International Journal of Science Education

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tsed20

Students' interest in particle physics: conceptualisation, instrument development, and evaluation using Rasch theory and analysis

Sarah Zoechling, Martin Hopf, Julia Woithe & Sascha Schmeling

To cite this article: Sarah Zoechling, Martin Hopf, Julia Woithe & Sascha Schmeling (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, DOI: 10.1080/09500693.2022.2122897

To link to this article: https://doi.org/10.1080/09500693.2022.2122897

Published online: 21 Sep 2022.

INTERNATIONAL JOURNAL OF SCIENCE EDUCATION
2022, VOL. 44, NO. 15, 2353–2380
https://doi.org/10.1080/09500693.2022.2122897

Students' interest in particle physics: conceptualisation, instrument development, and evaluation using Rasch theory and analysis

Sarah Zoechling [a,b], Martin Hopf [a], Julia Woithe [b] and Sascha Schmeling [b]

[a]Physics Education Research Group, University of Vienna, Vienna, Austria; [b]Teacher and Student Programmes Section, CERN, Geneva, Switzerland

ABSTRACT

Given the importance of fostering students' interest as a goal of physics education in meeting international science standards, empirical support for the theoretical description of the interest construct is essential. Empirical studies require the use of psychometrically sound measurement instruments. This study developed an instrument to measure students' interest in particle physics (IPPI). Drawing from previous research, we defined interest in particle physics, identified corresponding behaviours, and proposed a hierarchy of students' levels of interest in particle physics. Then, we developed the IPPI, using rating scale items that assessed the latent trait developed from our theory regarding the degree of interest in particle physics. We tested the IPPI in student think-aloud interviews and validated it in a field test on a sample comprising 99 German-speaking grade 9 students. A Rasch analysis provided evidence supporting the content, construct, statistical, and fit validity of the IPPI. We revised the hypothesised hierarchy of students' levels of interest in particle physics based on the item hierarchy revealed by the Rasch analysis. We associated each level with different contexts, such as socio-scientific issues. Knowing about these levels of interest in particle physics can help educators design their learning activities better and foster their students' interest.

ARTICLE HISTORY
Received 19 November 2021
Accepted 5 September 2022

KEYWORDS
Interest; particle physics; instrument development; Rasch analysis; physics education research

Introduction

Fostering interest in physics is a key component in national and international physics education standards (National Research Council, 2013; OECD, 2017). Empirical research has found that interest enhances persistence and achievement while engaging with an object (de Barba et al., 2016; Kauertz & Fischer, 2006; Nuutila et al., 2020). Interest plays an important role in shaping students' course and career choices (Ainley & Ainley, 2011; Maltese & Tai, 2011; Tyson et al., 2007). Nevertheless, students' interest in physics decreases over time despite various efforts invested by

CONTACT Sarah Zoechling sarah.zoechling@cern.ch Teacher and Student Programmes Section, CERN, Esplanade des Particules 1, 1211 Geneva 23, Switzerland

educators in making learning activities more interesting (Galton, 2009; Häußler et al., 1998; Osborne et al., 2003).

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 (Blankenburg et al., 2016; Dierks et al., 2016; Häußler et al., 1998; OECD, 2007, 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 et al., 1998; Sjøberg & Schreiner, 2012). 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, 1987; Häußler et al., 1996; Häußler et al., 1998; Rost et al., 1999; Sievers, 1999). In this sense, the context of a learning activity is also considered the combination of a 'focal event' and its corresponding fields of action (Duranti & Goodwin, 1992; Gilbert, 2006; Habig et al., 2018). 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-concepts (Cheung, 2018; Häußler et al., 1998; Häußler et al., 1998; Kalender et al., 2019; Lavonen et al., 2021; Nuutila et al., 2020; OECD, 2007, 2016; Sjøberg & Schreiner, 2012).

An important empirical study on interest in physics that considers the different aspects of physics and students' characteristics is 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 et al., 1998; Häußler et al., 1998; Rost et al., 1999; Sievers, 1999). In this study, students' specific interest profiles and characteristics were used to categorise them into different types of interest (Häußler et al., 1998). 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).

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. 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?'). Previous studies indicate that certain contexts are more interesting for students than others. In particular, most students demonstrate a high interest in the following contexts: the development of the universe (OECD, 2016), the possibility of life outside earth (Sjøberg & Schreiner, 2012), and the human body (Häußler et al., 1998). Our study aims to investigate students' interest in particle physics. As outlined above, different content of particle physics can be set in different contexts. We 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 is the core objective of this study.

First, we defined interest in particle physics and identified the corresponding behaviours that are aligned with the results of past empirical research, such as the IPN study (Häußler et al., 1998). Second, we derived a theoretical hierarchy of students' levels of interest in particle physics. Third, we developed and evaluated the psychometric properties of an instrument to measure particle physics interest (IPPI). The development procedures were guided by the application of Rasch theory and analysis. They included the creation of an initial item pool and the careful selection of items for the IPPI. Based on a Rasch analysis of the data collected with the IPPI, we propose a potential hierarchy of students' levels of interest in particle physics. We formulated three research questions to guide our study:

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?

Interest in physics

In this section, we describe the psychological construct 'interest' and distinguish it from similar constructs. We summarise the development of students' interest and describe the common assessment methods.

The psychological construct 'interest'

The structure of the psychological construct 'interest' was described by Krapp's (2002) 'Person-object-theory of interest'. Here, interest refers to the relationship between a person and an object (Krapp, 2002). Interest is multifaceted as it involves emotional, value-related, and cognitive-epistemic components (Krapp & Prenzel, 2011): The emotional component refers to the emotions associated with an object; the value-related component considers the significance of an object for a person; and the cognitive-epistemic component comprises the desire to understand better or learn or know more about an object.

Development of 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' (p. 32).

Operating interest can be caused by an existing individual interest or external factors, that is, the interestingness of an object. In the latter case, operating interest refers to 'situational interest'.

Hidi and Renninger (2006) introduced the 'Four-phase model of interest development', which 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 emotional, value-related, and cognitive-epistemic components of interest (Blankenburg & Scheersoi, 2018). Initially, the emotional component prevails, but it becomes less important from phase to phase, wherein the value-related and cognitive-epistemic components become more important (Hidi & Renninger, 2006).

Previous studies (e.g. OECD, 2007, 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 (2006) model, we assume 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

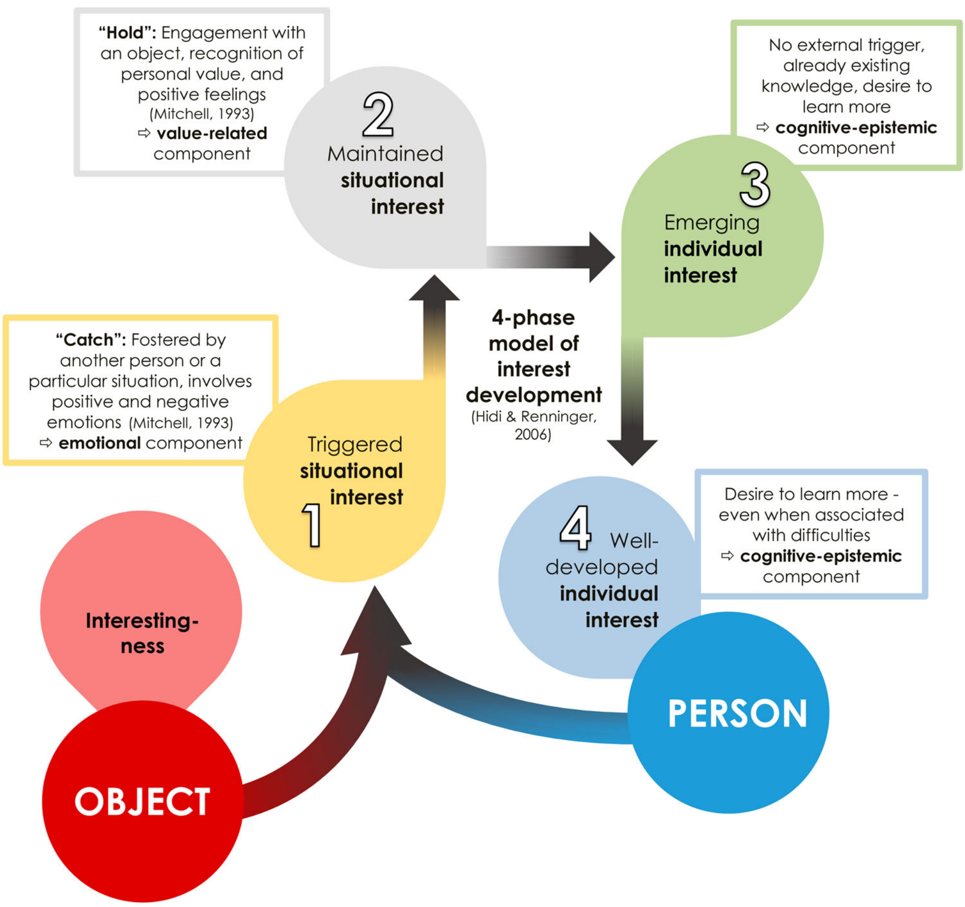

Figure 1. The situational interest is triggered by the interestingness of an object and eventually develops into individual interest in the framework of the 'Person-object theory of interest' (Krapp, 2002) and the 'Four-phase model of interest development' (Hidi & Renninger, 2006)

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.

The measurement of interest in physics

To measure students' interest in physics, various methods are commonly applied. Often, aspects of interest are assessed with open-ended questions or rating scale items (Krapp & Prenzel, 2011). Data regarding students' interest can be collected without presenting the students with any prior stimulus as done, for example, in the Programme for International Student Assessment (PISA; Frey et al., 2009; Mang et al., 2019) and the Relevance of Science Education (Sjøberg & Schreiner, 2012) studies. Students' interest can also be assessed during or after a learning sequence has been completed, such as in the PISA 2006 study (Frey et al., 2009) and by Rösler et al. (2018). Finally, data regarding students' interest can be collected by presenting the students with an introductory text to the assessed object of interest as done, for example, in the IPN study (Häußler et al., 1998).

When deciding on a method to measure students' interest in a certain content area, such as particle physics, we must take their prior experience with the content area into account. 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, in our opinion, the IPN study method is best suited to measure students' interest in particle physics, that is, to present the students with an introductory text and rating scale items. The introductory text aims to cover students' possible deficit in prior experience with the content area by giving them an idea of what it is about. We argue that students express operational interest in each item (see the section titled 'development of interest') while filling in such a measurement instrument. Operational 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. Operational interest may also be caused by the students' already existing individual interest.

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.

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 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 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).

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 prior experience (see section titled 'The measurement of interest in physics'). 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 1.[1] Häußler 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. Blankenburg et al., 2016; Dierks 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 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 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 1). 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 (Drechsel et al., 2011; Levrini et al., 2017; see the section titled 'Conceptualising the construct "interest in particle physics"'). 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

Table 1. Item categories, underlying contexts, and tasks as used in the IPN study (Häußler 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

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.

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 1 for each item category.

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 8–11) 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.

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. 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.

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 1–2, 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).

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 1–2 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 2a.

In comparison, for item I042, every category has an individual probability peak, as seen in Figure 2b. 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.

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 0.6–1.4 are satisfactory for rating scale data (Boone et al., 2014), we chose the smaller range (0.75–1.3) 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).

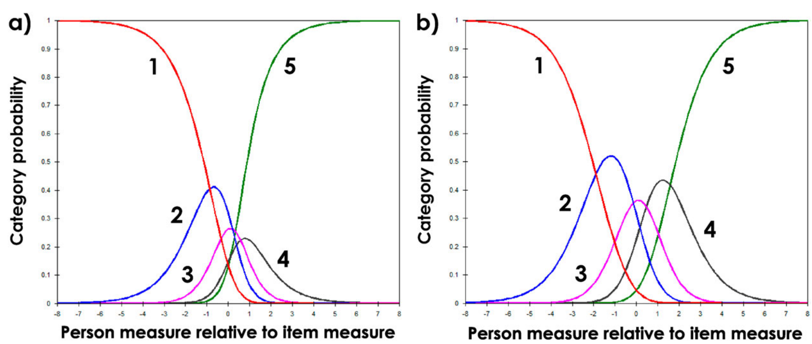

Figure 2. Probability of response for all five rating scale categories as a function of person minus item measure, a) item I081 with non-ordered thresholds (categories 3 and 4 are never the most probable), b) item I042 with ordered thresholds.

While analysing the fit statistics for each item (see Table 2), 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 3). Ideally, the items of an instrument are evenly spread so that they do not measure a similar portion of the trait. In

Table 2. 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.

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,

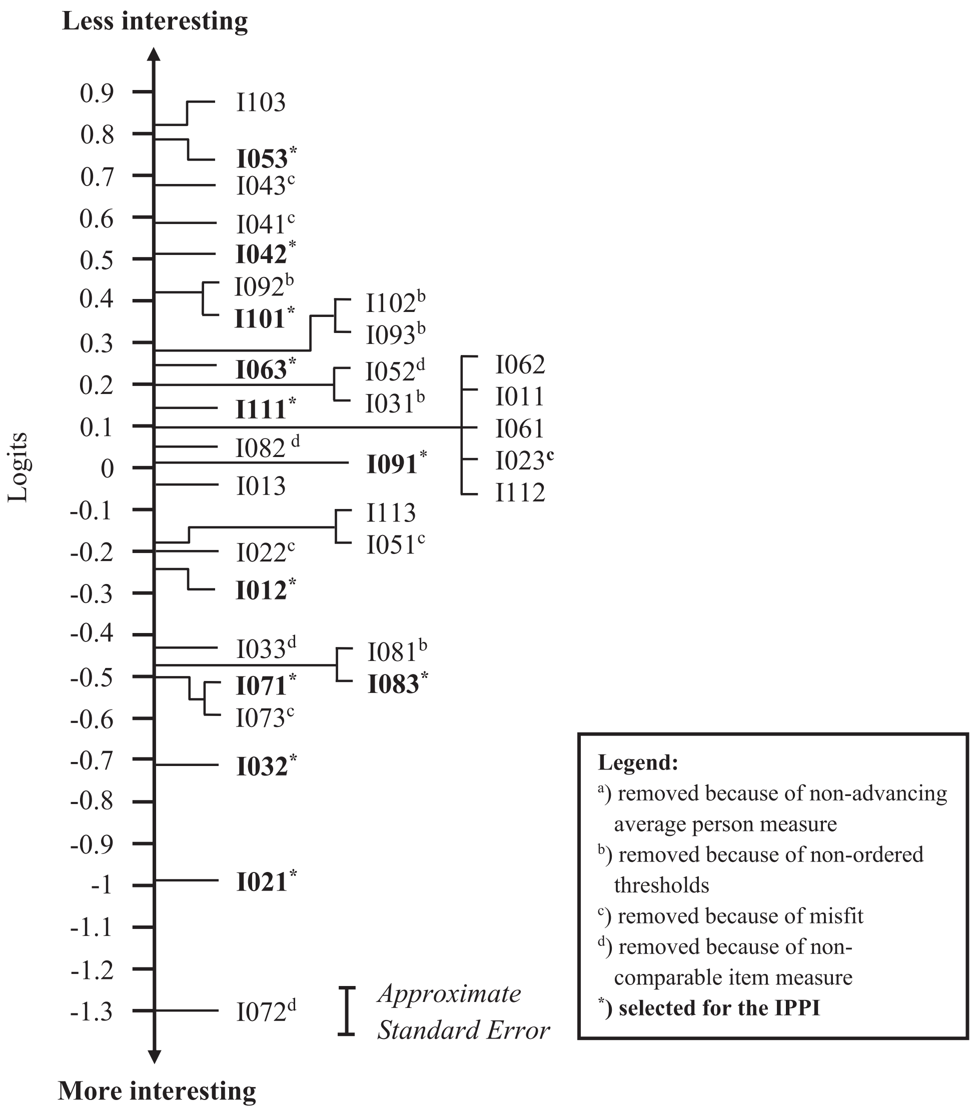

Figure 3. Wright Map of the initial item pool of the IPPI.

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.

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).

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 0.6–1.4 (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.

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.

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.

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.

Instrument functioning

Table 3 lists the Rasch item measures and statistics for the IPPI. The Outfit MNSQ values ranged from 0.79–1.36, which is within the acceptable range. All point measure correlation values were above the suggested 0.3 cut-off value (ranging from 0.59–0.71). 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 3. 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 4). 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 −1.02–0.94 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 −5.34–5.55 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 4 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.

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 4) and Table 3. 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

Table 4. 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.

	Item		Person		Outfit MNSQ	Infit MNSQ
Rasch analysis method	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*

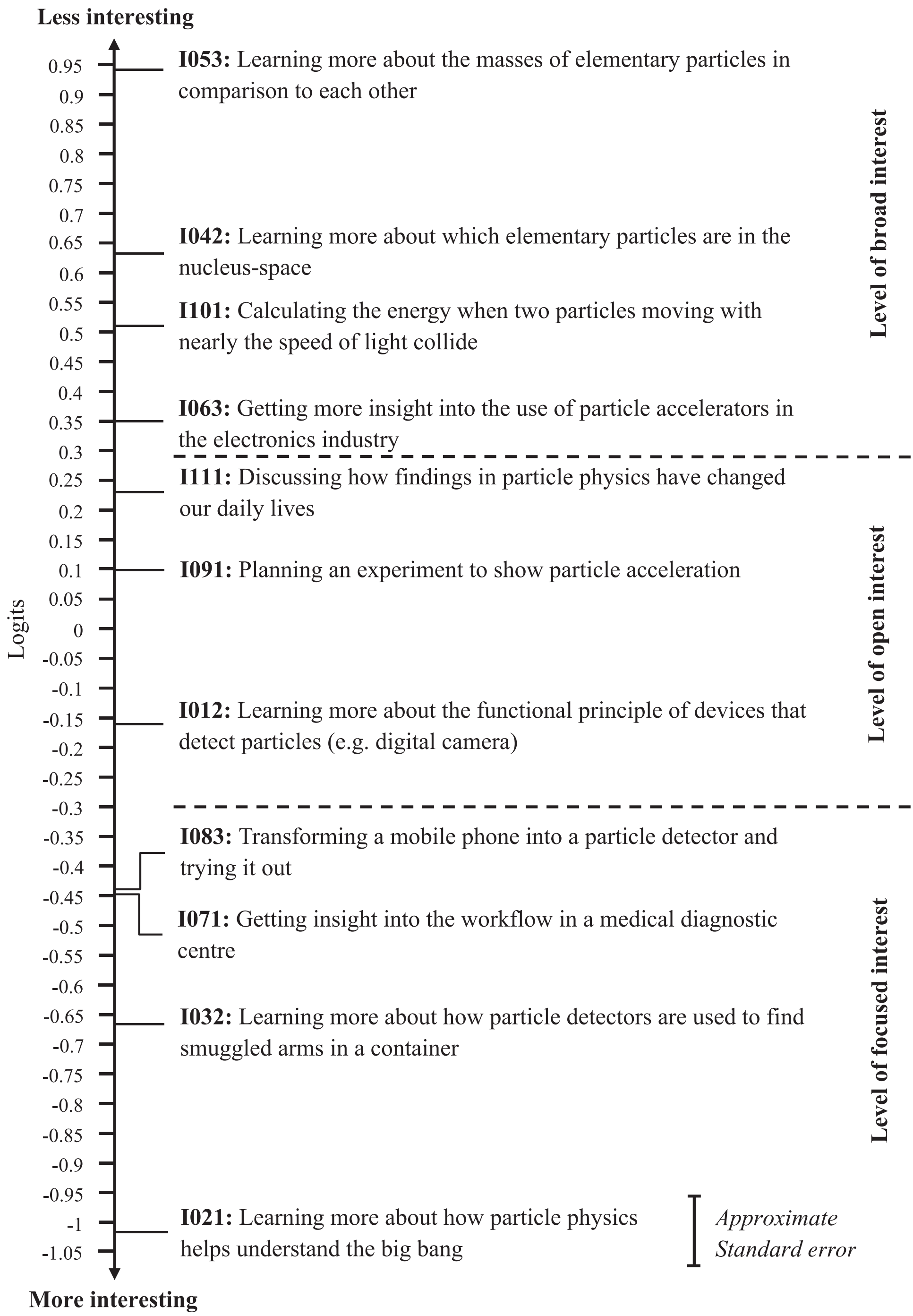

Figure 4. Wright Map of the IPPI.

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).

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.

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.

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 1–2, 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.

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 3). 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.

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 4). 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 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.

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 & 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 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.

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 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 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.

Note

1. 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 et al., 1998; Häußler et al., 1998; Rost et al., 1999; Sievers, 1999). In Table 1, 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.

Acknowledgements

We want to thank Dr. William Boone for reviewing a draft of this manuscript.

Data Availability Statement

The data that support the findings of this study are openly available in the Austrian Social Science Data Archive at https://doi.org/10.11587/OUDFJK.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Ethical statement

The study met the ethics/human subject requirements of the authors' institutions at the time the data was collected.

ORCID

Sarah Zoechling http://orcid.org/0000-0002-7069-5229
Martin Hopf http://orcid.org/0000-0003-3273-0000
Julia Woithe http://orcid.org/0000-0002-4772-2989
Sascha Schmeling http://orcid.org/0000-0003-2454-1699

References

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

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 & 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.

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.

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.
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.
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
de Barba, P. G., Kennedy, G. E., & Ainley, M. D. (2016). The role of students' motivation and participation in predicting performance in a MOOC. *Journal of Computer Assisted Learning*, *32*(3), 218–231. https://doi.org/10.1111/jcal.12130
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
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.
Ericsson, K. A., & Simon, H. A. (1993). *Protocol analysis: Verbal reports as data (revised edn.)*. Bradford Books/The MIT Press.
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. ISSN 1758-7956 (online).
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
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., 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.
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

Kalender, Z. Y., Marshman, E., Schunn, C. D., Nokes-Malach, T. J., & Singh, C. (2019). 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

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

Krapp, A. (2002). 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., & 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

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

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 & the Health Professions*, *41*(1), 25–43. https://doi.org/10.1177/0163278716676873

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.

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

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

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

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.

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. https://doi.org/10.1016/S0193-3973(01)00101-0

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

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. http://timssandpirls.bc.edu/timss2015/international-results/advanced/.

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

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. (2007). *PISA 2006 science competencies for tomorrow's world. Volume 1. Analysis. 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.

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

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

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., Sievers, K., Häußler, P., Hoffmann, L., & Langeheine, R. (1999). Struktur und Veränderung des Interesses an Physik bei Schülern der 6. Bis 10. Klassenstufe. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, *31*(1), 18–31. https://doi.org/10.1026//0049-8637.31.1.18

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.

Sievers, K. (1999). *Struktur und Veränderung von Physikinteressen bei Jugendlichen [Structure and change of physics interest among adolescents]*. Universität Kiel.

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.

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 (JESPAR)*, *12*(3), 243–270. https://doi.org/10.1080/10824660701601266

Vorholzer, A., von Aufschnaiter, C., & Kirschner, S. (2016). Entwicklung und Erprobung eines Tests zur Erfassung des Verständnisses experimenteller Denk- und Arbeitsweisen. *Zeitschrift für Didaktik der Naturwissenschaften*, *22*(1), 25–41. https://doi.org/10.1007/s40573-015-0039-3

Wright, B., & Stone, M. (1979). *Best test design*. MESA Press.

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

Appendices

Appendix A: Introductory Text

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 would you like to do the following? Please click "Next"!

Appendix B: Original German Item Wordings and English Paraphrases

#	Original German wording	English paraphrase
I011	Mehr darüber erfahren, wie ein Teilchenbeschleuniger funktioniert	Particle accelerator
I012	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	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	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	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	Mehr darüber erfahren, wie groß die Massen der Elementarteilchen im Vergleich zueinander sind	Masses of particles (quantitative)
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	Mehr Einblick erhalten, wie in der Elektronik-Industrie mit Teilchenbeschleunigern gearbeitet wird	Particle accelerators in the electronics industry
I071	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	Build a particle detector out of daily life objects (hands-on)
I082	Einen Elektromagneten bauen und damit die Bewegungsrichtung eines Teilchens verändern	Build an electromagnet to influence the direction of a particle (hands-on)
I083	Ein Handy in einen Teilchendetektor umbauen und ausprobieren	Transform a mobile phone into a particle detector (hands-on)
I091	Ein Experiment planen, um zu zeigen, wie Teilchen beschleunigt werden	Plan an experiment for particle acceleration (minds-on)

(*Continued*)

Continued.

#	Original German wording	English paraphrase
I092	Sich ein Experiment ausdenken, um zu zeigen, wie man den Aufbau eines Atoms erforschen kann	Plan 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	Plan an experiment to influence the direction of a particle (minds-on)
I101	Berechnen, wie groß die Energie beim Zusammenstoß zweier Teilchen ist, die sich mit nahezu Lichtgeschwindigkeit bewegen	Calculate the energy of a particle collision at the speed of light (minds-on)
I102	Berechnen, aus wie vielen Elementarteilchen ein menschliches Haar besteht	Calculate number of particles in human hair (minds-on)
I103	Die Massen verschiedener Elementarteilchen berechnen, weil man sie nicht einfach abwiegen kann	Calculate the mass of particles (minds-on)
I111	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)