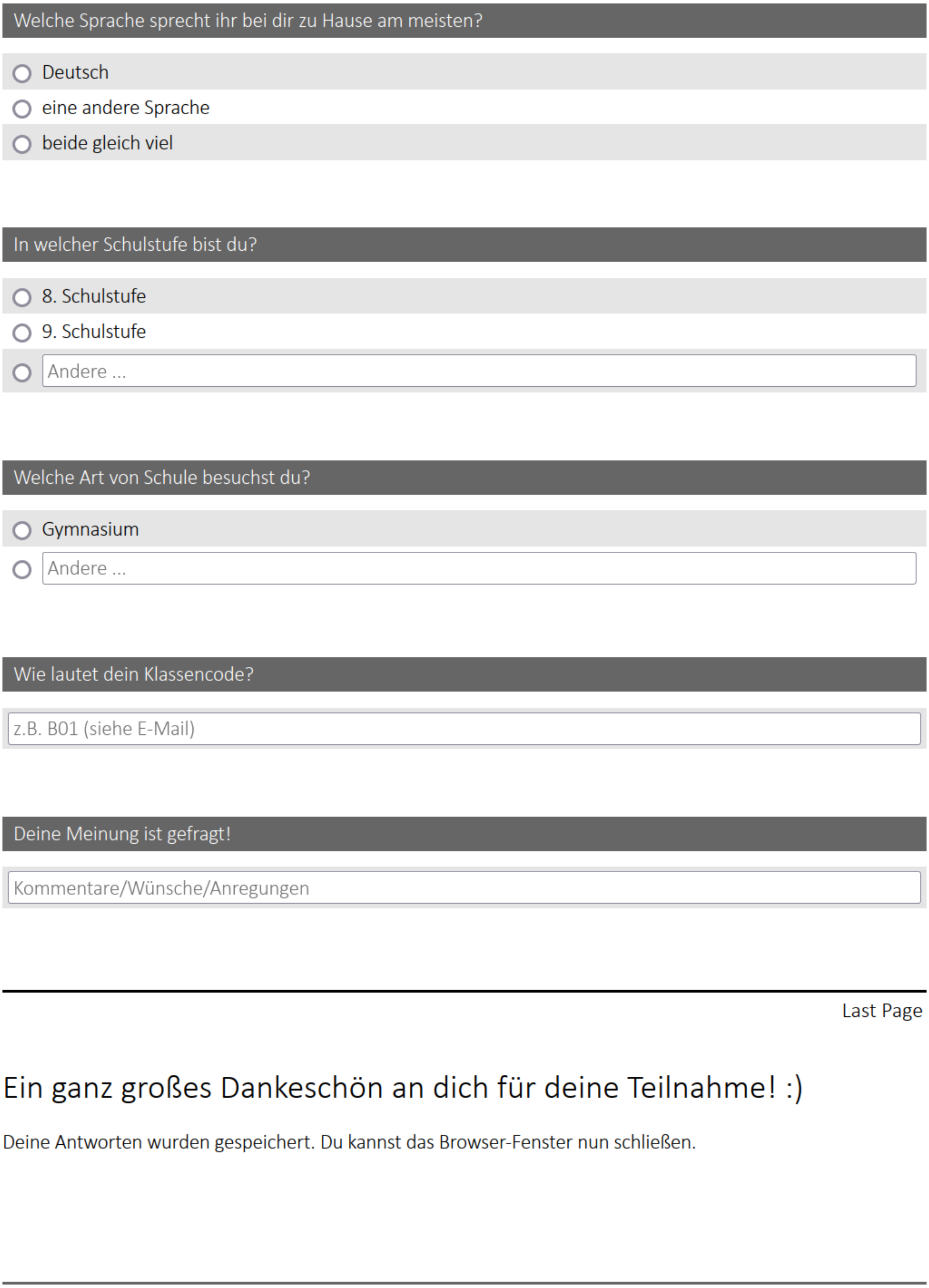
	stimme ganz zu	stimme eher zu	stimme eher nicht zu	stimme gar nicht zu
Ich glaube, dass ich anspruchsvollen Stoff im Physikunterricht leicht lernen kann.	○	○	○	○
Normalerweise kann ich Prüfungsfragen im Physikunterricht gut beantworten.	○	○	○	○
Mein*e Lehrer*in sieht mich als Physik-Person.	○	○	○	○
Ich lerne neuen Stoff im Physikunterricht schnell.	○	○	○	○
Meine Eltern sehen mich als Physik-Person.	○	○	○	○
Den Stoff im Physikunterricht finde ich einfach.	○	○	○	○
Meine Mitschüler*innen sehen mich als Physik-Person.	○	○	○	○
Wenn ich in Physik unterrichtet werde, verstehe ich neue Begriffe leicht.	○	○	○	○
Meine Freund*innen sehen mich als Physik-Person.	○	○	○	○
Es fällt mir leicht, neue Ideen im Physikunterricht zu verstehen.	○	○	○	○

Bitte kreuze dein Geschlecht an.

○ weiblich
○ männlich
○ keine Angabe

Wie alt bist du?

○ 14 Jahre
○ 15 Jahre
○ 16 Jahre
○ Andere ...



Welche Sprache sprecht ihr bei dir zu Hause am meisten?

- ○ Deutsch
- ○ eine andere Sprache
- ○ beide gleich viel

In welcher Schulstufe bist du?

- ○ 8. Schulstufe
- ○ 9. Schulstufe
- ○ Andere ...

Welche Art von Schule besuchst du?

- ○ Gymnasium
- ○ Andere ...

Wie lautet dein Klassencode?

z.B. B01 (siehe E-Mail)

Deine Meinung ist gefragt!

Kommentare/Wünsche/Anregungen

Last Page

Ein ganz großes Dankeschön an dich für deine Teilnahme! :)

Deine Antworten wurden gespeichert. Du kannst das Browser-Fenster nun schließen.

10.2.6. Version 2 of the questionnaire with new mechanics text

Page 01

Mechanik 1

MECHANIK oder *Was Kräfte bewirken*

Nicht alles, was wir im Alltag als Kraft bezeichnen, ist auch im physikalischen Sinne eine Kraft. So hast du bestimmt schon mal gehört, dass etwas «Kraft hat». Doch Kraft kann man nicht haben, man kann sie nur ausüben. Oft erkennen wir eine Kraft an ihrer Wirkung, zum Beispiel etwas in Bewegung zu versetzen oder zu verformen.

Manchmal sind wir Menschen nicht stark genug, um eine genügend große Kraft auszuüben. Manchmal wollen wir aber einfach von schwerer körperlicher Anstrengung befreit sein. Zum Glück wusste man schon in der Antike, wie man Kräfte verstärken kann. So erfand man schon vor 3000 Jahren den Flaschenzug, eine nützliche Konstruktion aus Seilen und Rollen, mit der man zum Beispiel schwere Tempelsäulen aufrichten konnte. Heute gibt es viele Geräte, die uns Anstrengungen abnehmen: Pumpen holen Grundwasser aus großer Tiefe an die Erdoberfläche. Hebebühnen heben Elefanten in Tierkliniken, Patient*innen in Krankenhäusern oder Autos in Autowerkstätten. Außerdem genügt ein leichtes Tippen auf das Gaspedal, um ein Fahrzeug auf 130 km/h zu beschleunigen!

Aber macht man sich immer klar, welche Energien man dabei erreicht? Ein frontaler Aufprall eines Autos auf eine Mauer bei einer Geschwindigkeit von nur 50 km/h ist zum Beispiel genauso gefährlich wie ein Sturz aus 10 Metern Höhe!

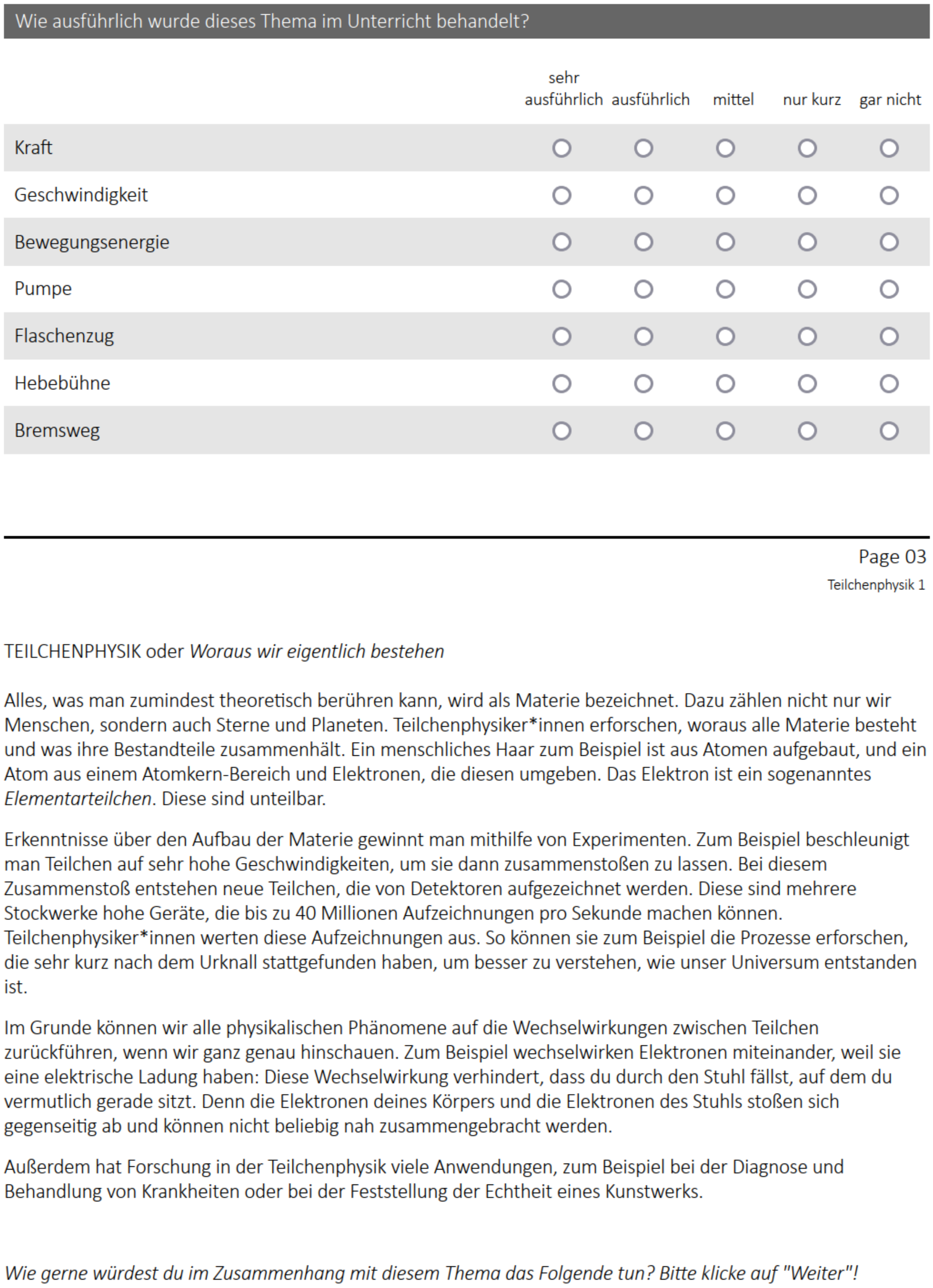
Um Fahrzeuge sicherer zu machen, werden Erkenntnisse der Mechanik genutzt. Zum Beispiel verhindert der Sicherheitsgurt, dass du bei einem Aufprall nach vorne fliegst. Außerdem gibt es im vorderen und hinteren Bereich jedes Fahrzeugs sogenannte Knautschzonen, in denen die Bewegungsenergie durch Verformung von Blechteilen aufgenommen wird. So wird die Wucht des Aufpralls bei einem Unfall verringert.

Wie gerne würdest du im Zusammenhang mit diesem Thema das Folgende tun? Bitte klicke auf "Weiter"!

Page 02
Mechanik 2

Mein Interesse daran ist

	sehr groß	groß	mittel	gering	sehr gering
Mehr darüber erfahren, wie Geräte funktionieren, die Kräfte verstärken (z.B. Flaschenzug, Hebebühne)	○	○	○	○	○
Mehr darüber erfahren, wie man Erdöl aus sehr großen Tiefen (z.B. aus 3000 Metern) heraufpumpen kann	○	○	○	○	○
Mehr darüber erfahren, wie die Wahrscheinlichkeit eines Autounfalls und die Schwere der Unfallfolgen mit zunehmender Geschwindigkeit wachsen	○	○	○	○	○
Mehr darüber erfahren, wie die Bewegungsenergie (Wucht) eines Fahrzeugs in andere Energieformen umgewandelt werden kann (z.B. in den Bremsen oder in der Knautschzone)	○	○	○	○	○
Mehr Einblick erhalten, wie die Bewegungsenergie eines Fahrzeugs aus seiner Geschwindigkeit berechnet werden kann	○	○	○	○	○
Mehr Einblick erhalten, welche kraftsparenden Geräte in einer Autowerkstatt verwendet werden	○	○	○	○	○
Mehr Einblick erhalten, welche künstlichen Organe (z.B. Herz als Blutpumpe) und Gelenke heute in der Medizin zur Verfügung stehen	○	○	○	○	○
Mit Rollen und Seilen verschiedene Flaschenzüge bauen und ausprobieren	○	○	○	○	○
Sich ein Sicherheitsfahrzeug ausdenken, in dem auch bei schweren Unfällen Fahrer und Beifahrer wenig oder nichts passiert	○	○	○	○	○
Darüber nachdenken, wie man aus dem Bremsweg eines Autos seine Geschwindigkeit vor dem Abbremsen berechnen kann	○	○	○	○	○
Sich mit Unfallstatistiken beschäftigen und über den Sinn von Geschwindigkeitsbegrenzungen diskutieren	○	○	○	○	○



Wie ausführlich wurde dieses Thema im Unterricht behandelt?

	sehr ausführlich	ausführlich	mittel	nur kurz	gar nicht
Kraft	○	○	○	○	○
Geschwindigkeit	○	○	○	○	○
Bewegungsenergie	○	○	○	○	○
Pumpe	○	○	○	○	○
Flaschenzug	○	○	○	○	○
Hebebühne	○	○	○	○	○
Bremsweg	○	○	○	○	○

Page 03
Teilchenphysik 1

TEILCHENPHYSIK oder *Woraus wir eigentlich bestehen*

Alles, was man zumindest theoretisch berühren kann, wird als Materie bezeichnet. Dazu zählen nicht nur wir Menschen, sondern auch Sterne und Planeten. Teilchenphysiker*innen erforschen, woraus alle Materie besteht und was ihre Bestandteile zusammenhält. Ein menschliches Haar zum Beispiel ist aus Atomen aufgebaut, und ein Atom aus einem Atomkern-Bereich und Elektronen, die diesen umgeben. Das Elektron ist ein sogenanntes *Elementarteilchen*. Diese sind unteilbar.

Erkenntnisse über den Aufbau der Materie gewinnt man mithilfe von Experimenten. Zum Beispiel beschleunigt man Teilchen auf sehr hohe Geschwindigkeiten, um sie dann zusammenstoßen zu lassen. Bei diesem Zusammenstoß entstehen neue Teilchen, die von Detektoren aufgezeichnet werden. Diese sind mehrere Stockwerke hohe Geräte, die bis zu 40 Millionen Aufzeichnungen pro Sekunde machen können. Teilchenphysiker*innen werten diese Aufzeichnungen aus. So können sie zum Beispiel die Prozesse erforschen, die sehr kurz nach dem Urknall stattgefunden haben, um besser zu verstehen, wie unser Universum entstanden ist.

Im Grunde können wir alle physikalischen Phänomene auf die Wechselwirkungen zwischen Teilchen zurückführen, wenn wir ganz genau hinschauen. Zum Beispiel wechselwirken Elektronen miteinander, weil sie eine elektrische Ladung haben: Diese Wechselwirkung verhindert, dass du durch den Stuhl fällst, auf dem du vermutlich gerade sitzt. Denn die Elektronen deines Körpers und die Elektronen des Stuhls stoßen sich gegenseitig ab und können nicht beliebig nah zusammengebracht werden.

Außerdem hat Forschung in der Teilchenphysik viele Anwendungen, zum Beispiel bei der Diagnose und Behandlung von Krankheiten oder bei der Feststellung der Echtheit eines Kunstwerks.

Wie gerne würdest du im Zusammenhang mit diesem Thema das Folgende tun? Bitte klicke auf "Weiter"!

Page 04
Teilchenphysik 2

Mein Interesse daran ist

	sehr groß	groß	mittel	gering	sehr gering
Mehr darüber erfahren, wie Geräte funktionieren, die Teilchen detektieren (z.B. Digitalkamera)	○	○	○	○	○
Mehr darüber erfahren, wie Teilchenphysik zum Verständnis des Urknalls beiträgt	○	○	○	○	○
Mehr darüber erfahren, wie mithilfe von Teilchendetektoren geschmuggelte Waffen in einem Container entdeckt werden können	○	○	○	○	○
Mehr darüber erfahren, welche Elementarteilchen man im Atomkern-Bereich findet	○	○	○	○	○
Mehr darüber erfahren, wie groß die Massen der Elementarteilchen im Vergleich zueinander sind	○	○	○	○	○
Mehr Einblick erhalten, wie in der Elektronik-Industrie mit Teilchenbeschleunigern gearbeitet wird	○	○	○	○	○
Mehr Einblick erhalten, wie in einem medizinischen Diagnose-Zentrum gearbeitet wird	○	○	○	○	○
Ein Handy in einen Teilchendetektor umbauen und ausprobieren	○	○	○	○	○
Ein Experiment planen, um zu zeigen, wie Teilchen beschleunigt werden	○	○	○	○	○
Berechnen, wie groß die Energie beim Zusammenstoß zweier Teilchen ist, die sich mit nahezu Lichtgeschwindigkeit bewegen	○	○	○	○	○
Darüber diskutieren, wie Erkenntnisse in der Teilchenphysik unser Alltagsleben verändert haben	○	○	○	○	○

Wie ausführlich wurde dieses Thema im Unterricht behandelt?

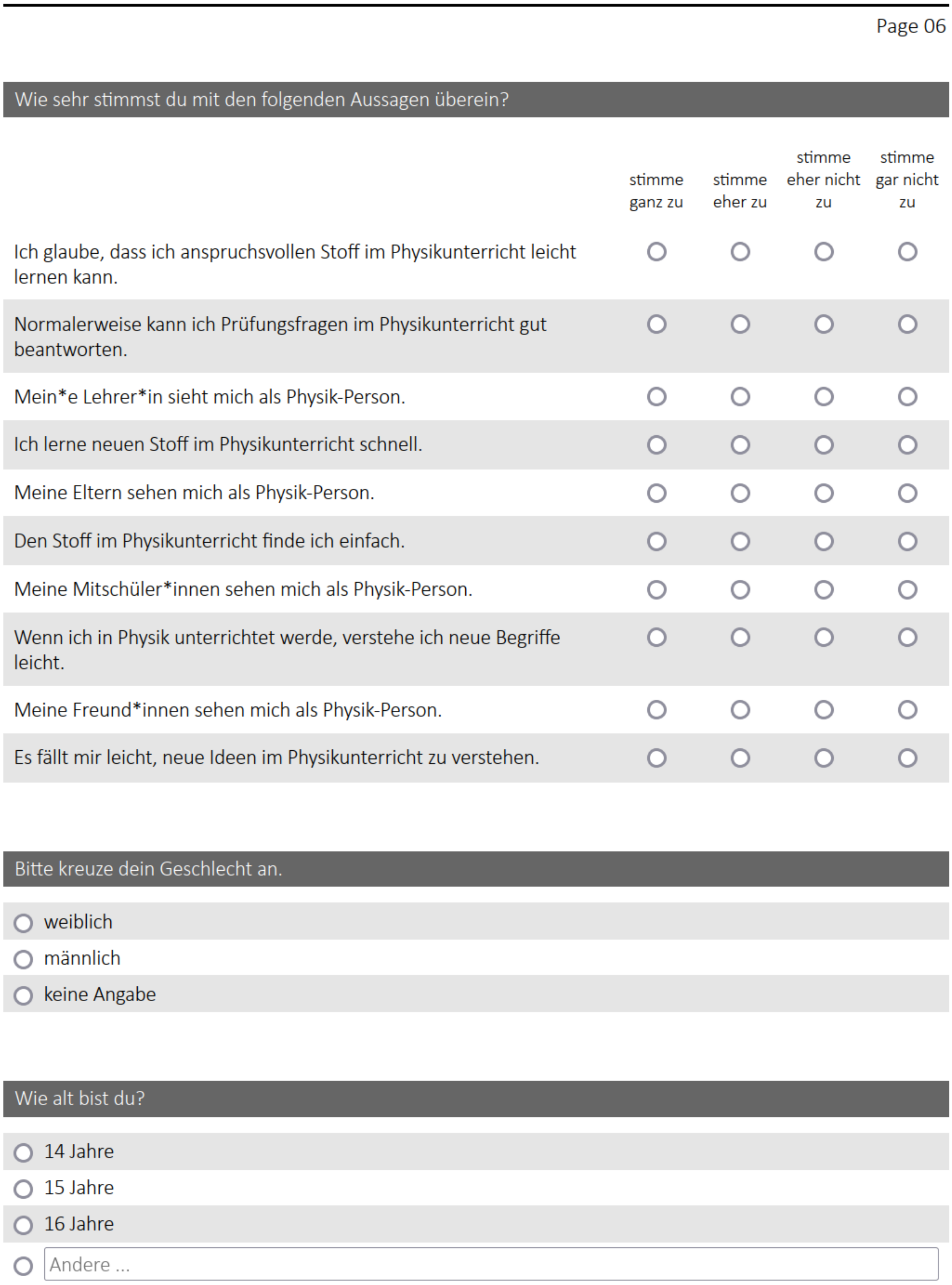
	sehr ausführlich	ausführlich	mittel	nur kurz	gar nicht
Elementarteilchen	○	○	○	○	○
Teilchendetektor	○	○	○	○	○
Teilchenbeschleuniger	○	○	○	○	○
Atomkern-Bereich	○	○	○	○	○
Elektron	○	○	○	○	○
Urknall	○	○	○	○	○
Lichtgeschwindigkeit	○	○	○	○	○

Page 05

Allgemeininteresse

Mein Interesse daran ist

	sehr groß	groß	mittel	gering	sehr gering
Teilchenphysik	○	○	○	○	○
Mechanik	○	○	○	○	○



Page 06

Wie sehr stimmst du mit den folgenden Aussagen überein?

	stimme ganz zu	stimme eher zu	stimme eher nicht zu	stimme gar nicht zu
Ich glaube, dass ich anspruchsvollen Stoff im Physikunterricht leicht lernen kann.	○	○	○	○
Normalerweise kann ich Prüfungsfragen im Physikunterricht gut beantworten.	○	○	○	○
Mein*e Lehrer*in sieht mich als Physik-Person.	○	○	○	○
Ich lerne neuen Stoff im Physikunterricht schnell.	○	○	○	○
Meine Eltern sehen mich als Physik-Person.	○	○	○	○
Den Stoff im Physikunterricht finde ich einfach.	○	○	○	○
Meine Mitschüler*innen sehen mich als Physik-Person.	○	○	○	○
Wenn ich in Physik unterrichtet werde, verstehe ich neue Begriffe leicht.	○	○	○	○
Meine Freund*innen sehen mich als Physik-Person.	○	○	○	○
Es fällt mir leicht, neue Ideen im Physikunterricht zu verstehen.	○	○	○	○

Bitte kreuze dein Geschlecht an.

- ○ weiblich
- ○ männlich
- ○ keine Angabe

Wie alt bist du?

- ○ 14 Jahre
- ○ 15 Jahre
- ○ 16 Jahre
- ○ Andere ...

Welche Sprache sprecht ihr bei dir zu Hause am meisten?

○ Deutsch
○ eine andere Sprache
○ beide gleich viel

In welcher Schulstufe bist du?

○ 8. Schulstufe
○ 9. Schulstufe
○ Andere ...

Welche Art von Schule besuchst du?

○ Gymnasium
○ Andere ...

Wie lautet dein Klassencode?

z.B. B01 (siehe E-Mail)

Deine Meinung ist gefragt!

Kommentare/Wünsche/Anregungen

Last Page

Ein ganz großes Dankeschön an dich für deine Teilnahme! :)

Deine Antworten wurden gespeichert. Du kannst das Browser-Fenster nun schließen.

10.2.7. Item tables of the groups of interest in mechanics and particle physics

The following tables present more detailed information about the items to assess interest in mechanics and particle physics for the different groups of interest in mechanics and particle physics, respectively, based on the sample size for calculating the mixed Rasch models (*N* = 1187). Total score refers to the total raw score of all respondents who answered the item. Total count refers to the total number of respondents who answered the item. Measure refers to the Rasch item measure in logit units. Lower and higher item measures represent more and less interesting items, respectively. Model SE refers to the standard error of the item measure in logit units. ZQ refers to a fit statistic based on quotient of log-likelihood ratios.

Table 38. Group 1_M

Item ID	Total score	Total count	Item measure [logits]	Model SE [logits]	ZQ
M01	1456	478	0.49	0.07	-0.02
M02	1452	478	0.45	0.07	1.43
M03	1766	478	-0.75	0.07	0.07
M04	1485	478	0.38	0.07	-0.28
M05	1415	478	0.68	0.07	-0.05
M06	1423	478	0.60	0.07	1.28
M07	1974	478	-1.66	0.07	-0.87
M08	1529	478	0.15	0.07	-1.01
M09	1797	478	-0.87	0.07	-0.89
M10	1426	478	0.60	0.07	0.04
M11	1588	478	-0.06	0.07	0.56

Table 39. Group 2_M

Item ID	Total score	Total count	Item measure [logits]	Model SE [logits]	ZQ
M01	1042	351	0.05	0.06	1.63
M02	1196	351	-0.38	0.06	0.87
M03	1205	351	-0.47	0.06	-0.45
M04	845	351	0.64	0.06	-0.88
M05	692	351	1.17	0.07	0.42
M06	862	351	0.61	0.06	0.59
M07	1507	351	-1.52	0.07	0.48
M08	1383	351	-0.97	0.06	-0.28
M09	1303	351	-0.77	0.06	-1.34
M10	710	351	1.11	0.07	-0.23
M11	871	351	0.53	0.06	-0.77

Table 40. Group 3_M

Item ID	Total score	Total count	Item measure [logits]	Model SE [logits]	ZQ
M01	499	201	0.34	0.07	1.09
M02	563	201	0.11	0.07	1.78
M03	742	201	-0.67	0.07	-0.58
M04	517	201	0.26	0.07	-0.47
M05	492	201	0.43	0.07	-1.26
M06	451	201	0.56	0.07	-0.51
M07	838	201	-1.27	0.09	-0.36
M08	409	201	0.62	0.09	0.56
M09	693	201	-0.45	0.07	0.74
M10	499	201	0.43	0.07	-0.82
M11	684	201	-0.35	0.07	-0.17

Table 41. Group 4_M

Item ID	Total score	Total count	Item measure [logits]	Model SE [logits]	ZQ
M01	542	157	-0.13	0.08	0.26
M02	511	157	0.03	0.08	1.24
M03	549	157	-0.26	0.08	-0.33
M04	555	157	-0.18	0.08	-0.88
M05	551	157	-0.20	0.08	-0.16
M06	503	157	0.11	0.08	-0.33
M07	401	157	0.52	0.08	1.30
M08	487	157	0.22	0.08	0.58
M09	527	157	-0.17	0.08	-0.50
M10	539	157	-0.14	0.08	-0.72
M11	484	157	0.21	0.08	-0.32

Table 42. Group 1_{PP}

Item ID	Total score	Total count	Item measure [logits]	Model SE [logits]	ZQ
PP01	1809	567	0.08	0.06	0.83
PP02	1932	567	-0.24	0.06	-0.72
PP03	2247	567	-1.10	0.06	-0.76
PP04	1625	567	0.62	0.06	-0.43
PP05	1363	567	1.40	0.06	0.15
PP06	1668	567	0.53	0.06	0.04
PP07	2168	567	-0.85	0.06	0.40
PP08	2195	567	-0.95	0.06	-1.27
PP09	1907	567	-0.15	0.06	-0.14
PP10	1662	567	0.49	0.06	0.94
PP11	1785	567	0.17	0.06	1.18

Table 43. Group 2_{PP}

Item ID	Total score	Total count	Item measure [logits]	Model SE [logits]	ZQ
PP01	1058	370	-0.15	0.05	-0.44
PP02	1109	370	-0.26	0.05	-0.10
PP03	1280	370	-0.61	0.05	0.02
PP04	828	370	0.28	0.05	-0.45
PP05	656	370	0.70	0.05	-0.87
PP06	764	370	0.39	0.05	-1.13
PP07	1210	370	-0.46	0.05	3.57
PP08	1201	370	-0.45	0.05	0.61
PP09	943	370	0.03	0.05	-0.06
PP10	781	370	0.37	0.05	-1.91
PP11	892	370	0.15	0.05	-0.92

Table 44. Group 3_{PP}

Item ID	Total score	Total count	Item measure [logits]	Model SE [logits]	ZQ
PP01	980	250	-0.02	0.09	0.15
PP02	1085	250	-0.81	0.09	-0.82
PP03	953	250	0.04	0.09	0.95
PP04	996	250	-0.11	0.09	-0.36
PP05	901	250	0.57	0.09	0.03
PP06	947	250	0.19	0.09	-0.10
PP07	793	250	1.07	0.09	1.47
PP08	1024	250	-0.41	0.09	-0.31
PP09	1020	250	-0.28	0.09	-0.47
PP10	1057	250	-0.52	0.09	-0.74
PP11	933	250	0.28	0.09	0.21

10.2.8. Wright Maps of the groups of interest in mechanics and particle physics

The following figures present the Wright Maps for the items to assess interest in mechanics and particle physics and the respondents assigned to the different groups of interest in mechanics and particle physics, respectively, based on the sample size for calculating the mixed Rasch models (N=1187). The y-axis shows the person and item measures, respectively, in logit units. On the left side of each Wright Map the relative number of students per logit range of person measures is shown. The person measures represent the students' degrees of interest in mechanics or particle physics, respectively. On the right side of each Wright Map the items to assess interest in mechanics or particle physics, respectively, are shown.

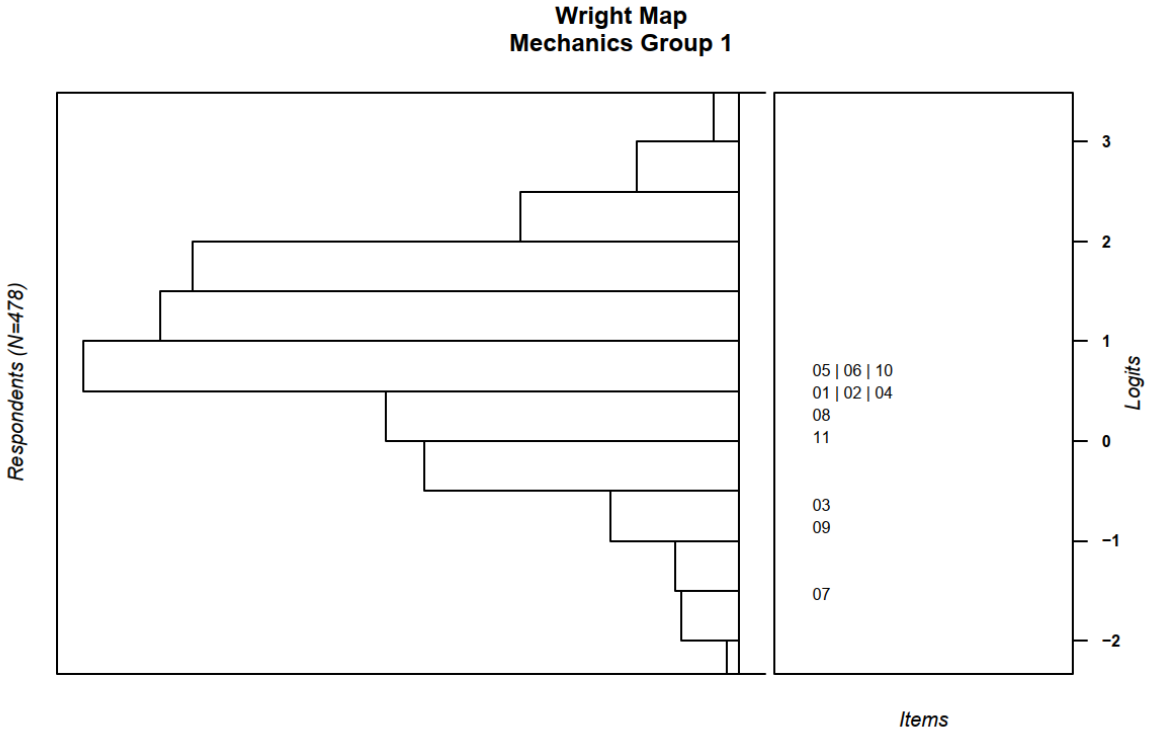


Figure 25. Group 1_M

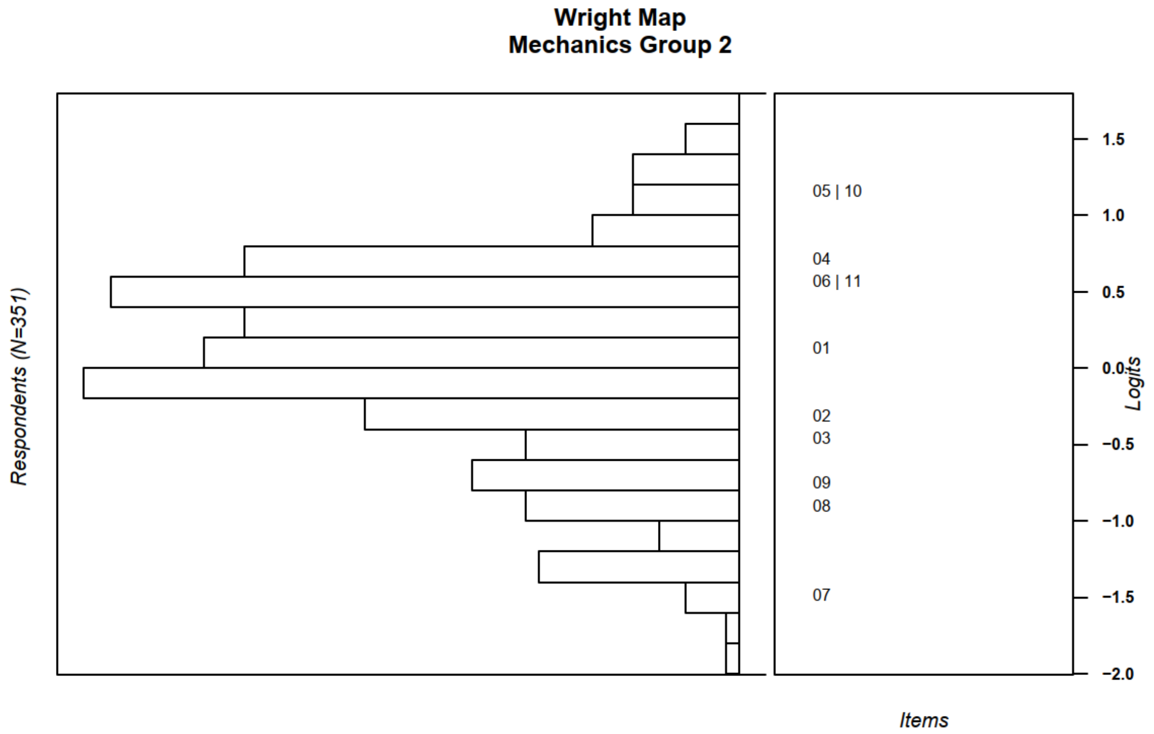


Figure 26. Group 2_M

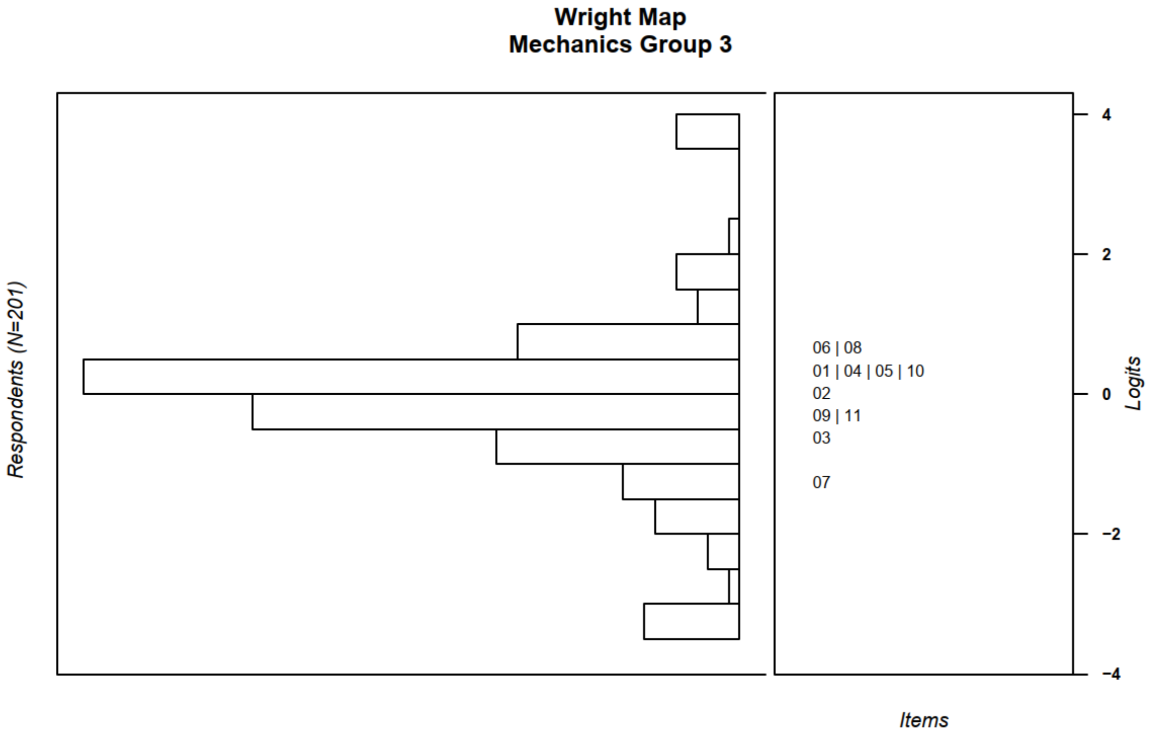


Figure 27. Group 3_M

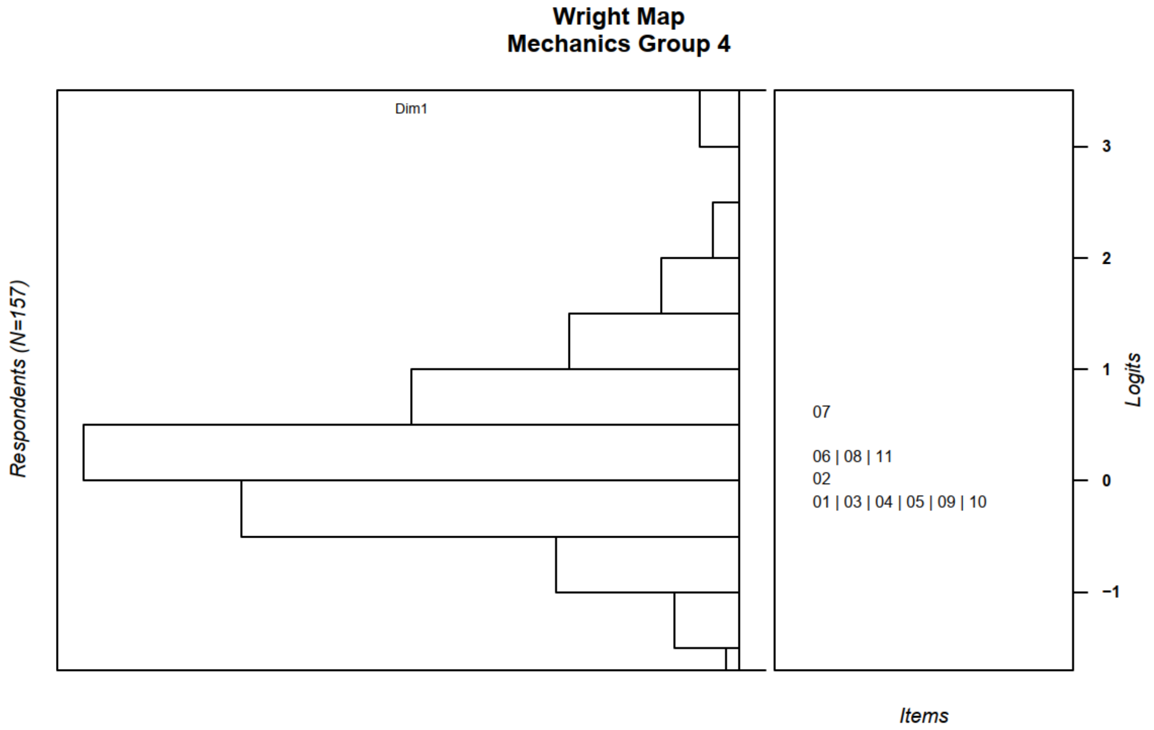


Figure 28. Group 4_M

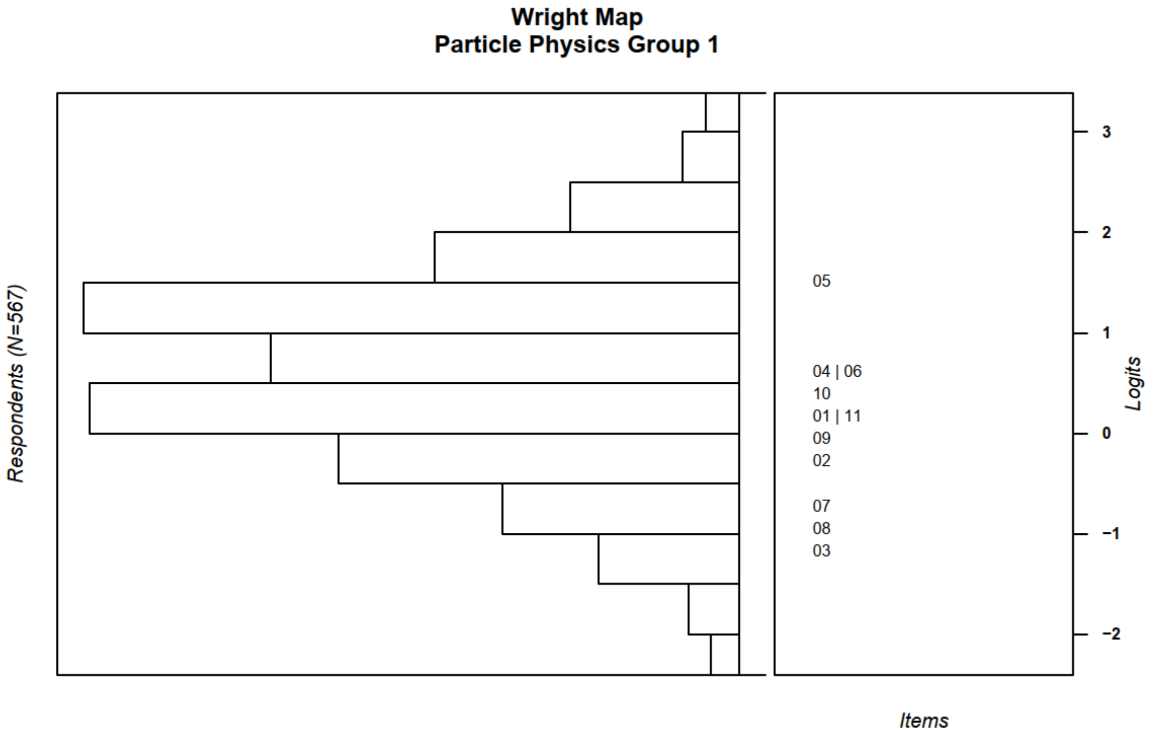


Figure 29. Group 1_{PP}

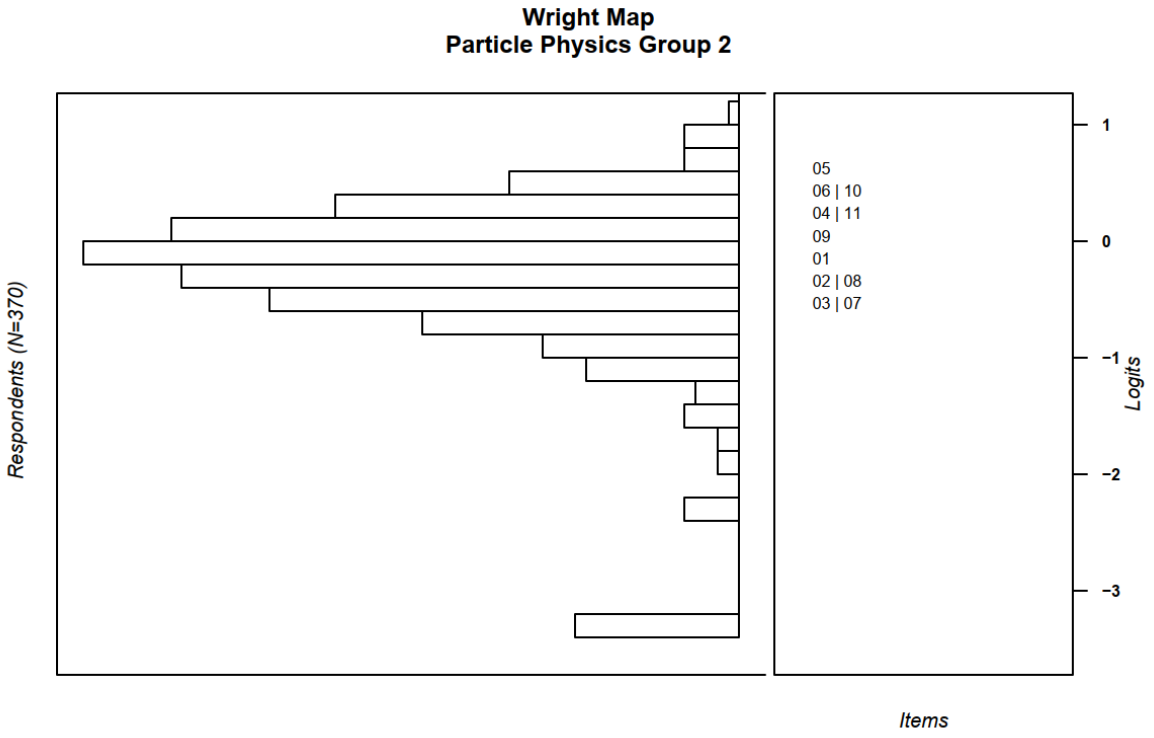


Figure 30. Group 2_{PP}

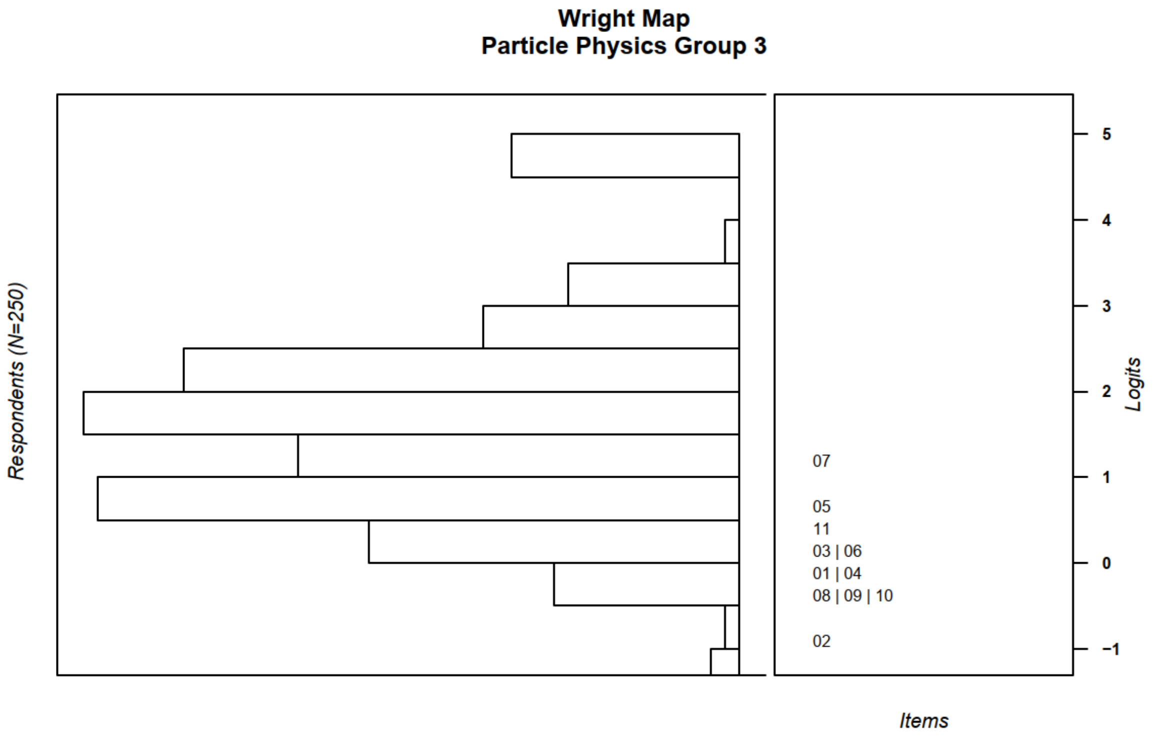


Figure 31. Group 3_{PP}

11. ACKNOWLEDGEMENTS

It would not have been possible without the support of numerous helpful advisors, colleagues, and friends to conclude not only my doctoral thesis but also an incredibly inspiring and educational time of my life. This time was characterised by the deep fascination with my research questions and methods, interesting discussions with fellow researchers, motivating encounters with students and teachers, and most importantly by the unique and collaborative spirit at CERN in which I was fortunate to take part.

From the very beginning, I was expertly supported by my university advisor, Univ.-Prof. Dr. Martin Hopf. I am incredibly grateful for his positive impact on my life since he aroused my interest in physics education research and gave me the opportunity to gain initial experience in this field at the University of Vienna. Moreover, he showed confidence in me to conduct the research for my diploma thesis and my doctoral thesis at CERN. He guided me throughout my research life with his profound and broad knowledge. I highly appreciate his patient and kind advice and his accurate yet realistic approach to physics education research, which I also aspire to apply in my own research career. Thank you so much, Martin!

I also want to thank my CERN supervisor, Dr. Sascha Schmeling, for his continuous support. When I started my first contract at CERN five years ago, I could never have dreamed of all the amazing opportunities he would present me with or the projects he would allow me to work on and that allowed me to grow, both personally and professionally. I am extremely grateful for these opportunities, and I am very much looking forward to our future collaboration.

I am particularly grateful to my former and current colleagues in the physics education research group at CERN for their various support. In especial, I want to thank Dr. Julia Woithe very much for her great co-supervision, constant support, interesting discussions, and invaluable help and suggestions throughout my doctoral studentship.

Special thanks are due to all the students and teachers who participated in my studies for their time, effort, and interest!

I am also incredibly grateful to all the amazing people that I was fortunate to meet throughout the last four years for making my doctoral studentship at CERN an enriching and unforgettable personal experience.

Finally, I want to thank my relatives, especially my mother, grandmother, and aunts, and my friends for their life-long support and for brightening up my life.

Vienna, May 2023