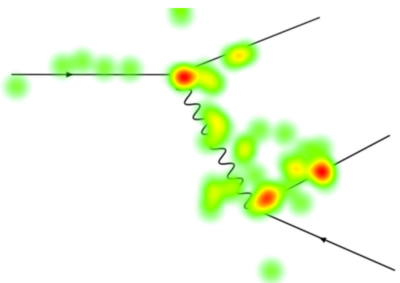
The Use of Feynman Diagrams in Physics Education: Opportunities, Challenges, Practices



DISSERTATION

for the award of the degree
"Doctor rerum naturalium" (Dr. rer. nat.)

of the GEORG-AUGUST-UNIVERSITÄT GÖTTINGEN

within the doctoral program *Physics*
of the Georg-August University School of Science (GAUSS)

submitted by
Merten Nikolay DAHLKEMPER

from ITZEHOE

Göttingen, 2024

Thesis Committee

Prof. Dr. Pascal KLEIN
Physik und ihre Didaktik, Fakultät für Physik, Universität Göttingen

Prof. Dr. Arnulf QUADT
II. Physikalisches Institut, Fakultät für Physik, Universität Göttingen

Prof. Dr. Andreas MÜLLER
Section de Physique, Faculté des Sciences, Université de Genève &
Institut universitaire de formation pour l'enseignement (IUFE)

Members of the Examination Board

Reviewer: Prof. Dr. Pascal KLEIN
Physik und ihre Didaktik, Fakultät für Physik, Universität Göttingen

Second Reviewer: Prof. Dr. Andreas MÜLLER
Section de Physique, Faculté des Sciences, Université de Genève &
Institut universitaire de formation pour l'enseignement (IUFE)

Further members of the Examination Board:

Prof. Dr. Arnulf QUADT
II. Physikalisches Institut, Fakultät für Physik, Universität Göttingen

Dr. Sascha Marc SCHMELING
Teacher and Student Programmes, CERN

Prof. Dr. Susanne SCHNEIDER
Physik und ihre Didaktik, Fakultät für Physik, Universität Göttingen

Prof. Dr. Kerstin STRECKER
Didaktik der Informatik, Fakultät für Mathematik und Informatik,
Universität Göttingen

Date of the oral examination: 14th of May 2024

"I think it is much more interesting to live not knowing than to have answers that might be wrong. [. . .] I don't feel frightened by not knowing things."

Richard P. Feynman[1]

[1]The quote is from the BBC documentary *The Pleasure of Finding Thing Out* (BBC, 1981, 48:25-48:58).

Abstract

The Use of Feynman Diagrams in Physics Education:
Opportunities, Challenges, Practices

This thesis explores the potential of Feynman diagrams as educational tools in upper secondary physics education, specifically for teaching particle physics concepts. It is set within the design-based research (DBR) framework, meaning learning materials incorporating Feynman diagrams were designed and evaluated while obtaining theoretical insights into learning processes.

Motivation Feynman diagrams are a widely used tool in particle physics and have, therefore, found their way into physics textbooks and popular culture. Students interested in particle physics will come across them. However, the complexity of these diagrams also raises concerns about their suitability as educational content.

On the theoretical side, this project addressed two key **research goals**:

1. Identifying the **opportunities and challenges** of using Feynman diagrams in particle physics education.
2. Developing **specific design principles** for designing and using learning materials that effectively leverage Feynman diagrams.

On the practical side, the outcome was an **online learning environment** that interested students can use to learn about particle physics.

Methods As mentioned in the beginning, the project followed the DBR framework. Interviews with four experts on particle physics education were conducted to investigate the opportunities and challenges of Feynman diagrams in physics education. Learning goals that can be achieved with Feynman diagrams were derived from these interviews and literature. A draft of the learning material was designed based on the learning goals and design principles from multimedia research. The material consisted of graphics and English-speaking textual explanations to convey concepts of particle physics with Feynman diagrams, which students typically took between 20 to 30 minutes to work through. This draft was evaluated in two rounds with 72 international students between 16 and 20 years at CERN and, after redesigning and creating a German version, in a third study with 33 students between 15 and 19

years in two German schools. The evaluation studies used eye-tracking to investigate students' visual strategies and cognitive processes. Think-aloud protocols and questionnaires about prior knowledge in particle physics, cognitive load, motivational factors, and conceptual understanding supplemented the eye-tracking data.

Results Four progressing educational **opportunities** could be derived from the expert interviews and relevant literature:

1. Feynman diagrams offer a clear and intuitive way to represent **charge conservation**, a crucial concept in physics, especially particle physics.
2. They can introduce the concept of **interaction particles**, a key concept of modern physics, contrasting it with the classical, non-local interaction view.
3. They help illustrate the superposition principle and the role of approximations in particle physics, unveiling the **quantum nature of particle physics** and, thereby, a more accurate picture of the field than usually painted in popular depictions.
4. By understanding how Feynman diagrams **connect theoretical and experimental particle physics**, students gain a deeper understanding of the nature of science.

The first and second eye-tracking studies revealed that students use different, diagram-dependent visual strategies to examine Feynman diagrams based on the depth of their processing. The eye-tracking data of the third study demonstrated that various types of diagrams have different needs of cognitive processing, that students can learn through examples of Feynman diagrams that visualise charge conservation to focus their attention on relevant parts, and that a focused and systematic viewing behaviour is beneficial to solve tasks with Feynman diagrams. Furthermore, responses to open comprehension questions in the third student study revealed students' difficulties with subject-specific terms and concepts (like "weak charge" or "interaction particle") but no inherently stable or widespread inadequate conceptions.

From these results, **challenges** that come with using the diagrams in education can be derived, which are classified into three categories:

1. Students have **difficulties with concepts** like charge conservation and interaction particles, requiring careful instruction and practice.
2. The **complexity of Feynman diagrams**, including non-sequential elements and arrows, is challenging for some students to decipher.
3. Using **technical terms in particle physics** confuses most students.

Implications Furthermore, certain **specific design principles** can be derived from the project to address these challenges:

1. Foster the **development of different strategies** (e.g., reading vertices forwards and backwards or identifying anti-particles) for examining Feynman diagrams to enhance their understanding.
2. Present the Feynman diagrams in a **well-defined progression**, starting from simpler diagrams and gradually increasing complexity.
3. **Introduce foundational terms and definitions** before using them in the explanations of diagrams to mitigate confusion.

Conclusion The results of this thesis demonstrate that Feynman diagrams can be valuable tools for teaching particle physics concepts in upper secondary education, provided the challenges are addressed with appropriate practices. However, further research should be done to explore the broader application of Feynman diagrams in science education.

Zusammenfassung

Die Verwendung von Feynman-Diagrammen in Lehr-/Lernkontexten:
Chancen, Herausforderungen, Praktiken

In dieser Arbeit wird dass Potential von Feynman-Diagrammen für das Physiklernen auf dem Niveau der Sekunderstufe II untersucht, insbesondere für die Vermittlung von Konzepten der Teilchenphysik. Sie ist im Rahmen der designbasierten Forschung (DBR) angesiedelt, was bedeutet, dass Lernmaterialien, die Feynman-Diagramme enthalten, entworfen und evaluiert wurden, während gleichzeitig theoretische Erkenntnisse über Lernprozesse gewonnen wurden.

Motivation Feynman-Diagramme sind ein weit verbreitetes Werkzeug in der Teilchenphysik und haben daher ihren Weg in die Physiklehrbücher und die Populärkultur gefunden. Jugendliche, die sich für Teilchenphysik interessieren, werden ihnen daher unweigerlich begegnen. Die Komplexität dieser Diagramme gibt jedoch auch Anlass zu Bedenken hinsichtlich ihrer Eignung als Lerninhalt.

Auf der theoretischen Seite wurden in diesem Projekt zwei zentrale Forschungsziele verfolgt:

1. Identifizierung der **Chancen und Herausforderungen** der Verwendung von Feynman-Diagrammen beim Lernen von Konzepten der Teilchenphysik.
2. Entwicklung von **spezifischen Gestaltungsprinzipien** zur Gestaltung und Verwendung von Lernmaterialien, die Feynman-Diagramme effektiv nutzen.

Auf der praktischen Ebene war das Ergebnis eine **Online-Lernumgebung**, die interessierte Schüler*innen nutzen können um über Teilchenphysik zu lernen.

Methoden Wie eingangs erwähnt, folgte das Projekt dem DBR-Ansatz. Es wurden Interviews mit vier Expert*innen für das Lehren und Lernen auf dem Gebiet der Teilchenphysik geführt, um die Chancen und Herausforderungen des Lernens mit Feynman-Diagrammen zu untersuchen. Die Lernziele, die mit Feynman-Diagrammen erreicht werden können, wurden aus diesen Interviews und der Literatur abgeleitet. Auf der Grundlage der Lernziele und Gestaltungsprinzipien aus der Multimediaforschung wurde ein Entwurf des Lernmaterials erstellt. Das Material bestand aus Grafiken und englischsprachigen Erklärungen in Textform zur Vermittlung von Konzepten der Teilchenphysik mit Feynman-Diagrammen, für deren Bearbeitung

die Schülerinnen und Schüler typischerweise zwischen 20 und 30 Minuten benötigten. Dieser Entwurf wurde in zwei Runden mit 72 internationalen Schüler*innen zwischen 16 und 20 Jahren am CERN und, nach Überarbeitung und Erstellung einer deutschen Version, in einer dritten Studie mit 33 Schüler*innen zwischen 15 und 19 Jahren an zwei deutschen Schulen evaluiert. In den Evaluationsstudien wurden mittels Eye-Tracking die visuellen Strategien und kognitiven Prozesse der Schüler*innen untersucht. Think-aloud-Protokolle und Fragebögen zu Vorwissen in Teilchenphysik, kognitiver Belastung, Motivationsfaktoren und konzeptionellem Verständnis wurden als Begleitdaten erhoben.

Ergebnisse Aus den Experteninterviews und der einschlägigen Literatur konnten vier aufeinander aufbauende didaktische Chancen abgeleitet werden:

1. Feynman-Diagramme bieten eine klare und intuitive Möglichkeit zur Darstellung der Ladungserhaltung, einem wichtigen Konzept in der Physik, insbesondere der Teilchenphysik.
2. Sie können das Konzept der Wechselwirkungsteilchen, einen Schlsselkonzept der modernen Physik, einführen und es mit der klassischen, nichtlokalen Wechselwirkungsansicht kontrastieren.
3. Sie helfen dabei, das Überlagerungsprinzip und die Rolle von Näherungen in der Teilchenphysik zu veranschaulichen, indem sie die Quantennatur der Teilchenphysik enthüllen und damit ein präziseres Bild des Feldes zeichnen als es üblicherweise in populären Darstellungen gezeichnet wird.
4. Durch das Verständnis, wie Feynman-Diagramme die theoretische und experimentelle Teilchenphysik miteinander verbinden, gewinnen die Schüler*innen ein tieferes Verständnis für die Natur der Naturwissenschaft.

Die erste und die zweite Eye-Tracking-Studie haben gezeigt, dass die Schüler*innen je nach Verarbeitungstiefe unterschiedliche, vom Diagramm abhängige visuelle Strategien zur Untersuchung von Feynman-Diagrammen verwenden. Die Eye-Tracking-Daten der dritten Studie zeigten, dass verschiedene Diagrammtypen unterschiedliche Anforderungen an die kognitive Verarbeitung stellen, dass Schüler*innen durch Beispiele von Feynman-Diagrammen lernen können, die Ladungserhaltung zu visualisieren, um ihre Aufmerksamkeit auf relevante Teile zu fokussieren, und dass ein fokussiertes und systematisches Betrachtungsverhalten von Vorteil ist, um Aufgaben mit Feynman-Diagrammen zu lösen. Darüber hinaus zeigten die Antworten auf offene Verständnisfragen in der dritten Studie, dass die Schüler*innen Schwierigkeiten mit fachspezifischen Begriffen und Konzepten (wie "schwache Ladung" oder "Wechselwirkungsteilchen") haben, aber keine inhärent stabilen oder weit verbreiteten inadäquaten Vorstellungen.

Aus diesen Ergebnissen lassen sich **Herausforderungen** ableiten, die mit der Verwendung der Diagramme im Lehr-/ Lernkontext einhergehen und in drei Kategorien eingeteilt werden:

1. Schüler*innen haben **Schwierigkeiten mit Konzepten** wie Ladungserhaltung und Wechselwirkungsteilchen, die eine sorgfältige Anleitung und Übung erfordern.
2. Die **Komplexität von Feynman-Diagrammen**, einschließlich nicht-sequenzieller Elemente und Pfeile, ist für einige Schüler*innen schwer zu bewältigen.
3. Die Verwendung von **Fachbegriffen in der Teilchenphysik** verwirrt die meisten Schüler*innen.

Implikationen Darüber hinaus lassen sich aus dem Projekt bestimmte **spezifische Gestaltungsprinzipien** ableiten, um diesen Herausforderungen zu begegnen:

1. Fördern Sie die Entwicklung verschiedener Strategien (z. B. das Vorwärts- und Rückwärtslesen von Vertices oder die Identifizierung von Antiteilchen im Diagramm), um die Feynman-Diagramme zu untersuchen und ihr Verständnis zu verbessern.
2. Präsentieren Sie die Diagramme in einer **klar definierten Abfolge**, beginnend mit einfacheren Diagrammen und allmählich ansteigender Komplexität.
3. **Führen Sie grundlegende Begriffe und Definitionen ein**, bevor sie bei der Erklärung von Diagrammen verwendet werden, um Verwirrung zu vermeiden.

Schlussfolgerung Die Ergebnisse dieser Arbeit zeigen, dass Feynman-Diagramme wertvolle Werkzeuge für den Unterricht von Konzepten der Teilchenphysik in der Sekundarstufe II sein können, vorausgesetzt, die Herausforderungen werden mit geeigneten Maßnahmen angegangen. Es sollten jedoch weitere Forschung durchgeführt werden, um die breitere Anwendung von Feynman-Diagrammen im naturwissenschaftlichen Unterricht zu untersuchen.

Acknowledgements

A PhD project is like climbing a mountain. The more I think about it, the more accurate it gets. When being on the foot of the mountain, it might seem pretty straightforward, maybe with a few difficulties laying ahead, but they are so far away that you don't care yet. But as soon as you start, you get aware that there are intricacies of the path you would have never foreseen. Sometimes, there are paths which do not seem to get you anywhere, and you have to walk back. There are long, boring stretches which have to be done, there are exhausting stretches, but which are also exciting. And then you stand on the top, you enjoy the view, but you also see that there are higher summits all around you. But the most important thing is: You should not climb a mountain alone. And I can call myself lucky to have so many people who joined me and supported me on this journey. I use these pages to thank them.

First of all, I am deeply grateful to my first university supervisor Pascal Klein, for taking the adventure of supervising me, as one of his first PhD students, remotely, for sharing his experience, for always having a sympathetic ear, and for always encouraging me. Just as important to the success of this PhD project was my second university supervisor, Andreas Müller, whom I'm deeply grateful for initiating the contact with CERN and, by that, making this journey possible in the first place, for fruitful discussions, for giving insightful feedback, and for deep insights in many fields of educational research. I thank my supervisor at CERN, Jeff Wiener, for sharing his experience, and for a lot of actionable feedback. I truly appreciate that you've been always open for discussions. I also want to thank Arnulf Quadt, who gave valuable feedback during the meetings of the thesis advisory committee. And of course, I am eternally grateful to Sascha Schmeling, who does everything he can to provide me and his whole team with the best possible working conditions. You are a great manager. I truly admire your skills.

A PhD journey would not be possible without the research groups one is part of. I was very lucky to be both part of the "Teacher and Student Programmes" section of CERN and the "Physics and its Didactics" group at the University of Göttingen.

I am incredibly grateful that I could be part of the most amazing team at CERN. It is a true pleasure to work with such wonderful, inspiring people. My thanks go at first at first those who took at least parts of the PhD journey together with me: First, I joined Anja Kranjc Horvat and Sarah Zoechling on their journey. They taught me things, both practically and life-advisory, which turned out super useful.

Later on, Yiota Chatzidaki, Ruadh Duggan, and Milena Vujanovic joined me on the journey and I am so grateful I could share my journey with them. But the group does not only consist of other PhD students. I want to thank Margherita, Niklas, Gernot, and Tobias for sharing their offices with me throughout the years and all the others who were lunch company, gave valuable feedback at PER seminars and private discussions and eye-tracking sessions, or were simply climbing or running companions and so much more: Thank you, Julia, Guillaume, Charlotte, Patrick, Anastasia. All of you made my experience at CERN a very special one and I am deeply grateful for your help, let it be through mental or research support.

I was lucky to spend four full months in my group in Göttingen but also met them regularly online and at multiple occasions on conferences and short trips to Göttingen. I am grateful for them to always welcome me and making me feel like being part of the group despite the distance. My thanks go to my fellow PhD students Larissa Hahn and Simon Lahme for fruitful discussions, valuable exchange of feedback, and especially to Simon for endless fun discussions about trains (well, not only trains, but in the end we always circle back to them). But I also want to thank all the other members of the group who gave valuable feedback during the group seminars.

But there were also many people outside of my working groups who helped me a lot with my research. I want to thank Michael Kobel, Philipp Lindenau, Oliver Passon, Jürgen Tolksdorf, and and Luis Castillo for fruitful discussions about Feynman diagrams and particle physics education As well as all the experts from the expert interviews, all the students and teachers from all around the world who came to CERN or let me come to their schools and do my studies with them, all the people who attended my conference talks and asked questions or gave fruitful remarks.

I also want to thank the German Federal Ministry of Education and Research who funded this PhD via the Gentner scholarship (grant no. 13E18CHA).

And last but certainly not least, a PhD journey is not decoupled from the rest of the life. Many of the above-mentioned people have become good friends over the years, which I am truly grateful for. I am also truly grateful to Moni, Kim, My, Anna, Lorenz, and Felix. There is a saying that "Friends are like stars. You may not always see them but you know they're there." You are my stars.

Und schließlich gilt mein Dank meinen Eltern. Ihr habt mir immer gesagt, wie sehr ihr an mich glaubt. Es ist unglaublich befreiend zu wissen, dass ihr immer für mich da seid. Dafür bin ich euch zutiefst dankbar. Ich glaube, ich sage euch das viel zu selten. Danke für alles!

Mais il y a une personne sans laquelle je n'en serais certainement pas là. Les mots me manquent pour exprimer à quel point je suis reconnaissant de t'avoir rencontrée pendant mon séjour au CERN, Oriane. Tu m'as toujours soutenu, tu m'as toujours dit que j'allais réussir, tu es la meilleure personne du monde. Je t'aime.

Contents

List of Figures

List of Tables

List of Abbreviations

ANOVA	**An**alysis **o**f **Va**riance
AOI	**A**rea **o**f **I**nterest
cf.	**c**ompare **f**or
CATLM	**C**ognitive-**A**ffective **T**heory of **L**earning with **M**ultimedia
CERN	European Organisation for Nuclear Research
CLT	**C**ognitive **L**oad **T**heory
CTML	**C**ognitive **Theory** of **M**ultimedia **L**earning
DBR	**D**esign-**B**ased **R**esearch
DeFT	**De**sign, **F**unctions, **T**asks
EL	**E**xtraneous Cognitive **L**oad
ET	**E**ye **T**racking
e.g.	for example
FD	**F**eynman **d**iagram
FTR	**F**ixation/**t**ransition **r**atio
GL	**G**ermane Cognitive **L**oad
GUT	**G**rand **u**nified **t**heory
HSSIP	**H**igh-**S**chool **S**tudent **I**nternship **P**rogramme
i.e.	that means
IL	**I**ntrinsic Cognitive **L**oad
LHC	**L**arge **H**adron **C**ollider
MER	**M**odel of **E**ducational **R**econstruction
MR	**M**ultiple **R**epresentations
NoS	**N**ature **o**f **S**cience
PER	**P**hysics **E**ducation **R**esearch
QCD	**Q**uantum **c**hromo**d**ynamics
QED	**Q**uantum **e**lectro**d**ynamics
RBAI	**R**esearch-**b**ased **a**ssessment **i**nstrument
SMPP	**S**tandard **M**odel of **P**article **P**hysics
$\mathbf{Sp\bar{p}S}$	**S**uper **P**roton-Anti**p**roton **S**ynchrotron
SSR	**S**ocial **S**emiotic **R**esources
s.t.	such that

List of Symbols

α	coefficient for internal consistency (Cronbach's alpha)
Γ	Decay rate
γ	Photon
δ	Ferguson's delta
μ^-	Muon
ν	Neutrino
$\overline{\nu}$	Anti-neutrino
ν_e	Electron neutrino
ν_μ	Muon neutrino
σ	Cross-section or standard deviation or standard error (context-dependent)
τ	Average lifetime
$\mathcal{A}$	Probability amplitude
D	Discriminatory index
d	effect size (*Cohen's d*)
df	degrees of freedom
e^-	Electron
e^+	Positron
F	F-statistic
H_s	Stationary entropy
H_t	Transition entropy
KR-20	Kuder-Richardson coefficient
M	Mean
P_{ij}	probability for a transition from AOI i to AOI j
p	test score or p-value (context-dependent)
R	Pearson's correlation coefficient
r_{pbi}	Point biserial coefficient
SD	Standard deviation
t	t-statistic or time (context-dependent)
W^+	W-plus boson
W^-	W-minus boson
Z^0	Z boson

Chapter 1

Introduction

When I was in 12th grade, I read the book *QED* by Richard Feynman[1]. I was so fascinated by the subject that I convinced my physics teacher to hand over one lesson from his physics course, in which I could try and teach parts of that approach to my fellow students. I quickly learned that I was not a Feynman: No one in the classroom understood a thing (including myself). That was probably not my first encounter with Feynman diagrams, but the first that I remember – and indeed, my first trial of teaching them.

Now, a few years down the road, in which I have learned many things – both about teaching and about Feynman diagrams, I am trying it again. This time, I am better prepared and want to know exactly what goes wrong. Back in the day, I could see it in the gazes of my fellow students that I had lost them. I am using more sophisticated methods in my current approach, but I am still analysing the students' gaze. On the title page, you see what this gaze on a Feynman diagram might look like. This is a connection between two very particular types of representations, one widely used in particle physics (as I will show in Section 2.1), one more and more used in physics education research (as I will show in Section 4.4). The two representations merged in that picture also share the characteristic that both of them are a source of overinterpretation (cf. Sections 2.1.3 and 5.2.4). This connection between Feynman diagrams and eye-tracking lies at the heart of this thesis. I am inviting you, dear reader of this thesis, to explore my second trial of teaching Feynman diagrams.

Feynman diagrams are only one of many examples where scientists use representations that seem accessible at first glance, maybe even ostensibly depicting how something "really looks", but get ever more complex the closer you examine them. Other examples include chemical formulas, quantum mechanical energy diagrams, and schematic depictions of proteins or antibodies. They all certainly have in common that, for the experienced eye, these representations are very valuable. In contrast, the inexperienced eye could easily misled by them into getting inadequate

[1] The book (Feynman, 1985) is based on a public lecture series Feynman gave, in which he explained his method of the path integral to do calculations in quantum electrodynamics. Feynman uses this method to motivate Feynman diagrams, which are not a central topic in the book.

ideas about the concepts which underlie these depictions. However, these depictions also provide a great chance to spark curiosity for the science topic as they are so ubiquitous in popular depictions of science. Therefore, my question is whether Feynman diagrams can also be successfully used in education, which can be seen as the link between popular depictions of science and the discipline.

This question became the guiding question of my PhD project. The project took the form of a design-based research (DBR) project (see Section 4.2) in which I developed learning material to teach concepts of particle physics to upper secondary students using Feynman diagrams. To do so, I searched for opportunities that Feynman diagrams give, investigated possible challenges they pose to learning, and suggested practices that can be used to teach with Feynman diagrams.

The structure of this thesis is as follows:

In Chapter 2, I give an overview of where Feynman diagrams come from (Section 2.1.1), how they work (Section 2.1.2), and how they are interpreted (Section 2.1.3). Following that, I give an overview of how particle physics as a topic is taught in school (Section 2.2) and provide a literature review of teaching approaches to Feynman diagrams (Section 2.2.2). The chapter ends with a brief review of the topic of the Nature of Science (Section 2.2.3), as this is a crucial subject matter at several stages of this work.

In Chapter 3, a concise but crucial chapter, I state and motivate the research goals derived from the guiding question already stated in this introduction.

Chapter 4 then reviews the most important theories that are used within this project, among them the model of educational reconstruction (Section 4.1), the DBR framework (Section 4.2) as well as the several approaches to learning with multiple representations (Section 4.3). In this chapter, I also review previous research with eye tracking in physics education research (Section 4.4) as this will be the essential research method used in this PhD project.

Chapter 5 then introduces the technicalities of the methods used throughout this project. It is made of three sections. In Section 5.1, I first introduce a detailed overview of how the DBR stages are addressed in this thesis and give an overview of the design principles used to design the learning material. Section 5.2 then introduces the technicalities of the eye tracking method, and Section 5.3 introduces the questionnaires I used throughout the studies within this thesis, before Section 5.4 introduces the statistical methods used throughout this thesis.

The heart of the thesis is Chapter 6. In this chapter, I present the various studies I conducted throughout the PhD project, their study-specific procedures, their results, and their study-specific discussion. General methods and overarching discussion points are addressed in the respective chapters of this thesis. In Section 6.1, I present an expert interview study I conducted to find out the opinion of experts in particle physics education about opportunities and challenges when teaching with Feynman diagrams. This study has been published and is therefore presented as an individual paper. The outcomes of that study, namely learning goals for teaching about particle physics with Feynman diagrams, are then presented in Section 6.2. In Section 6.3, I present another study that has already been published, namely a pilot study in which I explored the eye tracking method with Feynman diagrams for the first time. Section 6.4 then presents an exploratory study with the first draft of the learning material with high school students, in which I explored students' gaze patterns on Feynman diagrams. Section 6.5 then presents a direct follow-up on that study with an update of that learning material in which I investigated students' visual strategies when examining Feynman diagrams. The chapter concludes in Section 6.6 with the main study of this project, the third student study, in which I used a complete redesign of the learning material with high school students from two different German schools and investigated students' cognitive processes and conceptual difficulties when learning with Feynman diagrams.

In Chapter 7, I discuss the results from these studies along with the research goals formulated in Chapter 3. In Section 7.1, I discuss the opportunities and challenges of teaching particle physics with Feynman diagrams, whereas in Section 7.2, I suggest practices which should support these opportunities and overcome the challenges. In Section 7.3, I then discuss how the limitations of this project suggest directions for further research.

In Chapter 8, I make suggestions for a future redesign of the learning material and give ideas of how Feynman diagrams could also be used beyond the scope of the current project.

Chapter 9 then concludes the thesis by revisiting the opportunities, challenges and practices and answering whether Feynman diagrams *should* be used in particle physics education.

Of course, my failure to explain Feynman diagrams to my fellow students more than ten years ago is not the only reason I chose this topic. Another one is that Feynman diagrams are excellent examples of how science can be uncertain yet be the best knowledge generation system. This is not a paradox; the uncertainty makes science so robust. Feynman meant this when saying that he doesn't "feel frightened by not knowing things". It was in this spirit that I wrote this thesis. I will not provide definitive answers to the questions I pose; I will give tentative, evidence-based answers and point out directions for further research.

Chapter 2

Particle Physics

In this chapter, I will introduce the most critical concepts from particle physics used within this thesis and give an overview of how particle physics, especially Feynman diagrams, is addressed in school physics.

2.1 Feynman diagrams

Feynman diagrams are among the most famous representations in physics. In this section, I will briefly explain how they came to be and how they are used in physics. Since this is not a thesis in particle physics, the account is not rigorous at every point. Instead, it should put the significance into perspective and give an idea of its use in particle physics.

2.1.1 Origin of Feynman diagrams

Feynman diagrams play a vital part in the development of quantum field theory. This section shall briefly overview the most critical steps in this development to put these graphical tools into perspective. Since this work is not a historical one, I refer the interested reader to more comprehensive works about the origins and development of quantum field theory for further information (Cao, 2019; Darrigol, 2019; Kuhlmann, 2023; Schweber, 1994). Furthermore, I refer to the thorough review by Kaiser (2005) for all the intricacies of how Feynman diagrams have been used since their invention.

By the mid of the 1920s, a coherent theory to describe the movement of subatomic particles had been developed known as quantum mechanics, formulated as wave mechanics by Erwin Schrödinger and matrix mechanics by Werner Heisenberg, which soon proved to be two equivalent formulations. Applying quantum mechanics to electrodynamics, which had already been formulated as a field theory, was the next step. This would be a formulation of quantum mechanics consistent with special relativity. Paul Dirac coined the term *quantum electrodynamics* (QED) for this research program (Dirac, 1927). There have been essentially two approaches to this quest:

One, in the tradition of Schrödinger, de Broglie, and Jordan, put fields in the centre of their theory where particles were excitations of these fields (Heisenberg & Pauli, 1929; Jordan & Pauli, 1928). Another approach, put forward mainly by Dirac, puts particles in the centre of the theory in his so-called "hole-theoretical" approach (Dirac, 1930).

Both approaches were quite successful and led to many predictions in the field of quantum electrodynamics, such as the theory of the beta decay and the theory of nuclear forces which are mediated by "force carriers" (Yukawa, 1935) evolved from the field view. In contrast, correct computations of particle processes, such as electron-positron pair production, bremsstrahlung or Compton scattering, could be made using Dirac's hole-theoretic approach, i.e., the particle picture.

Despite these successes, both approaches still had conceptual problems as higher-order terms yielded infinities in the calculations as stated, among others, by Oppenheimer (1930). Several ways were tried to eliminate or circumvent these infinities, the most promising one being the technique of renormalisation, which describes the redefinition of parameters such as the charge of the electron. An often overlooked contribution to renormalisation theory was made by Swiss physicist Ernst Stueckelberg in Geneva (Lacki, 2017). By the time of 1948, two approaches to the formulation of a renormalised theory of QED were completed: one independently from the other by Tomonaga (1943, 1946) and Schwinger (1948, 1949a, 1949b) who followed the field theoretical approach and one by Feynman (1949a, 1949b, 1950) who was in line with the particle picture. All three received the Nobel Prize in Physics in 1965 for this work (Nobel Prize Outreach AB, 2024a). Feynman and Schwinger presented their results at the same conference in Pocono (Pennsylvania), which was the second of a series of three meetings to discuss recent breakthroughs in quantum field theory (Kaiser, 2005, p. 43 ff.), in the spring of 1948. However, they did not fully understand each other's approach. Still, they were convinced of their respective validity as the results of their approaches, for example, in calculating the Lamb shift, matched up (Schweber, 1994, p. 444 f.) It was subsequently proven by Dyson (1949a) that both approaches are equivalent.

One particularity of Feynman's approach was using a "space-time view", in which he used graphical tools today known as Feynman diagrams. However, Feynman's presentation at the Pocono meeting was described as rushed and not accompanied by a clear set of rules on how to apply these tools to the day-to-day work of theoretical physicists. Therefore, at that point, it was unclear to most participants how to use these new graphical tools in their daily work. It was not before Dyson (1949b) compiled a comprehensive set of rules for the use of the diagrams for calculations that they became more and more widespread for calculations, first just for QED processes (i.e., processes that are only intermediated by the electromagnetic interaction).[1] But

[1] Dyson's papers on the diagrams were published even before Feynman's and long before Feynman published his rigorous derivation of the diagrammatic approach (Feynman, 1950).

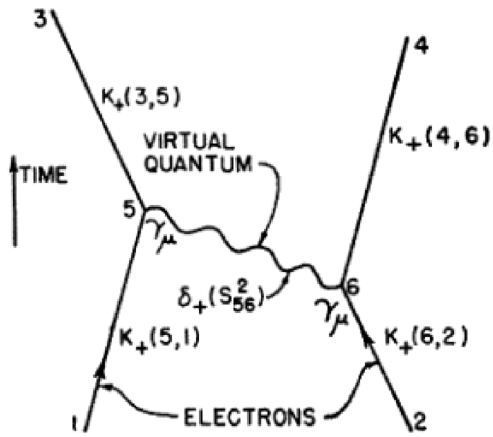


FIGURE 2.1: One of the first published Feynman diagrams as printed in Feynman (1949a). The diagram depicts one contribution to calculating an electron-electron scattering process mediated by a photon.

with the success of QED, other quantum field theories were modelled using a similar framework. These theories were the quantum chromodynamic (QCD), which explains the strong interaction, i.e., processes between quarks mediated by gluons, and the theory of the weak interaction, which later was merged with QED to become the electroweak theory. The latter theory described processes involving the weak interaction, mostly particle transformations[2] mediated by the W^+, W^-, and Z^0 bosons. Nowadays, it is hard to think about particle physics without Feynman diagrams. They are the standard way to compute properties of particle processes, which are needed to interpret data obtained from particle detectors.

Figure 2.1 shows one of the first diagrams printed in Feynman's paper (Feynman, 1949a). This diagram shows one contribution to the calculation of a process where two electrons interact with each other. The terms at the lines denote mathematical terms that need to be multiplied and integrated in a way that is presented in the next section.

2.1.2 Working mechanism of Feynman diagrams

A technical description of how Feynman diagrams work can be found in any particle physics textbook (e.g., Griffiths (2008)) or as an overview, for example, in Harlander (2021). The following account is mainly based on these two sources. I will not go into the technical details but rather give an idea of what Feynman diagrams are and why they are important in particle physics.

In quantum mechanics, the deterministic view known from classical physics is replaced by a probabilistic view. This means that every process happens with a certain

[2]In this work, the term "particle transformation" is usually used rather than the more common term "particle decay" to emphasise the fact that the products of such a process generally have not been a part of the initial particles. On the contrary, many particle transformations discussed in this work are transformations of one elementary particle, i.e., particles without a substructure, into others.

probability. "Process" is meant here in a comprehensive sense and can mean, for example, the movement of a particle, the interaction of one particle with another, or the transformation of a particle into another. The term "particle" is used within this work for any subatomic particles, which are classified within this work into matter particles[3] (such as protons, neutrons, electrons), antiparticles (in general, each matter particle has an antiparticle; the antiparticle of the electron is called positron, for example), and interaction particles (such as photons). In particular, two (probabilistic) quantities are interesting for particle physicists. One is the *cross-section* σ, which is an effective quantity of an area's size for a scattering process between two particles. The other one is the *decay rate* Γ, which measures how many particles of a given species transform[4] per unit time. The inverse of the latter quantity would be the *average lifetime* τ of a particle.

I won't give a detailed introduction to how exactly the calculations are carried out, but sketch the most important steps in following Griffiths (2008, chapter 6). In principle, the calculation consists of two steps: the calculation of the square of the absolute value of the probability amplitude $|\mathcal{A}|^2$ and the integration over the phase space of final states, i.e., all possible momentum configurations in which the process can happen. As an example, the cross-section of a scattering process with two particles with four-momenta[5] p_1 and p_2 in the initial state and two particles with four-momenta p_3 and p_4 in the final state looks like

$$\sigma \sim \int |\mathcal{A}|^2 \delta^4(p_1 + p_2 - p_3 - p_4) \times \delta(p_3^2 - m_3^2c^2)\delta(p_4^2 - m_4^2c^2)\theta\left(\frac{E_3}{c}\right)\theta\left(\frac{E_4}{c}\right) \mathrm{d}^4p_3\mathrm{d}^4p_4. \tag{2.1}$$

A more general form of this formula is called "Fermi's golden rule", and its rigorous derivation would go far beyond the scope of this work. The δ distributions ensure momentum and energy conservation (or, in short, four-momentum conservation): The factor $\delta^4(p_1 + p_2 - p_3 - p_4)$ vanishes whenever the sum of the four-momenta of the particles in the final state is not equal to the sum of the particles in the initial state and the factors $\delta(p_j^2 - m_j^2c^2)$ (with $j = 3, 4$) vanish whenever $p_j^2 = E_j^2 - \vec{p}_j^{\,2}c^2 \neq m^2c^4$, i.e., when energy conservation does not hold[6]. The Heaviside functions $\theta(E_j/c)$ ensure that the energies of the outgoing particles are always positive. This integral

[3] The term "matter particle" is primarily used to distinguish fermions from the gauge bosons, not to say that these particles "make up matter". Some particles do not fit into these three categories, like the Higgs boson. This is a deliberate choice to simplify the categories of particles.

[4] or decay, see above

[5] A four-momentum p^μ combines the energy and momentum of a particle into a four-dimensional vector whose metric is invariant under Lorentz transformations. The 0^{th} (or "temporal") component is E/c with the relativistic energy $E = \gamma mc^2$ with the Lorentz factor $\gamma = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}$ and the three "spatial" components are $\gamma m\vec{v}$, which are notated in the following as $\vec{p}$. Energy and momentum conservation are then encoded by the definition of the scalar product of a four-momentum with itself $p_\mu p^\nu = g_{\mu\nu}p^\mu p^\nu = \frac{E^2}{c^2} - \vec{p}^2 = m^2c^2$ (with the metric $g_{\mu\nu} = \mathrm{diag}(1, -1, -1, -1)$) being an invariant.

[6] or, to be more precise, when the particles "do not lie on the mass shell", we will see a bit later what that means.

can be easily calculated (the precise derivation is skipped here) and is solved as

$$\sigma \sim \frac{|\mathcal{A}|^2 |\vec{p}_f|}{(E_1 + E_2)^2 |\vec{p}_i|}, \tag{2.2}$$

where $\vec{p}_i = \vec{p}_1 = \vec{p}_2$ and $\vec{p}_f = \vec{p}_3 = \vec{p}_4$ because the problem is taken in the center-of-mass frame.

The only thing left is calculating the probability amplitude $\mathcal{A}$. This is done using Feynman diagrams. There are two ways how to motivate this approach. The rather historical one which follows Feynman's motivation is sketched out, for example, in Harlander (2021) and briefly introduced here. Feynman introduced the so-called path integral to calculate the probability amplitude of a particle process. This integral sums up the wave function of every path a particle can take from a given initial state to a particular final state.

This integral turns out to be solvable only when particles don't interact with each other, i.e., in the "free theory". This would only describe an undisturbed movement of a particle. However, any other particle process involves interactions of particles with each other. If an interaction between the particles is considered, the integral can no longer be solved exactly. Instead, physicists can systematically account for the effects of interactions, constructing a perturbation series that accounts for progressively higher-order corrections to the calculations. Each term in this perturbation series can then be translated into a diagram representing the interactions occurring in the particle process. This translation is possible in both directions. Therefore, drawing a set of diagrams and translating them into mathematical expressions according to a specific set of rules is also possible, which can then be used to organise calculations.

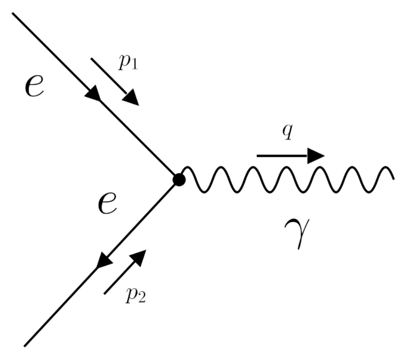


FIGURE 2.2: A fundamental vertex of QED. The line with an arrow to the right stands for an electron, the line with an arrow to the left for a positron and the wavy line for a photon. The notations for electron and positron are both e because the arrow already denotes the antiparticle in the negative time direction. The additional arrows denoting the momenta are usually added when using the diagram for calculations.

This set of rules is called *Feynman rules*. These rules dictate how each diagram element contributes to the probability amplitude $\mathcal{A}$. There is a fixed set of possible so-called *vertices*, which are points where three lines meet [7]. The lines can represent very different particles. In general, a line with an arrow to the right represents a matter particle, such as an electron, and a line with an arrow to the left represents an

[7] This is always valid for quantum electrodynamics, but in other quantum field theories like quantum chromodynamics for the strong interaction or the theory of the weak interaction, vertices exist where four lines meet. Such cases are not further considered in this work.

anti-particle[8] such as a positron, and a wavy line represents an interaction particle, such as a photon[9]. Such a vertex is shown in Figure 2.2. This vertex can be turned in all directions, representing either an electron and a positron annihilating and transforming into a photon (such as shown in the figure), an electron which emits a photon, an electron which absorbs a photon, a positron which emits a photon, a positron which absorbs a photon, or a photon which transforms into an electron-positron pair.

The precise Feynman rules for how to transform a diagram into the probability amplitude $\mathcal{A}$ differ from theory to theory (i.e. whether it is a diagram of QED, QCD, electroweak theory), but in a simplified version, i.e. a "toy theory" for particles without spin, they are the following (Griffiths, 2008, p. 213 f.):

Any vertex contributes a factor

$$-ig \tag{2.3}$$

which is determined by the strength of the respective interaction at that vertex. In the case of the electromagnetic interaction, the factor is $g_e = \sqrt{4\pi\alpha}$ with the fine structure constant $\alpha \approx \frac{1}{137}$. Each internal line, i.e., a line which connects two vertices and hence is not present in the initial or final state, contributes a so-called *propagator*

$$\frac{i}{q^2 - m^2c^2} \tag{2.4}$$

where q is the four-momentum and m is the particle's mass represented by the internal line. In the next step, the conservation of four-momentum is accounted for by introducing a delta distribution for every vertex in analogy to equation 2.1, for example, as

$$\delta^4(p_1 + p_2 - q) \tag{2.5}$$

if, in the example of Figure 2.2, the electron and positron are assigned incoming momenta p_1 and p_2 and the photon an outgoing momentum q (incoming and outgoing in this case referring to the vertex in question). Finally, the expression has to be integrated over each of the internal momenta d^4q. In detail, the rules are a bit more complicated. In particular, particles described in the Standard Model of Particle Physics have spin (fermions, like electrons, muons, quarks and other "matter particles", have half-integer spin, while bosons, like photons and other interaction particles, have whole-integer spin)[10] and are distinguished into particles and antiparticles. Particles and antiparticles are only discriminated in external lines, while in internal lines, they are treated equally.

[8] The idea that antiparticles can be represented by particles that go backwards in time was first mentioned by Stückelberg (1942) long before Feynman diagrams were introduced.

[9] There are also curly lines representing gluons (which are also interaction particles) and pointed or dashed lines representing scalar bosons, such as the Higgs boson. These are not used in this thesis.

[10] To be more precise, fermions have a half-integer and bosons an integer spin. The Higgs boson, called a "scalar boson", for example, has spin 0. But the Higgs boson brings many more intricacies that go far beyond this work's scope.

It is interesting to note that, in contrast to the outer lines where $p^2 = m^2c^2$ was ensured (Equation 2.1), the propagator factor 2.4 diverges at $q^2 = m^2c^2$. Therefore, particles represented by internal lines are often called being "off their mass-shell". These particles are also called *virtual particles*. Some authors state that energy conservation does not hold for internal lines and motivate this with Heisenberg's uncertainty principle $\Delta E \Delta t \geq \frac{\hbar}{2}$ stating that if the time the virtual particle exists is small enough, the uncertainty in energy can be arbitrarily high, s.t. energy conservation can be violated (e.g., Jones (2002)). However, this explanation is inadequate since energy and momentum are conserved as enforced by Equation 2.5. A more accurate explanation is that the mass of a virtual particle is not sharply defined. They are rather virtual states with a "resonance" at the mass of the corresponding real particle. From Equation 2.4, it is visible that the probability amplitude gets higher when the energy-momentum relation – which is clearly defined from energy and momentum conservation – comes closer to the rest mass of the corresponding real particle, i.e., when $q^2 - m^2c^2 \rightarrow 0$. This is reflected by the fact that processes in which the "mass" of the virtual interaction particle is far away from the "actual mass" of that interaction particle have a low probability.

As stated before, the Feynman diagram approach is a perturbative series approach. Therefore, calculating an amplitude $\mathcal{A}$ does not end with calculating the contribution of one diagram. Instead, all diagrams contributing to a particular process must be drawn, and their respective contributions must be summarised. That sum is then squared to obtain $|\mathcal{A}|^2$, which can then be inserted into equation 2.2 to calculate the cross-section of a process, for example. However, there are two problems with this process:

For every process, infinitely many diagrams can be drawn. For example, Figure 2.3 shows four diagrams that all contribute to the same process where an electron and a positron transform into a muon and an anti-muon. The top left diagram shows the dominant contribution for low energies. However, since the Z^0 boson, which is the interaction particle in the top right diagram, has a non-zero mass, the contribution of that diagram is dominant if the energy of the initial particles gets close to the rest energy of the Z^0, since the corresponding propagator term (equation 2.4) diverges near $q^2 = m_{Z^0}^2c^2$. This is further discussed in Section 6.2.4. The two bottom diagrams show higher-order terms of that process. This means these diagrams have more than two vertices, present in the simplest possible diagram that can be drawn for the process. Similarly, for any given diagram, more diagrams can be drawn with even more vertices. However, since each vertex contributes a vertex factor (equation 2.3), which is less than one, diagrams of higher order contribute less to the final amplitude than those of lower order. This is the principle of a perturbation series.

The second problem is more concerning. When calculating the contributions of diagrams of higher order, the integrals do not converge. Instead, they give infinities. This issue is what had blocked the development of quantum electrodynamics since

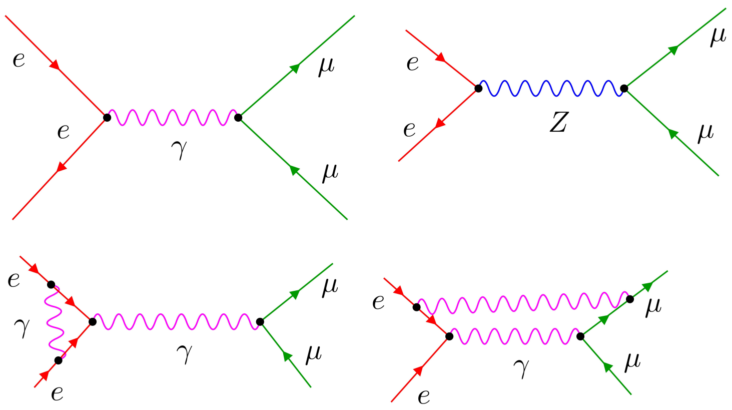


FIGURE 2.3: Example diagrams for the process $e^+e^- \to \mu^+\mu^-$.

its introduction at the end of the 1920s (see above). It took physicists almost 20 years to find a workable solution. The solution involved a so-called *renormalisation* of the masses and charges involved in the calculation. Simply put, the idea is that the "bare" masses and charges of the particles are shielded by an infinite amount of virtual particles surrounding the original particles. This results in infinite corrections to the bare values, s.t., the measured quantities are finite. These corrections depend on the energy at which the process takes place, which means that the strength of interactions depends on the energy scale, which is called "running coupling constants". The procedure of renormalisation has been met with scepticism by physicists, first and foremost by those who played significant roles in inventing it (Dirac, 1963; Feynman, 1985; van Kampen & Meijer, 2000), but is widely accepted nowadays, also since it has proven its usefulness in other fields. The idea of running couplings has opened the possibility of a "Grand Unified Theory" (GUT) in which all coupling constants have the same strength at a particular energy scale. However, this episode is an excellent example of the fact that science still has open questions. Undoubtedly, Feynman was – among many other things – referring to this particular issue in his quote, which is printed at the beginning of this thesis.

2.1.3 Interpretation of Feynman diagrams

There is an ongoing debate on whether Feynman diagrams are merely calculational tools or whether they represent actual processes, which goes back to the differing interpretations of the diagrams by Dyson and Feynman, dubbed by Kaiser (2005, p. 175 ff.) as the "Feynman-Dyson split". To some extent, this split is traceable even further back to the split between the field-theoretic and the particle approach mentioned in Section 2.1.1, as Feynman was a strong advocate of the particle approach (Feynman, 1966). Since Dyson (1949a, 1949b) proved the mathematical equivalence of the two methods, the question of their interpretation became merely philosophical. Kaiser (2005, p.263 ff.) sketches the introduction of Feynman diagrams within textbooks for students of particle physics: While many authors claim to follow Dyson's rigorous derivation of the diagrams, in fact, they liberally use them also before the derivation of the rules and by that let them "speak for themselves" (Kaiser, 2005, p. 265). This went so far that 't Hooft and Veltman (1974) stated that "[f]ew physicists object nowadays to the idea that diagrams contain more truth than the underlying formalism" and that "diagrams form the basis from which everything must be derived" ('t Hooft & Veltman, 1974, p. 178), i.e., they turned Dyson's idea that the diagrams are helpful only because they can be rigorously derived from first principles, upside down and put them in the beginning.

The arguments against a representational function of Feynman diagrams are further explicated by, e.g., Kuhlmann (2010), Passon (2019), and Weingard (1988). Niels Bohr mentioned the oldest argument at the Pocono meeting in which Feynman diagrams were first represented. According to Heisenberg's uncertainty principle, any notion of a trajectory of subatomic particles is meaningless (Kaiser, 2005, p. 47). The second argument concerns the superposition of the diagrams. Therefore, it is never a single diagram that is observed. Weingard (1988) was one of the first who exploited this argument by arguing that there are always several Feynman diagrams contributing to a process and, hence, the "virtual particle type and number are not sharp". He goes on to make two technical arguments about the derivation of Feynman rules without creation and annihilation operators and the introduction of so-called "ghost fields" [11]. Passon (2019) explicates Weingard's first argument with a case study of the discovery of the Higgs boson. He argues that, technically, a "direct observation" of the process $H \to \gamma\gamma$ has never been done because this process is always in superposition with the "background", i.e. all the other processes which create 2 photons. Since all Feynman diagrams in which two photons are created, including those in which a (virtual) Higgs bosons transforms (via different ways) into two photons have to be summed up and then squared to calculate the cross-section of the process (cf. equation 2.1), there is no single "$H \to \gamma\gamma$-diagram" which could be observed. A third argument against "literal reading", i.e., the reading of a diagram as a picture of a process, is that of "topological equivalence" (Passon, 2019; Passon et al., 2020): Since Feynman diagrams can be deformed into each other as long as the initial and final state stays the same, the "stories" that can be told about what happens inside the "black box" between the initial and final state can be very different for the same equivalence class of Feynman diagrams. Figure 2.4 shows an example of three topologically equivalent first-order diagrams for Compton scattering ($e^- + \gamma \to e^- + \gamma$). In diagram a, an electron absorbs a photon and emits it after some time; in diagram b, the electron first emits a photon and absorbs the initial photon after the emission, and in diagram c, the initial photon transforms into an electron-positron pair. At the same time, the positron annihilates with the initial electron, transforming into a photon. In the calculation, diagrams b and c contribute the same mathematical

[11] These "ghost fields", called *Faddeev-Popov ghost fields*, are so technical that they are merely mentioned in a footnote in Griffiths (2008, p. 288).

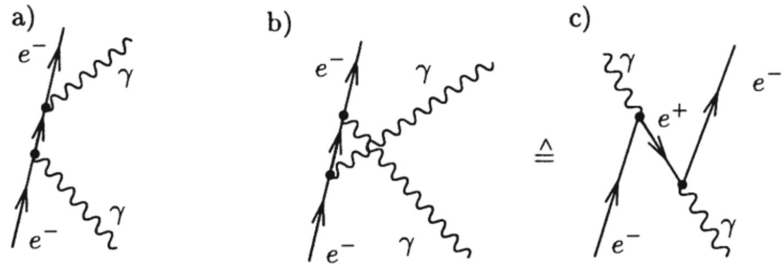


FIGURE 2.4: Example for three diagrams describing the same physical process. Diagrams b and c are "topologically equivalent". In contrast to most other diagrams in this work, the time flows from bottom to top. The image is taken from Passon (2019).

term; hence, only one of them has to be considered. This is what is called "topological equivalence". Passon et al. (2020) compare this situation with drawings of electrical circuits where the same circuits can be drawn very differently, each representing the same physical circuit. In both cases, there are just some topological features of the depictions that matter (in the case of the Feynman diagrams, these are the initial and final state and the number and types of vertices).

However, there are also arguments which support a representational function of Feynman diagrams, and with that to some extent a "literal reading". Harré (1988) takes on a pragmatic view and asks where the "photonic talk" originated. He traces the success of the diagrams to the "power of the material practices of a scientific community - those involved in doing experimental work – to shape concepts in the basic theory". Since the physical apparatus experimental physicists work with produces tracks, he argues, it is only natural that the external and internal lines are referred to as particles. Meynell (2008) argues that Feynman diagrams are "in some sense pictures" but notes that this question is independent of any question about the "epistemological interest" of Feynman diagrams, meaning that the question of whether Feynman diagrams depict something "true" remains untouched by the question of whether a single Feynman diagram represents something. Meynell distinguishes the term "representation" from the term "denotation". She argues that representations do not necessarily represent a real object or something that is, in principle, within the realms of our perception in contrast to a denotation, which is simply a picture of what is happening. In that sense, Feynman diagrams would be representational pictures of what happens between the initial and final state of a particle process – without stating that the particles represented in between actually exist.

Stöltzner (2018), drawing on Wüthrich (2011, 2012), takes the middle ground between these two positions and argues how Feynman diagrams can be seen as models to mediate between mathematics and physics. He distinguishes four functions the diagrams have for physicists. Throughout this work, I will use diagrams in all four of these functions.

- A single Feynman diagram viewed in isolation is a mere representation of a mathematical term in a perturbation series, as introduced in Section 2.1.2 (equations 2.3 -2.5). Often, the single elements are labelled with the terms they represent, such as in Figures 2.1 or 2.2.

- A set of diagrams for a particular type of physical process. Physicists usually use one specific diagram to represent the process, knowing that other diagrams also contribute to the observed cross-section of the process. For example, Figures 5.1, 8.1, 8.2, and 8.3 all show diagrams of the lowest order representing different processes.
- All diagrams up to a certain order of corrections. These are drawn to explain how a particular cross-section's calculation must be modified to obtain more precise results. Figure 2.3 shows the first two corrections for a specific process.
- The whole infinite series of Feynman diagrams. This cannot, in principle, be drawn as it requires infinitely many diagrams and is a purely theoretical construct.

From this account, it is clear that Feynman diagrams have a highly discipline-specific meaning that might not be directly accessible to learners. As it will be explained in Section 4.3, it requires practice for learners to discern these different meanings. All this should lead students to understand better the role of a model in modern science, a crucial feature of the *Nature of Science*.

2.2 Particle Physics in School

In a recent direction paper for the future of modern physics education, Kersting et al. (2023) have mapped out possible future directions in contemporary physics education. They argue that teaching particle physics in schools is crucial in the modern era of rapidly advancing physics. With recent groundbreaking discoveries like the Higgs boson and gravitational waves and quantum technologies gaining ever-increasing attention, it is evident that physics education needs to evolve to keep pace with these developments. They argue that by integrating particle physics into school curricula, students can engage with the forefront of scientific discovery, fostering a deeper appreciation for the universe's complexities and inspiring the next generation of physicists. In this section, I sketch out what has been done in particle physics education throughout the last decades and focus mainly on Feynman diagrams. In the end, I will give an overview of the topic of the Nature of Science (NoS), as this is a critical overarching framework which allows the embedding of modern physics topics into physics teaching, even if it is not an explicit part of the curriculum.

2.2.1 General approaches

Particle physics as a high school topic was first introduced in the early 1980s with Swinbank (1992) citing a Dutch school curriculum development project from 1981 and 1984. Also, in 1984, the first conference on teaching modern physics was held at CERN (Aubrecht, 1986). Since then, various other calls for including particle physics have been made, e.g. by Barlow (1992), Chatterjee (2002), Kobel (2003), and Tuzón and Solbes (2016). Some proposals even aim at lower secondary (Wiener et al., 2015, 2017b) or even primary (Alexopoulos et al., 2018; Pavlidou & Lazzeroni, 2016) schools. A recent curriculum study (Kranjc Horvat et al., 2022) has investigated 27 high-school physics curricula concerning the extent to which particle physics is included. They found that particle physics had a dedicated chapter in 12 of the 27 investigated curricula. In comparison, particle physics only occurred in the other 15 curricula as part of different topics, such as quantum physics.

Following the above-mentioned calls to integrate physics into high school curricula, many proposals have been made to suggest how to do so. I will give a comprehensive overview of the proposals that include Feynman diagrams in Section 2.2.2; therefore, I only exemplify some of the proposals in this section. Arguably, one of the most comprehensive educational reconstructions of the Standard Model of Particle Physics (SMPP) in the German-speaking community has been carried out by Kobel and Lindenau (2020) and Lindenau and Kobel (2019) and is presented comprehensively in Kobel et al. (2022). In the teaching & learning material, they introduce particle physics using fundamental interactions, charges, and elementary particles as central concepts that are intertwined. Hobson (2010, 2011) started from quantum mechanics and introduced quantum field theory, thus choosing an approach rather close to theoretical physics. In his approach, he argues that particles are mere artefacts, with fields being the fundamental objects. This notion is also found in other educational approaches (e.g., by Allday (1997), Bertozzi (2013), Daniel (2006), and Organtini (2011)).

Furthermore, Wiener et al. (2017a) focuses on explaining the concept of colour charge and proposes a framework to introduce the colour charge in contrast to previous learning material. This framework uses striped instead of complementary colours to denote "anticolours", which are part of the charge concept in the strong interaction. This reduces the cognitive load as students do not have to remember complementary colours. A teaching approach that included the experimentalist side of particle physics was made by Polen (2019), even though their proposal does not explain how they introduced this aspect of particle physics. A comprehensive overview of how particle accelerators, in particular the Large Hadron Collider (LHC), can be introduced in the physics classroom, even if particle physics is not a dedicated topic, was compiled by Wiener et al. (2016). An analogue to the educational reconstruction of the SMPP presented above is the learning material presented by Kobel et al. (2021), which focuses on the research method, i.e. particle accelerators and detectors.

Some approaches focus less on conveying particular concepts but rather on giving students a playful introduction to the topic (Alexopoulos et al., 2018; Lindenau & Winkler, 2020; MacDonald & Bean, 2009; McGinness et al., 2019; Pascolini & Pietroni, 2002; Pavlidou & Lazzeroni, 2016; Riva et al., 2022). In contrast, other approaches focused on approaching the subject through arts (Andrews & Nikolopoulos, 2018; Nikolopoulos & Pardalaki, 2020). Yet other methods are more advanced and let students get hands-on data from actual particle physics experiments and explore the daily work of a particle physicist (Bilow & Cecire, 2022; Cecire et al., 2014).

Regarding the research of students' conceptions, there is a vast body of research into students' conceptions of quantum physics, which is focused chiefly on wave-particle duality, wave functions, and atoms (Krijtenburg-Lewerissa et al., 2017). Several proposals have been made to help students form more coherent mental models of concepts from quantum physics, among them quantum visualisations such as PhET (McKagan et al., 2008) or Quantum Composer (Zaman Ahmed et al., 2021), quantum games (Seskir et al., 2022), or innovative learning sequences such as one based on Feynman's sum over paths approach (Malgieri et al., 2017) which is used in Feynman's original derivation of his diagrams (cf. Section 2.1.2). On the particle physics side, there are just two studies: Tuzón and Solbes (2016) have investigated students' conceptions of Spanish high-school students about the structure and interactions of matter. The results they found were very variable with having a classical picture in general. However, fragments of a more modern picture are easily confused with more classical ideas. Gourlay (2016) studied the conceptions of British year-12 high school students using the method of concept mapping. They found that students mostly knew about the quark types up, down, and strange, particle annihilation, and the electron being a type of lepton. However, they found misconceptions regarding other kinds of leptons and about matter being constituted of antiparticles. Surprisingly, they also found misconceptions regarding annihilation pair production, while the definition of annihilation was among the most common correct conceptions.

2.2.2 Feynman diagrams as educational tools

Only a few years after Feynman diagrams were introduced, they eventually made it to particle physics textbooks (Kaiser (2005, p. 253 ff.), see Jauch and Rohrlich (1955) for the first textbook on quantum electrodynamics). Here, they were introduced to teaching QED and, later also, other quantum field theories to university students of particle physics. With being more and more widespread, they also became popular in physics outreach material, as illustrations of physics, and even in popular culture (Cendrowski, 2008; Cham, 2007; Munroe, 2015b). Also, Feynman contributed to the diagrams' spread by painting them on a van (Jepsen, 2014).

With the rise of particle physics as a curriculum topic in school (cf. 2.2), they were also used as educational tools, as they even made their way into physics textbooks for secondary schools (Diehl et al., 2011; Grehn & Krause, 2011; Sexl et al., 2019). In this section, I review various approaches to how Feynman diagrams are explained in literature published in physics education journals or how they are used to explain concepts from particle physics. The approaches are primarily targeted at the secondary level, even though some might be suitable only for the university level.

- Lambourne (1992) highlights the interplay between theory and experiment, but also the mathematical nature of the diagrams and shows how the diagrams are used to predict physics.
- Allday (1997) explains the principle of fundamental interactions and interaction particles using Feynman diagrams. He builds on the image of an interaction particle being exchanged, which can be motivated by drawing a Feynman diagram.
- Dunne (2001) explains Feynman diagrams step by step by introducing fundamental vertices and highlights their use as a formal tool rather than an informal illustration.
- Jones (2002) uses Feynman diagrams to illustrate virtual particles. He explains them using Heisenberg's uncertainty particle and describes the nature of forces in quantum physics with the concept of interaction particles.
- Daniel (2006) introduces quantum fields as operators and Feynman diagrams as calculation tools. This approach is contrasted by Bertozzi (2013) with an approach by Hobson (2005), which treats fields as fundamental objects very close to a classical electromagnetic field.
- Organtini (2011) outlines an introductory lecture on particle physics for the general public but with a mathematical background. Thus, he gives a mathematical derivation of Feynman diagrams and then motivates why they can also be interpreted physically.
- Johansson and Watkins (2012) use Feynman diagrams to explain the Standard Model of Particle Physics starting from the elementary particles. They introduce Feynman diagrams as representations of particle interactions but make the point that they are used in calculations.
- Woithe et al. (2017) use fundamental vertices of Feynman diagrams to explain the single terms of the Lagrangian, which describes the Standard Model.
- Lindenau and Kobel (2019) use Feynman diagrams to visualise particle interactions, first with a black box at the place of the interaction particle, and then reveal the nature of this interaction particle by including it explicitly. This educational reconstruction is shown in Figure 2.5 (top).
- Harlander (2021) sketches the theoretical way from the path integral to Feynman diagrams and explains the theoretical foundation quite comprehensively on a level suitable for advanced upper secondary to undergraduate students.

- Berg and Hoekzema (2006) and Hoekzema et al. (2005) introduce a learning unit for secondary students which features a modified version of Feynman diagrams which they call "reaction diagrams". These reaction diagrams do not include intermediary, i.e., virtual particles and no interaction particles, and have only one arrow direction. In this learning unit, they explain the conservation laws and symmetries built into Feynman diagrams. This educational reconstruction is shown in Figure 2.5 (bottom).
- Passon et al. (2020) point out a conceptual similarity to circuit diagrams in that two diagrams can be "topologically equivalent", which means that even though they differ in their geometrical depiction, they still represent the same thing (in the case of a Feynman diagram a particular mathematical expression, in the case of the circuit diagram one specific circuit).
- Wagner (2020) sees Feynman diagrams as a possibility in the teaching of particle physics to talk about possible particle processes as well as the interference between different diagrams but also points out possible misconceptions arising from the use of the diagrams (see below).
- Launer (2020, p. 276 f.) proposes a teaching and learning sequence in which he introduces the basics of particle physics for upper secondary school with a focus on the concept of interaction particles, in which he also introduces Feynman diagrams to motivate that particle processes are probabilistic.

Besides approaches for explanations and educational reconstructions of Feynman diagrams, there are also "hands-on" approaches that introduce Feynman diagrams in a rather playful way that aims to engage the learners. These approaches are primarily targeted at the secondary level – with the notable exception of the last example – to get students in touch with Feynman diagrams for the first time, but not to use them for particle physics calculations.

- Kontokostas and Kalkanis (2013) developed an activity in which students combine fundamental vertices to draw diagrams representing electron-positron-photon interactions.
- Day et al. (2022) built on this activity, but instead of giving the students ready-made fundamental vertices, they gave them rules on how to draw the diagrams to foster "lateral thinking", i.e., to think "about a problem with effortful creativity and a tangential approach".
- Pascolini and Pietroni (2002) use toys where the rules to create Feynman diagrams are mechanically built in as metaphors for particle processes.
- Lindenau and Winkler (2020) playfully introduce Feynman diagrams by creating a game called "Feynman-Rhombino", which includes different fundamental vertices which should be put together according to some rules modelled on the actual rules how the diagrams work.

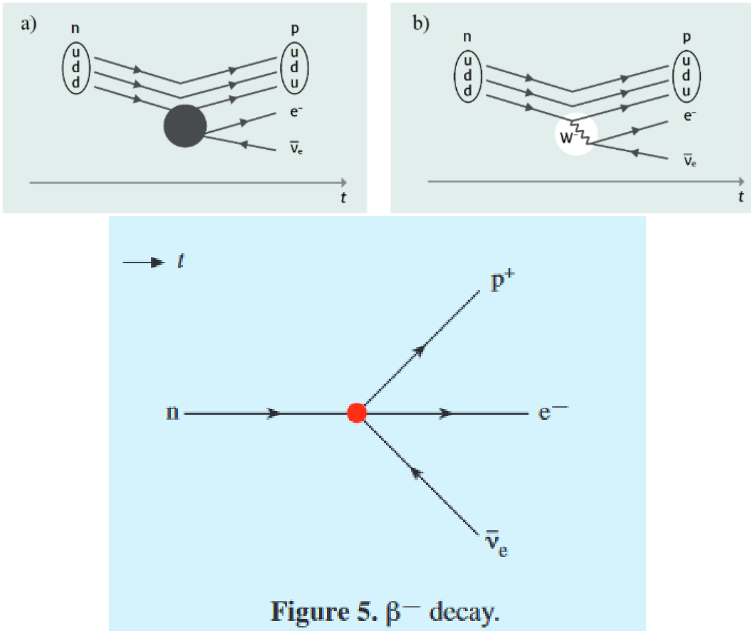


FIGURE 2.5: Two examples for educational reconstructions of Feynman diagrams. The top one, using a "black box", is presented in Lindenau and Kobel (2019) where it is cited from Kobel et al. (2022). The bottom shows a "reaction diagram" from Berg and Hoekzema (2006). Details in the text.

- Riva et al. (2022) created another game called "Tutti Quantum" based on a similar principle.
- Harlander et al. (2020) created a computer program called "FeynGame", which can be used to either draw Feynman diagrams and check them for correctness and even automatically print the mathematical expression the diagram represents or to draw one diagram for a given process. The purpose of this program is, on the one hand, to familiarise oneself with the mechanisms of a Feynman diagram, find out which diagrams are possible, and draw diagrams which can even be used professionally. Most of the diagrams seen in this thesis, such as those in Figures 8.1-8.3, are drawn with this program.

As lined out in Section 2.1.3, the interpretation of Feynman diagrams and, hence, their suitability for education is disputed. Passon et al. (2018) criticise multiple of the approaches mentioned above. In general, they criticise the depiction of particle physics in these educational approaches as oversimplified. In particular, they criticise a "literal reading", as defined on page 13. Also, they criticise "very vivid" wording, which may lead the reader to believe that the diagrams depict particle paths in space and time, even though they acknowledge that the mathematical complexity is emphasised. There are nearly always remarks that caution is needed when interpreting the diagrams. Another criticism is that virtual particles are described as mediators for particle interactions. At the same time, they are "just artefacts of a specific solving technique" and hence a "calculational device" only (Passon et al., 2018, p. 14).

The current work aims to create learning material to explain concepts from particle physics using Feynman diagrams, which partly addresses the considerations from Section 2.1.3. However, I will mainly build on the argument by Stöltzner (2018) that there are four ways in which Feynman diagrams are used in practice.

2.2.3 Nature of Science

While many aspects of particle physics allow a direct connection to other physics topics, like Wiener et al. (2016) did by providing the LHC as a context to discuss topics like electricity or magnetism, it is also worth using it to examine the broader picture of science. This aspect of science teaching is subsumed under the term *Nature of Science* (NoS), which is at the core of scientific literacy and which is defined by Schwartz et al. (2004) as referring "to the values and underlying assumptions that are intrinsic to scientific knowledge, including the influences and limitations that result from science as a human endeavour". There is no general agreement on what exactly constitutes NoS from a philosophical, historical, or sociological point of view. However, in this work, I will follow the suggestion of Abd-El-Khalick (2012) and adopt the pragmatic "consensus view" of NoS, which provides lists of aspects relevant to the school science curriculum. These lists are not all identical, but the largest number of aspects can be found on all of them. This work uses a list from Lederman (2013), which is not meant to be comprehensive but rather to give a frame of reference. This list constitutes six aspects of NoS, which are

- Observation and Inference: Students should distinguish between observations, which are directly accessible to the senses or their extensions without interpretation, and inferences, which go beyond sensory perception to develop explanations or models essential for understanding complex phenomena in science.
- Scientific Laws and Theories: The hierarchical view of theories evolving into laws is incorrect; laws describe observable relationships, while theories explain. Both are legitimate forms of scientific knowledge, and theories are not intended to become laws.
- Creativity and Imagination in Science: Scientific knowledge is partially empirical but also involves creativity in inventing explanations. Concepts in science, like atoms and species, are functional theoretical models rather than exact replicas of reality.
- Subjectivity and Theory-Laden Nature: Scientific knowledge is influenced by scientists' beliefs, experiences, and theoretical commitments, shaping the questions investigated and interpretations made. Hypotheses and theories guide observations, and science rarely starts with neutral observations.

- Science in Cultural Context: Science is practised within broader cultural contexts, impacting and being influenced by social, political, and philosophical factors. Western science's acceptance of practices like acupuncture depends on providing explanations within the cultural framework.
- Tentativeness of Scientific Knowledge: Scientific knowledge, including facts, theories, and laws, is tentative and subject to change based on new evidence or reinterpretation. This tentativeness arises from science's inferential, creative, and culturally embedded nature, as well as logical arguments supporting the idea of constant revision. Science never claims to have an absolute answer; uncertainty is instead an inherent fact of science. However, it is vital to learn how to interpret this uncertainty. Rosenberg et al. (2022), for example, suggests introducing Bayesian reasoning in the classroom to make sense of uncertainty and, by that, build trust in science.

Lederman (2013) states in his review on the research about teaching NoS that conceptions about NoS are "best learned through explicit, reflective instruction as opposed to implicitly through simply 'doing' science". Park et al. (2019) argue that modern physics, on the example of the detection of gravitational waves, provides a well-suited context to teach the nature of science to secondary students. Within this work, I point out NoS aspects that could and should be connected to the various discussed concepts.

Chapter 3

Research Goals

As elaborated in section 2.1, Feynman diagrams are a central tool widely used in particle physics. Since particle physics is a fascinating topic for high school students, which has the potential to spark interest in physics (cf. Section 2.2), there is a need to address the topic of Feynman diagrams with high school students. As stated in Section 2.2.2, many suggestions about embedding Feynman diagrams in physics education have been made. However, only few of these suggestions were evaluated concerning their effectiveness in teaching particle physics concepts. However, it is crucial to do so as students will inevitably come across this sort of diagram, so it is worth investigating their educational value.

Therefore I have identified a research gap in developing and evaluating learning material which aimed on teaching particle physics concepts using Feynman diagrams. The guiding question for this project was the following:

Can we use Feynman diagrams for teaching particle physics concepts to upper secondary students? If yes, how?

This guiding question is two-fold. In the first part, I ask *whether* Feynman diagrams are suited to teach particle physics concepts. This includes identifying opportunities and challenges of using them in teaching and analysing whether the opportunities outweigh the challenges and whether the challenges can be adequately addressed. The second question depends on answering the first with *yes*. It requires an analysis of how exactly students perceive Feynman diagrams and learning materials involving them. From this analysis, this project aims to formulate recommendations for practitioners about how to use Feynman diagrams in teaching.

3.1 Research Goal 1: Educational purposes

To answer the first part of the guiding question, finding the educational purposes of Feynman diagrams was necessary. The educational purposes can be judged most accurately by people who have both expert knowledge in the field of particle physics and practical knowledge in teaching particle physics. These people are called "particle physics education experts." Therefore, the first research question of the first research goal was

Which educational purposes, as seen by particle physics education experts, do Feynman diagrams serve when using them for teaching particle physics to upper secondary students?

The answer to this question is learning goals, which can be achieved using learning material about particle physics focusing on Feynman diagrams.

However, to gauge the benefits of using Feynman diagrams, finding challenges that Feynman diagrams might pose to educators is also essential. These challenges are exhibited mainly by the conceptual difficulties that Feynman diagrams pose to learners. Therefore, the second research question of this learning goal is stated as

Which conceptual difficulties do Feynman diagrams pose when using them for teaching particle physics to upper secondary students?

The answer to this question includes possible inadequate conceptions about particles and their interactions and difficulties that might result from the diagrams' properties.

3.2 Research Goal 2: How to teach with FD?

The second research goal is to find practical solutions to the challenges examined in the first research goal. These practical solutions are found in designing interactive learning material for students. As the first research goal was two-fold, consequentially, this is also two-fold. The research question is, therefore, stated as

Which principles for the design of learning material for upper secondary students
a) support the educational purposes and
b) minimise potential difficulties ?

As the investigation requires both a design and an evaluation of learning material, the framework of *Design-based research* (DBR) was identified as the most suitable research framework to address these research goals. This framework is presented in the next chapter alongside the other theoretical considerations necessary for investigating the questions stated in this chapter.

Chapter 4

Theoretical Framework

The project presented in this thesis is an educational reconstruction of the topic of Feynman diagrams within the design-based research paradigm. Considerations from multimedia theory and the theory of social semiotic resources serve as design principles, while eye tracking was the predominant research method. In this chapter, I present the fundamentals of these concepts.

4.1 Model of Educational Reconstruction

This project aims to find a way how to teach Feynman diagrams that follow the physical perspective (cf. Section 2.1.2), are compatible with other physics curriculum topics, and are understandable on the secondary level. A framework to determine "as to whether it is worthwhile and possible to teach particular content areas of science" (Duit et al., 2012, p. 19) is the *model of educational reconstruction* (MER, or, in its German original, *didaktische Rekonstruktion*), first developed in the 1990s by Kattmann et al. (1996, 1997). This framework has been widely accepted within the German-speaking physics education community and has become a standard method for practitioners (Kircher et al., 2015, p. 108 ff.). The model returns to the ideas of "educational reduction" and "elementarization". The subject matter has to be "broken down" into its units, then reconstructed. This results in a new content structure for the subject matter, which does not necessarily reflect the original content structure. Therefore, this model is more than a mere "reduction" of the content. The critical components of the model are shown in a triangle, which is cited in Figure 4.1. As described before, the reconstruction starts from a clarification of the science content (1), which then has to be accompanied by students' perspectives through research on teaching and learning (2), which together leads to the design of teaching and learning environments (3). These, however, need to be evaluated. The findings from the evaluation then inform steps (1) and (2). In this manner, the process of educational reconstruction is an iterative one.

The model of educational reconstruction has been successfully utilised in various topics. Höttecke and Schecker (2021) have published a book with educational reconstructions for the most crucial curriculum topics of physics curricula. Besides, it has

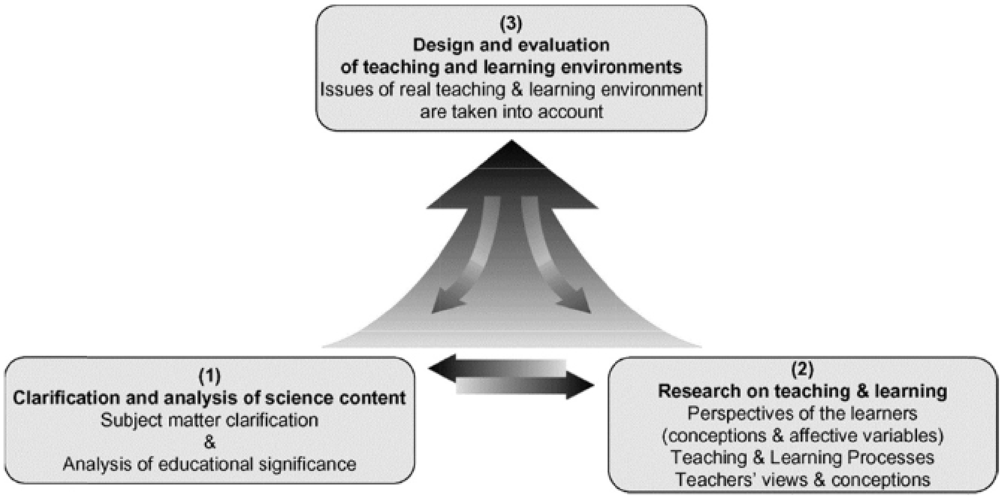


FIGURE 4.1: The components of the model of educational reconstruction as Duit et al. (2012) presents it.

been used for modern physics topics especially, among them chaos theory (Duit & Komorek, 1997), special relativity (Kamphorst et al., 2021), general relativity (Kersting, 2019a), or particle physics (Lindenau & Kobel, 2019; Wiener et al., 2017b).

4.2 Design-Based-Research

The overall framework for the work presented in this thesis is that of Design-Based Research (DBR). DBR has been invented to bridge the research-practice gap, i.e., to put results from education research directly into practice (Anderson & Shattuck, 2012; Design-Based Research Collective, 2003; Wilhelm & Hopf, 2014). The aim of DBR is twofold: On the one hand, the results of a DBR project shall inform domain-specific learning theories. On the other hand, a DBR project should result in a product which is ready to use in practice. These products might cover a wide range: It can be teaching-learning sequences (Haagen-Schützenhöfer & Hopf, 2020), textbooks (Finta et al., 2021), or online learning environments (Kersting et al., 2018). The DBR approach has been used in a wide range of topics. In physics education, it has been used, among others, to develop teaching learning sequences in the field of optics (Haagen-Schützenhöfer & Hopf, 2020) and electrical circuits (Burde & Wilhelm, 2020a, 2020b). Recently, DBR approaches have been used to develop teaching learning sequences in PhD projects about the teaching of general relativity (Kersting, 2019b), astronomy (Langendorf, 2022), and particle physics (Wiener, 2017).

As Haagen-Schützenhöfer and Hopf (2020) have shown in Figure 4.2, the general process of DBR consists of two cycles. The starting point of any DBR project is a practical problem. From there, hypotheses are formed based on previous research and existing learning theories which inform the creation of draft material. This step is usually based on existing general learning theories that consider cognitive and

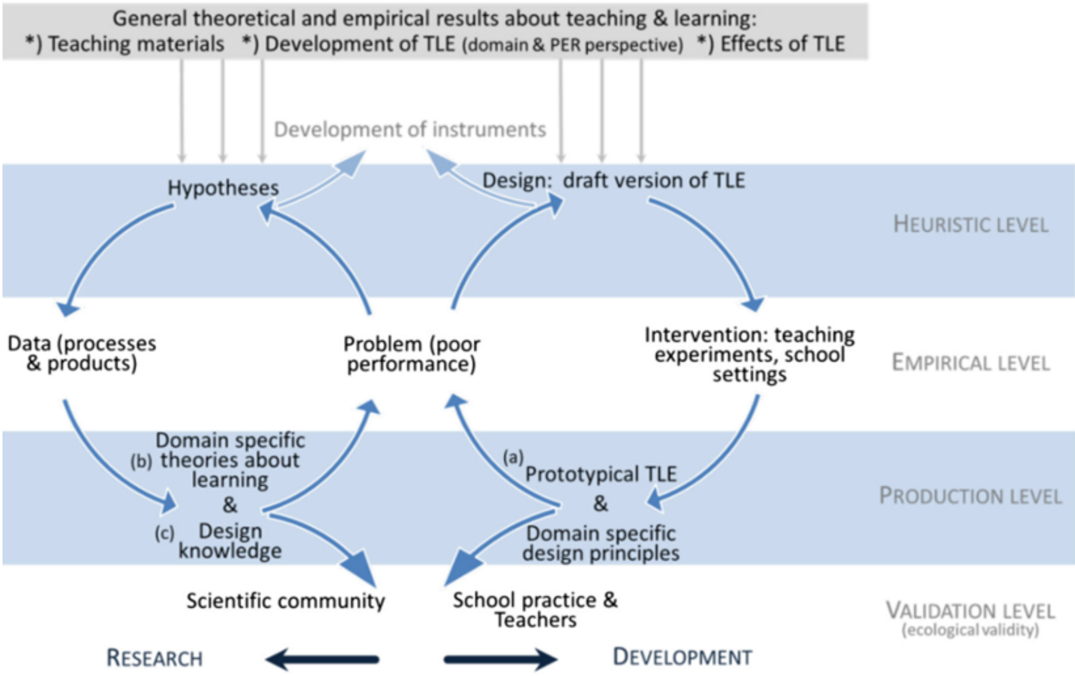


FIGURE 4.2: A visualisation of the DBR process. The image is cited from Haagen-Schützenhöfer and Hopf (2020). The abbreviation "TLE" stands for "teaching and learning environment".

affective variables that support and hinder learning and domain-specific theories that consider elements specific to the particular domain. A common framework that considers the subject matter and the learners' perspective to develop educational material is the framework of educational reconstruction. This framework is further described below. This first phase of designing the material usually uses a consultation of the relevant literature and discussions with other education researchers as well as educators and subject specialists from the field. The material that is designed in the process described above is then tested in an ecologically valid environment (empirical level in Figure 4.2), i.e., in use cases and with the target group the material is designed for. It can also be presented to people with experience in teaching who know about students' understanding from practice. The outputs of the DBR process are then domain-specific theories which can be used by the scientific community as well as a prototype of the product which was developed that can be used by practitioners. As DBR is a cyclical process, the process continues with a prototype redesign based on the insights gained from the first cycle. Due to its iterative nature, the model of educational reconstruction and the framework of design-based research are complementary and very well suited to be used together. It is described in Section 5.1.1 which stage of the DBR cycle is addressed in which part of this thesis.

4.3 Multiple Representations in Physics Education

When learning science, learners often have to learn from multiple different types of media, most commonly from text and pictures, but also videos, or interactive media such as simulations. In this section, I present the most influential theories on how learning with multimedia takes place and which considerations for the practice can be derived from these.

Functions of Multiple Representations. In multimedia learning environments, multiple representations (MR) play a significant role. Multiple representations can have three key functions in learning: They can complement each other, constrain each other, or construct a deeper understanding (Ainsworth, 1999; Opfermann et al., 2017). For example, in the case of Feynman diagrams, they can be complemented by annotations of charges or their interpretation can be constrained by visualisations of experimental data that represent the underlying processes, such as event displays. Similarly, Feynman diagrams can contribute to a deeper understanding of particle physics by using them to explain unfamiliar data visualisations, such as the Z resonance curve (for the latter two examples, cf. Figure 8.3). However, there are several cognitive tasks associated with learning. First, learners need to understand how information is encoded into the representational form. They also need to understand how the representation relates to the domain and know how to select the correct representation and how to construct it. These functions and cognitive tasks are described by Ainsworth (2006) together with specific design heuristics in the "Design, Functions, Tasks"-Framework (DeFT) for multiple representations. These design heuristics consider the number of used multiple representations, the conveyed information, the form, sequence, and mode of translation between the various representations.

Cognitive Load Theory. One of the most fundamental cognitive theories of learning is the *Cognitive Load Theory* (CLT) by Chandler and Sweller (1991). The CLT states that the cognitive load while learning from instructional media consists of three different kinds of cognitive load (Paas et al., 2003; Sweller, 2010): *extraneous cognitive load* (EL), *intrinsic cognitive load* (IL), and *germane cognitive load* (GL). The intrinsic load refers to the load caused by the learning content itself. The IL is a function of the type of task and the learner's prior knowledge. For a given task and learner, the IL is determined by *element interactivity* (Sweller, 1994, 2010), which is the interconnectedness of different elements that should be learned. For example, it is low for memorising facts, such as dates or chemical symbols, but it is higher when these dates or symbols have to be connected. Higher element interactivity results in a higher IL because the working memory has to make connections between different elements. Extraneous, also referred to as ineffective cognitive load (Paas et al., 2003), is independent of the contents of the learned materials, as it is the load imposed upon the learner by the ineffective design of the learning material. EL occurs, for example, when learners have to look for information. Section 4.3 shows a list of certain measures that mostly aim to reduce extraneous load. The third category, germane cognitive load, is also called effective cognitive load (Paas et al., 2003). This is the load which learners dedicate to the learning process. It is not independent of IL and EL, but in contrast to those, it depends on learner characteristics (Sweller, 2010). High IL in connection with a low EL will result in high GL as working memory capacities can be devoted to learning the material. In contrast, a high EL will result in a low GL, as the resources are needed to deal with the instructional material. However, an essential factor which determines GL is motivation. This factor is further discussed below. Leppink et al. (2013) have developed and validated an instrument which measures the different kinds of cognitive load. An adapted version of this instrument is used within parts of this thesis (cf. Section 6.6.2).

Cognitive Theory of Multimedia Learning. Mayer (2002a) presented a comprehensive theory about how learning is most effective when incorporating different media, such as text and pictures. His theory is based on three assumptions. The dual-channel assumption states that humans process visual and auditory information via different channels (Baddeley, 1992; Paivio, 1990). The limited-capacity assumption is based on the CLT and states that these channels can only process a limited amount of information (Baddeley, 1992; Chandler & Sweller, 1991). The third assumption states that learners need active processing of incoming information to learn meaningfully (Wittrock, 1989). This assumption is closely related to the constructivist view of learning (Duit, 1996), which assumes that knowledge has to be constructed by the learners themselves by engaging with the subject matter in contrast to a view focused on receiving knowledge. The theory resulting from the three assumptions is the *Cognitive Theory of Multimedia Learning* (CTML, Mayer (2002a, 2005)). The theory postulates that five different cognitive processes take place while learning with multimedia: The students (i) **select verbal information**, which means paying attention to parts of the written text or narration provided in the learning material, (ii) **select non-verbal information**, which means paying attention to parts of the pictures, animations, videos etc. provided by the learning material, (iii) **organise words**, which means that they organise the selected verbal information into a mental representation, (iv) **organise images**, which means that they create a mental representation of the non-verbal information, and, finally (v) **integrate** the mental representations of verbal and non-verbal information with each other and with prior knowledge from long-term memory. From the theory, Mayer has derived and tested design principles which should guide designing instructional materials and should foster learning from these materials by reducing extraneous cognitive load. These design principles are presented in Section 4.3.

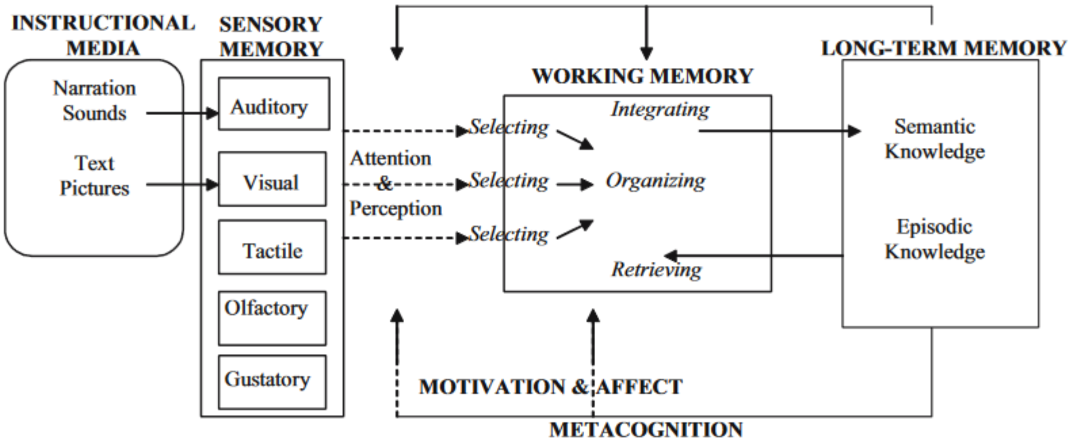


FIGURE 4.3: A scheme of the CATLM. The model is an extension of the CTML, therefore the latter is not visualised here. The figure is cited from Moreno and Mayer (2007).

Cognitive-Affective Theory of Learning with Multimedia. As already mentioned in the paragraph about CLT, motivational factors play a crucial role in learners' ability to assign mental capacities to learning. This fact was accounted for by expanding the CTML into the *Cognitive-Affective Theory of Learning with Multimedia* (CATLM, Moreno and Mayer (2007)), which also takes affective and metacognitive factors into account. Moreno and Mayer (2007) present the CATLM as depicted in Figure 4.3. Most parts of this scheme are already described in the paragraph about the CTML. The CATLM adds on these cognitive processes that the cognitive processes are partially guided by the prior knowledge, as illustrated by the top arrows in Figure 4.3 and that metacognitive skills are used to regulate cognitive processes, as illustrated by the arrows at the bottom. Metacognitive skills mean the awareness of strengths and limitations of one's knowledge, strategies, affect and motivation. In their theory, Moreno and Mayer (2007) redefined the different types of cognitive load, mentioned above. They distinguish between *extraneous processing*, which is defined in the same way as extraneous cognitive load, but has a particular subcategory, named *representational holding*, which describes the process of holding a representation in the working memory during the meaning-making process. The other parts of cognitive load are named *essential* and *generative processing*. These are related to the intrinsic and germane cognitive load described above. Essential processing is defined as the cognitive processes used to select new information. It is thus a function of the complexity of the learning material, defined above as element interactivity, and prior knowledge of the learner. Generative processing is defined as the sense-making processes, such as organizing and integrating processes. Moreno and Mayer have derived five design principles based on this theory which expand the design principles from CTML and apply especially to interactive learning environments. These design principles are presented alongside the design principles from CTML in Section 4.3.

Design principles The CTLM and CATML described above provide a list of design principles for multimedia learning environments (Mayer, 2002a; Moreno & Mayer, 2007).

- The **multimedia effect** states that learners learn better from text and pictures than from text alone.
- The **spatial contiguity effect** states that corresponding text and picture should be presented close to each other.
- The **temporal contiguity effect** states that corresponding narrated text and pictures should be presented simultaneously rather than successively.
- The **coherence effect** states that irrelevant text and pictures have to be excluded to create a more effective learning experience.
- The **modality effect** states that using different modalities, i.e., pictures and narrated text instead of picture and printed text has a positive effect.
- The **redundancy effect** states that adding redundant information, such as narration and on-screen text, might hurt learning.
- The **pre-training effect** states that learning about necessary terms and definitions before the actual content is beneficial for learning.
- The **signalling effect** states that including certain pointers, such as structuring words in the accompanying text or narration or picture elements that highlight certain elements, are beneficial for learning.
- The **personalisation principle** states that explanatory text should be in a conversational rather than a formal style.
- The **guided activity principle** states that students should engage with a pedagogical agent which helps guide their cognitive processing.
- The **reflection principle** states that students should reflect upon their answers to foster organization and integration processes.
- The **feedback principle** states that students need explanatory rather than corrective feedback.
- The **pacing principle** states that students need to determine their own pacing to allow students to process smaller chunks of information.

Representational Competencies. To successfully learn from visual representations, learners must have *representational competencies*. However, to acquire these representational competencies, they have to learn about them. Therefore, visual representations have a dual role. This is what Rau (2017) dubbed the "representation dilemma". She provides a framework of how the different representational competencies are acquired through visual representations while learning science content. There are two broad categories of these competencies: conceptual and perceptual. Conceptual competencies involve the visual understanding of visual representations, e.g., the ability to connect visual representations to concepts, and connectional understanding, e.g., the ability to connect several multiple representations. These skills are acquired through verbally mediated sense-making using particular or multiple visual representations. Conceptual competencies have to be distinguish from perceptual competencies, though they are interrelated. These include visual fluency, e.g. efficiency in connecting visual representations to concepts, and connectional fluency, e.g efficiency in connecting multiple representations. Perceptual competencies are acquired through nonverbal inductive learning processes while participating in disciplinary discourse. Besides that, Rau (2017) also defines meta-representational competencies, which include, e.g., the ability to choose visual representations based on task demands or context. These are acquired both through verbally mediated sense-making and nonverbal inductive learning processes. As with cognitive load, representational competencies are also defined slightly differently by different authors. For example, conceptual and perceptual competencies are subsumed under the term representational fluency, while meta-representational competencies are summarised under the term representational flexibility by De Cock (2012).

Conceptual challenges posed by visual representations The challenges posed by this representation dilemma are further investigated, e.g., by Dorris and Rau (2022), who investigate the conceptual challenges posed by a very peculiar representation in chemistry, the atomic orbital energy diagram (AOED). This diagram poses challenges that are in some way similar to those posed by a Feynman diagram, for example, by using arrows to denote particular properties of particles different than the direction of motion (in the case of the AOED, the spin state, in the case of the FD the difference particle/anti-particle), or by using a representation that might suggest some spatial order. However, it is just a conventional notation of energy states (in the case of AOED) or time-ordering between initial and final states and virtual states in between (in the case of FD).

Chunking mechanisms. Visual fluency can colloquially be described as "seeing at a glance". More elaborately, it involves *perceptual chunking*, which describes the ability to retrieve corresponding "chunks" from long-term memory based on relevant visual features. These chunks are defined as collections of elements having strong associations with one another but weak associations with elements within other chunks (Gobet et al., 2001; Sweller, 1993). These chunks were first investigated by a field in psychology called "Gestalt psychology" (Humphrey, 1924; Wertheimer, 1923). The study of chunking mechanisms goes back to the study of expertise, particularly chess (Chase & Simon, 1973). Gobet et al. (2001) distinguishes perceptual chunking from *goal-oriented chunking*, which is a deliberate process, whereas perceptual chunking is more of an automatic and continuous process. The acquisition of visual fluency and the development of visual chunks in long-term memory is a crucial step to acquire perceptual fluency and thus expertise in a field.

Social Semiotic Resources Another perspective on multiple representations is the framework of Social Semiotic Resources (SSR, Airey and Linder (2017)). This framework includes more than what is classically referred to as representations; it also includes resources like experimental practices or specific laboratory equipment. The main characteristic of SSR is that it focuses on group meaning-making, i.e., common practices which foster communication within a particular social group, in the case relevant to this work within particle physicists. Examples include graphs, sketches, specific language, etc. For this reason, the Feynman diagram is a particularly well-suited example of an SSR since it has been developed and evolved within the group of particle physicists and is now a common tool used for different kinds of calculations (cf. Section 2.1). Each SSR has a set of disciplinary-specific meaning potentials which are called *disciplinary affordance* (Airey, 2015; Fredlund et al., 2012) and which are defined as "agreed meaning making functions that a semiotic resource fulfils for the disciplinary community" (Airey & Linder, 2017). As illustrated in simplified physics concepts, a critical constellation of SSR is essential for grasping disciplinary knowledge effectively. This constellation, however, is often not spontaneously accessible to students since in the everyday use of physicists, only a single SSR is used which functions as a *disciplinary shorthand*, necessitating explicit teaching to interpret and coordinate these resources fluently for understanding. Achieving mastery of these resources, termed representational competence (cf. Section 4.3), is likened to linguistic fluency, emphasizing understanding and adept use within physics contexts. Fluency in semiotic resources, similar to language fluency, evolves through repeated practice and educational strategies. SSR might not only have disciplinary affordances but also *pedagogical affordance*, which is defined as "aptness of a semiotic resource for the teaching and learning of some particular educational content" (Airey, 2015). The relationship between disciplinary and pedagogical affordance is mostly an inverse one, i.e., an increase in the pedagogical affordance of an SSR leads to a decrease in the disciplinary affordance because it does not serve as disciplinary shorthand anymore. The increase of pedagogical affordance can be achieved by *unpacking* the high disciplinary affordance of an SSR (Fredlund et al., 2014).

4.4 Eye Tracking in Physics Education Research

Since learning with multimedia relies on learners dedicating their attention to visual stimuli which can be presented to them in a controlled environment, a method that is particularly well suited to investigate the cognitive processes connected with it is to analyse their eye movements. Indeed, eye-tracking has a decades-long history of being used to study learning with multimedia (Alemdag & Cagiltay, 2018; Hyönä, 2010; Mayer, 2010). Specifically in physics education research, however, it is relatively new. In a recent review, Hahn and Klein (2022) found that eye-tracking was used in a study dedicated to PER for the first time in 2005 (van Gog et al., 2005). In this section, I present how eye tracking is used in PER to uncover learning processes.

4.4.1 Description of the Eye Tracking Method

One of the first methods to measure eye movements for psychological research was described by Delabarre (1898):

"*After considering and testing numerous possibilities, forced to the conclusion that only by firmly attaching some object to the surface of the eye or to the eye-muscles, as a support to a mirror or to a thread for moving levers, could my object be obtained. [...] I made a few plaster casts over the cornea of an artificial eye. [...] Then I made the eyeball anaesthetic by applying two or three drops of a two to three per cent solution of cocaine, and on fitting the cast over the cornea found that it held there perfectly without pain or discomfort.*"

He continues describing this method in every detail and concludes with the remark

"*As to whether there is any danger to the eye to be feared from using it in this manner, I cannot say with assurance. [...] I have also found it necessary to allow a considerable interval to elapse between experiments, - usually a week.*"

Research and technology have come a long way. Fortunately, today's methods of detecting eye movements are far from being as invasive as Delabarre's, so we can say with certainty that there is no danger to the eye by using modern eye-tracking methods. These modern eye trackers are usually video-based. Figure 4.4 shows a picture of such an eye-tracker. They work by recording the eyes, detecting the pupil and creating a light reflex on the so-called *cornea*. By measuring the vector between this corneal reflex and the pupil, they can estimate the gaze point (Hansen & Ji, 2010). The resulting eye movement can be categorized into *fixations* and *saccades*. The former are movements where the gaze essentially rests within a radius of $2^\circ - 5^\circ$ for about a tenth of a second, whereas the latter are very fast movements which last between a hundredth and a tenth of a second, during which vision is suppressed (Duchowski, 2007, p. 42 ff.). Other eye-movement types, such as smooth pursuits or nystagmus, exist but won't be further discussed here.

In eye-tracking analysis, it is assumed that fixations are the expression of someone's attention which is based on the theory by (Just & Carpenter, 1980) and grounded in reading research. Saccades are then the movement of the gaze, and hence, attention, to another point. This assumes that the central part of visual attention is in the *foveal vision*, i.e., the centre-most 5° of the gaze.

FIGURE 4.4: An example of a video-based eye tracker. The eye tracker is mounted below a computer screen. The model shown here is used in the studies presented in this thesis (cf. Section 5.2). The image was taken with a camera sensitive in the near-infrared range, which is why the near-infrared light emitted by the eye tracker is visible.

4.4.2 Use Cases

In a review of 113 ET studies that investigated learning processes throughout different domains from 2000 to 2012, Lai et al. (2013) found that ET mainly was used to examine information processing patterns and instructional strategies' effects. Hahn and Klein (2022) have conducted a systematic review investigating how ET has been used in different areas of physics education research for over ten years. They investigated 33 journal articles concerning the learning domains, the methodological implementation as well as the results found in the studies. They categorised learning scenarios where ET was used in assessment, knowledge construction, and lab work. The fields of application range from the investigation of learning processes during learning with simulations (Hoyer & Girwidz, 2020) to the inquiry of assessment scenarios with multiple representations (Rosengrant et al., 2009), the influence of active manipulations or instructive assistance (Klein et al., 2019; Madsen et al., 2013; Rouinfar et al., 2014) to the use of eye-tracking metrics as predictors for the correctness of answers when solving tasks (Küchemann et al., 2021). Learning difficulties when working with kinematic diagrams, such as point-interval confusion, were also confirmed by analysing gaze data (Klein et al., 2021a).

In addition to this confirmatory character, the potential of eye tracking is also shown in exploring the visual handling of representations for the first time. In the quantum physics context, a recent ET study investigated students' strategies using an interactive learning tool called "QuantumComposer" (Küchemann et al., 2023). In the astronomy context, the gaze behaviour of university students in handling the Hertzsprung-Russell diagram was analysed (Langendorf et al., 2022). In the context of vector fields, visual strategies have been identified that experts use to assess divergence; they predominantly view the diagrams systematically with vertical and horizontal saccades (Klein et al., 2018). Especially in the latter case, it is crucial to combine the eye-tracking method with other indicators for understanding the representations at hand, such as verbal data (van Gog et al. (2005), cf. Section 5.2). Furthermore, eye tracking is used in the design of learning materials to test and extend existing theories, such as CTML (Alemdag & Cagiltay, 2018; Jarodzka et al., 2017; Mayer, 2010).

Thus, ET is a common tool for investigating learning processes when learning with representations, specifically with multiple representations. In the next section, I will give an overview of how ET is used specifically, i.e., which ET measures there are and how they can be interpreted.

4.4.3 Measures

One of the fundamental problems of research with ET is that ET data are hard to interpret. Jarodzka et al. (2017) argue that it is essential to work along existing theories and to operationalise measures carefully to get meaningful results .

The most common ET metrics, according to Hahn and Klein (2022), are:

- fixation counts, which are used as a measure of attention on a certain area - usually the so-called *Areas of Interest* (AOI),
- fixation duration, also named dwell time or visit duration, which is usually defined as the time the participant spends inside a certain AOI and as the latter one used as a measure of attention,
- mean fixation duration, which is the average length of a fixation inside an AOI and used as an indicator of processing demand,
- transition count, which is the number of transitions between different AOIs and is often used as a measure of integrating processes,
- saccadic length, which is the average length of a saccade, and
- saccadic angle, which can be used to trace certain viewing behaviour.

With regards to the relation between response accuracy and ET metrics, Hahn and Klein (2022) found that there is broad evidence that the total viewing time on a stimulus is uncorrelated with response accuracy (Han et al., 2017; Klein et al., 2020; Klein et al., 2019; Susac et al., 2021). In contrast, the viewing time on a correct option or relevant parts of a stimulus is a good predictor for response accuracy (Chiou et al., 2022; Han et al., 2017; Kekule & Viiri, 2018; Klein et al., 2020; Klein et al., 2018; Rouinfar et al., 2014; Susac et al., 2020). The picture is more mixed for transition counts: In knowledge construction scenarios, they found two studies (Chen & She, 2020; Chiou et al., 2022) in which higher transition counts were correlated with high response accuracy in knowledge construction scenarios. This result, however, was not reported in assessment scenarios (Ibrahim & Ding, 2021; Klein et al., 2021a).

Gegenfurtner et al. (2011) have conducted a meta-analysis of 65 articles reporting on 73 data sources from ET studies that investigated experts' vs novices' gaze patterns in various domains, like sports, medicine, or transportation. They found that expertise in visual domains shows up mainly through shorter fixation durations, more fixations on task-relevant areas, fewer fixations on task-redundant areas, and longer saccades and shorter times to first fixate relevant information. Alemdag and Cagiltay (2018) have evaluated 58 ET studies on multimedia learning to find out which cognitive processes introduced in Section 4.3 are connected to which ET measures. They found selecting processes were associated with the time to the first fixation on the relevant item, the place of the first five fixations, the number of times a relevant item was fixated during the first seconds, or the proportion of fixations on a specific diagram or text. Organising processes were associated with the total fixation count or duration on a relevant item, as well as the average fixation duration. The latter, however, was also used to determine the learner's cognitive load, and hence the processing difficulty. Integrating processes mainly were associated with measures based on transitions, i.e., the number of transitions between two relevant elements, or the sum of saccade paths between text and picture. Also, some studies qualitatively analysed the scan paths, i.e., the ordered sequences of elements. Some studies also quantified these scan paths using transition entropy (Krejtz et al., 2016).

Representation-specific measures While the ET results reported above refer to learning processes in general, it is also possible to investigate the learning with specific representations, like Langendorf et al. (2022) did with the Hertzsprung-Russell-Diagram from astronomy and Klein et al. (2019) and Klein et al. (2018), Klein et al. (2021b) did extensively with vector fields. Hahn and Klein (2023) have shown how to distinguish between representation-specific and -unspecific competencies using hierarchical cluster analysis. They compared two models with one taking into account fixation and transition counts on stimulus options and the other taking into account representation-specific measures like saccade angles. The second model performed better in predicting performance indicators. An analogue measure which could quantify the cognitive processes involved with examining Feynman diagrams will be presented in Section 5.2.7 in the form of "triple transition ratio". This measure quantifies the amount of transitions between particles in a vertex relative to the transitions between any three locally defined AOIs on a stimulus.

Apart from the above-mentioned commonly used measures, there are also novel measures which will be introduced in the following:

Fixation/Transition ratio. Rodemer et al. (2020) investigated case comparisons in chemistry education i.e., that students had to compare different chemical structure equations, which are discipline-specific representations like Feynman diagrams. In this investigation, they distinguished *comparative* and *focused* viewing behaviour. Since fixations are associated with focusing attention while transitions are associated with comparing processes, they introduced a new measure called the *fixation/transition ratio* (FTR). A high FTR is then associated with a relatively focused behaviour, while low numbers measure comparative behaviour. They found that comparative behaviour was exhibited especially by advanced students.

Entropy-based measures. As mentioned above, scan path comparison was used qualitatively to investigate integrating processes. These scan paths can also be quantified by calculating their so-called *entropy* as shown by Krejtz et al. (2014). The entropy measure used here is based on the concept of information entropy by Shannon (1948), which measures the amount of information within a given random variable. A high entropy value signifies a "high degree of randomness", whereas a low value of this quantity signifies a "high degree of order". To apply this concept to ET scan paths, they should be interpreted as a Markov chain, which is a stochastic system of transitions between different states, where each state only depends on the previous one. There are different entropy measures in use. If the concept is to be applied to transitions, it should be used as the *transition entropy* H_t, which is calculated as

$$H_t = -\sum_i \left(P_i \sum_j \left(P_{ij} \log_2(P_{ij}) \right) \right), \tag{4.1}$$

where P_{ij} is the probability of a transition from AOI i to AOI j and P_i is the probability of the fixation occurring in AOI i. A high value of transition entropy signifies a more even transition spread between stimulus elements. A higher entropy therefore means more randomness in the transitions and therefore a more explorative viewing behaviour. In contrast, a low value of transition entropy means a more strategic viewing behaviour (Jordan & Slater, 2009; Krejtz et al., 2016; Shic et al., 2008). Mozaffari Chanijani et al. (2016) used transition entropy to investigate differences in representational competence between experts, intermediates, and novices and found that the transition entropy of novices was much larger than that of experts. A detailed explanation of how the measure is calculated is given in Section 5.2.7.

A stationary form of this measure (Krejtz et al., 2014) also exists in the form

$$H_s = -\sum_i \left(P_i \log_2(P_i) \right), \tag{4.2}$$

which measures the spread of attention distribution. A higher value means a more equal spread of attention across the stimulus.

Chapter 5

Methods and Materials

5.1 Design Process

5.1.1 Stages of the DBR project

As the project described within this thesis is a design-based research project based on the model of educational reconstruction, it consists of several stages. In this section, I explain which stages are addressed at which points in this thesis. The project started with stating the practical problem, which is presented in the Introduction (Chapter 1), further motivated in the chapter about Feynman diagrams and their use in education (Chapter 2), and explicated in the chapter about research goals (Chapter 3). The next step of the project was the clarification of science content: the role of Feynman diagrams in particle physics (Sections 2.1.1, 2.1.2) and difficulties in their interpretation (Section 2.1.3). I also analysed how Feynman diagrams were previously treated in the PER literature (Section 2.2.2). Furthermore, research on learning with multimedia served as the theoretical background for this thesis (Section 4.3). The design of the prototype involved an interview study with experts on teaching particle physics (Section 6.1). Based on the research on multimedia, adequate design principles were chosen (Section 5.1.2), and learning goals were formulated (Section 6.2). The prototype of the learning material, which was designed based on these design principles and the first two learning goals, was then evaluated using two eye-tracking studies (the use of eye tracking is motivated in Section 4.4). Before these student studies started, I tried eye tracking with working group members in Göttingen and at CERN. This small pilot study is described in Section 6.3 and served mainly as methodological guidance. The two evaluation studies, the first and second student studies, were initially planned to be one study. Due to research-practical reasons, the first part became an exploratory preparatory study for the second one. The second student study is described in Section 6.5. The material was then redesigned, partly based on the experiences and results collected in the first two student studies. This second iteration of the learning material is presented in Section 6.6.3. This material version was subsequently tested in the third student study study using eye-tracking and a questionnaire. This study is described in Section 6.6. All these stages are given as an overview in Table 5.1 together with a

Stage in DBR	Stage in MER	How is it addressed?	Section
Practical problem		How to teach particle physics with Feynman diagrams?	1, 2, 3
Theoretical basis	Clarification of science content	Working mechanisms and interpretation of FD	2.1.2, 2.1.3
	Previous research	Previous research on teaching particle physics	2.2
	Research on learning and teaching	Research on learning with multimedia	4.3
Design of a prototype	Perspectives of practitioners	Expert interviews	6.1
	Design of learning material	Choosing adequate design principles	5.1.2
		Choosing adequate learning goals	6.2
First design cycle	Evaluation of first iteration of learning material	First and second student study: Research on students' strategies	6.4, 6.5
Redesign	Design of learning material	Redesign of the material	6.6.3
Second design cycle	Evaluation of second iteration of learning material	Third student study: Research on students' cognitive processes	6.6
Design knowledge about FDs	Outcomes from the project	Possible learning goals	7.1.1
Domain-specific learning theories		Theories about complexity of FDs	7.1.2
Domain-specific design-principles		Suggestions for practices	7.2, 8

TABLE 5.1: Stages of the DBR and MER process as explained in Sections 4.2 and 4.1

mapping to the respective stage in the DBR and the MER process as introduced in Sections 4.1 and 4.2. There are three different outcomes from this project. First, the learning goals are presented (Section 6.2) and discussed concerning their suitability (Section 7.1.1). Second, the results yield theories about the complexity of Feynman diagrams, which are discussed in Section 7.1.2, and last, there are suggestions for practice (Section 7.2) as well as future directions for a possible second redesign of the learning material (Chapter 8).

5.1.2 Design Principles

As described in the previous section, one of the two bases for designing the learning material was the choice of design principles. As presented in Section 4.3, the CTML and CATLM provide a comprehensive list of design principles which should be followed when designing instructional material. Due to the general conditions of the project, it is not always possible to follow all of the design principles. Therefore, in this section, I review the design principles and motivate which ones I used for the project and why. Besides the two multimedia theories, the framework of SSR also played an essential role in the project as it provides a framework for how learners acquire representational competencies when using disciplinary representations.

Therefore, the design principles used for this project can be categorised into three general design principles: The reduction of cognitive load, as suggested by the CLT and CTML, the engagement of learners, or, as it is dubbed in the framework of CATLM, the fostering of generative processing (Mayer, 2008), and the increase of accessibility of disciplinary representations, dubbed as "unpacking" by the framework of SSR, or as "managing essential processing" in the framework of CATLM (Mayer, 2008; Moreno & Mayer, 2007).

Reduction of Extraneous Cognitive Load

The **spatial contiguity principle** states text that complements a picture must be presented within the picture instead of alongside it. This principle is implemented, e.g., by presenting particle names of Feynman diagrams or their charges directly in the diagram instead of next to it.

The **signalling principle** states that learning outcomes are higher if important information is highlighted, using, e.g., pointer words in the explanations. According to the DeFT framework of multiple representations (Ainsworth, 2006), "a way must be found to signal the mapping between representations without over-burdening learners by making translation complex”. de Koning et al. (2009) provided a framework on how cues support selecting, organizing and integrating information. This design principle is implemented by highlighting important information where necessary. Also, it is used by highlighting relevant chunks of diagrams, which could help students interpret them.

Explanations should be **summarised**. The occurrence of summaries might contribute to the effectiveness of science videos (Kulgemeyer, 2020; Kulgemeyer & Schecker, 2013). However, these summaries should be minimal, i.e., free from irrelevant details and coherent (Anderson et al., 1995). In the learning material, this principle is implemented by placing summaries at the end of every section.

The modality principle, temporal contiguity principle, and redundancy principle were not implemented as these principles refer to instructional materials where spoken language is used: the modality principle suggests using the auditive and visual channel; the temporal contiguity principle suggests corresponding information to be presented synchronously, which refer to spoken language to be presented at the same time, and the redundancy principle suggests not to show the same information via different channels. I opted for purely visual material to be more flexible and implement changes more easily. Therefore, no information was presented using the auditive channel. These principles should be implemented in a future version of the learning material, which could include creating an educational video.

Engagement of Learners

The CATLM states that essential processing is fostered by increasing learners' motivation and activating their metacognition, i.e., their knowledge about what they know. This is done by using several different principles.

One of the most basic principles is the **multimedia principle**, which states that students learn better from text and pictures than from text alone. This is implemented by always accompanying explanations with corresponding pictures, usually Feynman diagrams.

To increase learners' engagement with the material, conversational language is suggested, directly addressing the learners and explaining the content from a narrator's perspective. This is stated as the **personalisation principle**. In the learning material, this is realised by using language on the supposed level of learners. Even though no "pedagogical agent" vocalizes the text, an image of an alpaca is included, supposedly narrating the content.

The **guided activity principle** states that learners' generative processing is enhanced by actively engaging with the material, e.g. through interactive questions. This principle is implemented by using quiz questions, which the students answer while working on the learning material.

Learners' metacognitive skills can be activated by fostering active reflection, as stated in the **reflection principle**. In the learning material, as it was used within this project, the students were asked to verbalize their thought processes. The verbalisation was recorded and thus could also be used to complement the eye-tracking data.

The **Feedback principle** states that learners should get explanatory instead of summative feedback. This was implemented in the second iteration of the learning material by adding explanations dependent on the students' answers.

Increasing the accessibility of disciplinary representations

The **coherence principle** states that only necessary information should be presented. This principle is implemented by designing the learning goals progressively (cf. Section 6.2, s.t. at each stage in the learning material, the learners only need a small amount of information simultaneously.

The **pre-training principle** states that necessary terms and definitions must be introduced before the actual content of the learning material. This is implemented by introducing the working mechanism of a Feynman diagram right before this information is used to explain particle physics concepts.

The **pacing principle** emphasises the importance of learners controlling their own pace while going through the learning material. Mayer (2008) formulates it as the "segmenting principle", i.e., that learning material should be presented in "learner-paced segments". This principle is implemented by successively presenting parts of the learning material, broken down according to its learning goals, where learners can continue at their own pace.

Choosing the right level of diagram. Since Feynman diagrams can describe many different processes, there are – in principle, infinitely many – types of diagrams of varying degrees of accessibility. In the following part, the accessibility of the diagrams is examined. An example of a particularly inaccessible diagram would be a so-called *penguin diagram* (Ellis et al., 1977). This type of diagram, an example of which is depicted in Figure 5.1, describes a contribution to the calculation of a transformation between two quark flavours mediated by a W or Z boson via an intermediate quark loop, typically under the influence of the strong interaction. In the example, a bottom quark (notated as "b" on the left) transforms into a strange quark (notated as "s" on the right) under the influence of the strong interaction. This diagram is a very good example of what can make a particular Feynman diagram inaccessible:

- Some lines are vertical; hence, it is unclear whether it's an incoming or outgoing particle – which is irrelevant for a virtual particle, but learners might need clarity about this.
- Some vertices are on top of each other, so there is no clear distinction of what "happens first", - which is not needed for the physics to make sense, but in the explanation of the diagram, this poses a challenge.
- There are three different line types. These stand for quarks (solid lines), W bosons (the squiggly line on top), and gluons (the vertical spring-like line at the bottom).
- The solid lines do not have arrows as in a standard FD. This diagram is also valid if all the particles are anti-particles.
- Some lines have multiple annotations, meaning this diagram considers processes with all possibilities.
- Some annotations are meant for several lines, while one line does not have an annotation.

These elements make this diagram an example of a representation with high disciplinary affordance (cf. Section 4.3), which must be unpacked for new learners to use. Figure 5.2 shows two examples of how diagrams with lower disciplinary and higher pedagogical affordance could look like. These diagrams were used in two different studies (later described in Sections 6.4 and 6.6.2). The diagrams are both made more

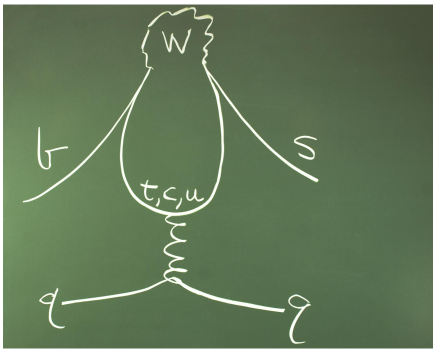


FIGURE 5.1: An example for a *penguin diagram* as introduced by Ellis et al. (1977). Image Source: Marcelloni (2007)

accessible than usual disciplinary depictions, while they also differ from each other in some crucial aspects:

- The charges are annotated following the spatial contiguity principle, meaning learners can work with the diagrams to check for charge conservation without remembering the charges of certain particles.
- In the left diagram, there is only one line type and only one arrow direction to show that time goes from left to right. In contrast, in the right diagram, there are already different line types and arrow directions to distinguish matter particles, antiparticles, and interaction particles.
- In the left diagram, particles are denoted by letters, whereas in the right diagram, particles are denoted by lines. In this aspect, the left diagram violates the coherence principle as the letters are unnecessary for understanding the diagram.
- The left diagram shows a "one-level" diagram, meaning that all initial particles transform into one intermediate particle, which then transforms into the final state. In contrast, the diagram on the right is a "two-level" diagram. Here, the particle in the initial state transforms into one particle from the final state and an intermediate particle, which subsequently transforms into particles from the final state. The diagrams denoted as "one-level" are typical first-order contributions of scattering processes, whereas the "two-level" diagrams are typical contributions to decay processes[1]. Therefore, in the following, diagrams of the left type are referred to as *scattering diagrams*, and diagrams of the right type as *decay diagrams*.

[1]Even though with learners, the term "decay" should be replaced by the term "transformation" (see footnote 2 on page 7), I choose the term "decay" to denote diagrams of this type as this is the technical term for this sort of process. In contrast, transformation would be ambiguous as this term would also describe processes involved with the left diagram.

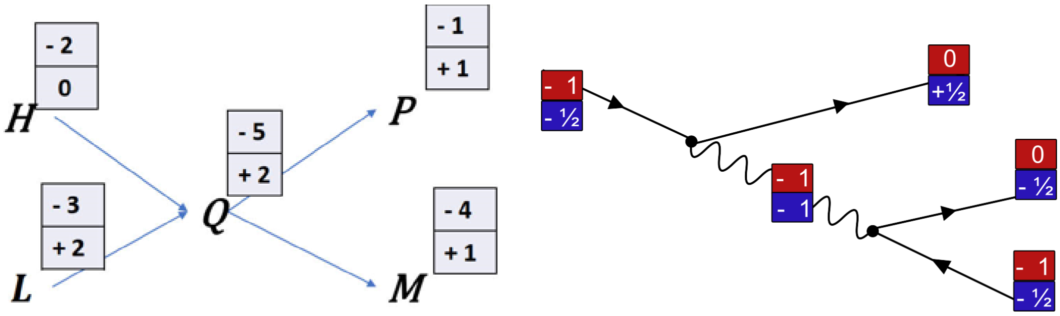


FIGURE 5.2: Two examples for an unpacking of diagrams.

5.2 Eye Tracking

As described in Section 5.1.1, the student studies used to evaluate the learning material were conducted as eye-tracking studies. The different advances made using Eye Tracking in PER are described in Section 4.4. This section describes the technicalities of how the data are taken and analysed.

The eye-tracking studies were used to find students' strategies and cognitive processes when examining the diagrams. The answers to these questions were then used to partly answer the second question posed in the first research goal stated in Section 3.1 as well as the question of the second research goal stated in Section 3.2. The precise research questions are stated when describing the respective studies in Sections 6.3, 6.4, 6.5, and 6.6.

5.2.1 Hardware and Software

The data were taken with a so-called remote eye tracker, which means that the eye tracker was mounted under a computer screen while the study participant sat in front of the screen. There was no head restraint as the participants should get distracted by the eye tracking part as little as possible. In particular, I used a *Tobii Pro Fusion eye-tracker* (Tobii Pro AB, 2019). In the pilot study and the first two student studies, I had one eye tracker at a time, while in the third student study, there were up to four eye trackers in parallel use.

The software used to calibrate the eye tracker and to take the data was *Tobii Pro Lab* (Tobii AB, 2023) in different versions from v1.152 in the pilot study (Section 6.3) via v1.171 in the first (Section 6.4) and second student study (Section 6.5) to v1.217 in the third student study (Section 6.6). The first step of the analysis was also done using Pro Lab, namely defining areas of interest (AOIs) and creating heat maps. All statistical analyses of metrics and transition-based data were done using the *R* software package (R Core Team, 2023).

When using Tobii Pro Lab to take the data, participants only saw the learning material on a computer screen, while the researcher (usually me) saw the participant's gaze on another screen. The participants used their mouse to continue the learning material at their own pace. However, they only saw the cursor when needed to pick an answer from a single-choice question. In studies with think-aloud protocol (see Section 5.2.3 below), the eye-tracking software also recorded their voice.

5.2.2 Task creation, participant recruitment, and data calibration

As eye-tracking data is susceptible to the task participants get to solve, the design of the tasks and the resulting stimuli, i.e., the visuals that participants look at, is a crucial step of the eye-tracking studies presented in this thesis.

The task evolved from study to study, but the principle remained the same: Students were asked to check diagrams for the conservation of charges. This was a plausible task for students as, in the learning material, they were taught about charge conservation. The purpose of this task was then to prompt students to examine the diagram in detail. The exact prompt changed from asking students to check for the conservation of two charges separately to deducing the charges of a particular diagram and finding a diagram with conserved charges. Also, the choice of diagrams varied and is further described in Section 5.1.2 and the respective descriptions of the studies.

The participants were secondary school students aged 15 to 19. The recruitment happened in two ways. In the first and second student study (described in Sections 6.4 and 6.5), the students were at CERN, while in the third study (Section 6.6), they were in school. They were asked to participate in educational research at CERN to make particle physics more accessible to high school students like them. Eye tracking was also mentioned upfront. An example information text used in the first study reads as follows: "*In the study there will be explanatory texts and illustrations about a particular aspect of particle physics and also some tasks based on these explanations. While you read these texts, illustrations and tasks, your eye movements will be recorded. After that, you will see the recordings of your eye movements and you will be asked a few questions about your thoughts and strategies. In the end, you will fill out a questionnaire about how you perceived the activity.*"

Calibration of the eye tracker took place within the software right before the data taking started using a 9-point calibration. This means that participants were asked to follow a dot around the screen, which appeared at nine different places on the screen. The calibration was accepted if the accuracy and precision were better than 1°; otherwise, the calibration was repeated.

5.2.3 Think-aloud protocols

As eye-tracking data often cannot be interpreted on their own but usually needs another source of data together with it, I took verbal data to complement the eye-tracking data. The procedure evolved over the studies: During the first student

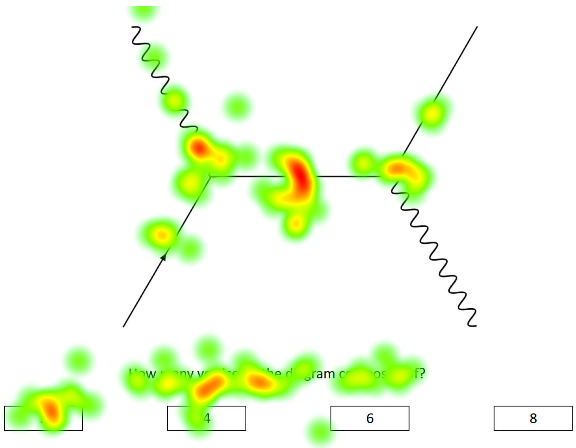


FIGURE 5.3: An example for a heatmap created of data taken in the pilot study described in Section 6.3. Red areas stand for high average attention, and greener areas stand for lower attention.

study, verbal data were taken using both think-aloud and cued retrospective think-aloud, which means that students were asked to verbalize their thought processes while solving the tasks (think-aloud). They were shown a recording of their eye movements afterwards while being asked to verbalize their thought processes again. This double think-aloud was to validate the students' statements with the retrospective think-aloud. However, this method turned out to be impractical since students often did not remember what their thoughts were while solving the task. Also, showing the eye-tracking data did not help; it was instead a source of confusion because they were surprised about how their gaze patterns looked. For this reason, the think-aloud protocol was altered in the second student study. In this study, students could solve the task independently, but directly after they indicated their answer, they were asked why they chose it. In some cases, they revealed a flaw in their original reasoning as they reflected on it, which aligns with the reflection principle presented in Section 4.3.

In the third student study, verbal data were not possible to collect, as eye-tracking data were collected from several participants at the same time, and verbal data would have influenced the other participants. For this reason, the choice was made to omit taking verbal data.

5.2.4 Heat Maps

To obtain a qualitative data picture, so-called *heat maps* were created. These are seemingly simple visualizations of the respective stimuli with an overlayed map of the average attention distribution of the participants. Figure 5.3 shows an example of such a heatmap. The heat map shows how much time a specific subset of participants spent on which areas of a stimulus. The more time is spent on a particular area of the stimulus, the warmer (i.e., redder) the colour in the heat map. It is easy to read more into a heatmap than what is visualised. The only thing that a heat map shows is the spatial distribution of participants' gaze, and there is usually no temporal information and no causal information encoded, e.g. it does not necessarily show which part of the stimulus participants found "most interesting". To give an example of this difficulty in interpreting, the heat map in Figure 5.3 is made from data taken in the pilot study (cf. Section 6.3). The question reads "*How many vertices is the diagram composed of?*" with the correct answer "2" being the leftmost option. The redder areas on the left and the right can be explained as they are referred to in the question, and the participants have learned what vertices are right before. This interpretation would already need some backing, e.g. through verbal data. The "hotspot" on the correct answer option can be explained as most participants gave the correct answer to this question, which is in line with earlier studies which have found that time on correct answer options is related to high response accuracy (cf. Section 4.4.3). The hotspot in the centre, however, is hard to interpret. It could be because there is an arrow on that line, which would also explain the smaller hotspots on the top right and the bottom left lines. However, it could also be an artefact because a so-called fixation cross was directly before that stimulus. Therefore, the first fixations of all the participants were in the centre of the stimulus. This shows that the interpretation of ET data requires knowledge of the context of the data, as well as guidance by theory.

5.2.5 Areas of Interest

The advantage of data visualizations using heat maps is that they do not need any stimulus interpretation. However, as explained in the previous section, the interpretation of the visualization does need to be done.

To obtain quantitative results from ET data, the stimulus must first be interpreted to assign meaning to fixations and saccades. This interpretation is made using *areas of interest* (AOIs) – which are already briefly introduced in Section 4.4.3. An AOI assigns a label to the bare spatial coordinates of gaze points. An example of the definitions of AOIs is given in Figure 5.4. This example is typical for AOI definitions throughout the studies in this thesis, particularly the second and third student study (cf. Sections 6.5 and 6.6). There are usually two levels of AOIs: the *global* and the *local* level. The AOIs on the global level denote large stimulus elements, in the example of Figure 5.4 these are the text on the top left, the alpaca image on the bottom left, the question including the answer options on the bottom and the whole diagram on the right. The AOIs on the local level denote the more detailed elements, such as the single answer options and each charge annotation within the diagram.

The global and local-level AOIs refer to different sorts of hypotheses. The data produced using local-level AOIs are used to test hypotheses referring to representation-specific processes and strategies, while the data produced with global-scale AOIs are used to test general hypotheses, such as processing demands on tasks.

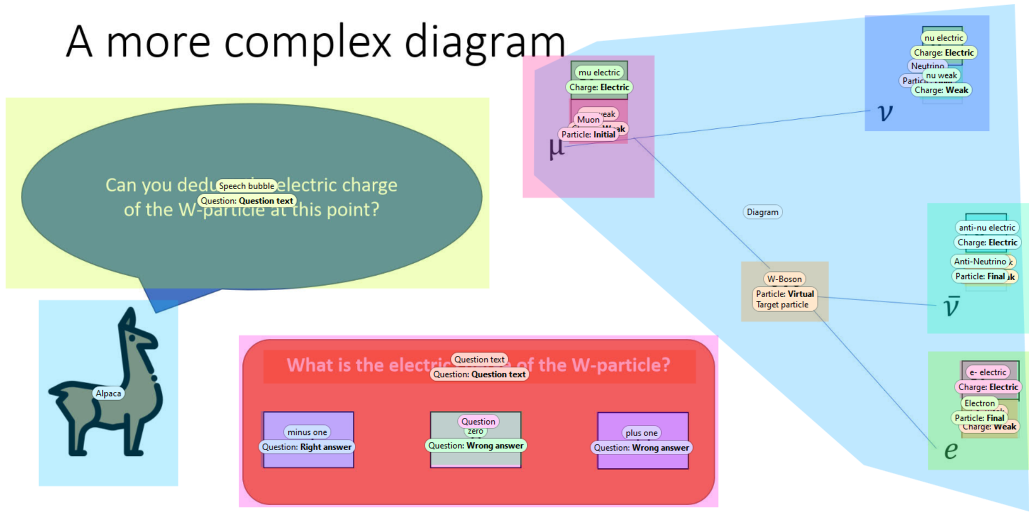


FIGURE 5.4: Definitions of AOIs for a stimulus used in the second student study, further described in Section 6.5. The stimulus has global AOIs, which cover the large areas of the stimulus, such as the question, the diagram, or the alpaca, and local AOIs which cover smaller areas, such as the single answer options, the charge annotations, or the individual particle names.

5.2.6 Fixation measures

As mentioned in Section 4.4.3, there are different fixation measures, i.e., metrics which aggregate fixations within AOIs. Within this project, the most common fixation measure is the total dwell time on an AOI, as this is among the most researched measures. For the sake of simplicity, this measure is usually taken as the total dwell time, i.e., no difference is made between saccades and fixations within an AOI. The dwelling time is then calculated as the total time between the entry into and exit from that AOI. In some cases, an AOI's average fixation duration is calculated as the total time spent on fixations on the AOI divided by the number of fixations within that AOI. The conversion from raw data into fixations and saccades is made within the software using a filter designed by the manufacturer (Tobii, 2023). This filter classifies eye movements with a velocity of under 30 °/s as fixations and faster movements as saccades.

5.2.7 Transition measures

Another class of measures are transition-based measures, which are based on the number of transitions between several stimulus elements designated by AOIs. As mentioned in Section 4.4.3, these measures are usually associated with integrating processes. Transition measures can be calculated based on scan path sequences, i.e., the sequence of fixated AOIs. These measures can be categorised into measures that omit sequential information, i.e., consider the number of transitions between certain AOIs and those that retain at least part of the sequential information. An example

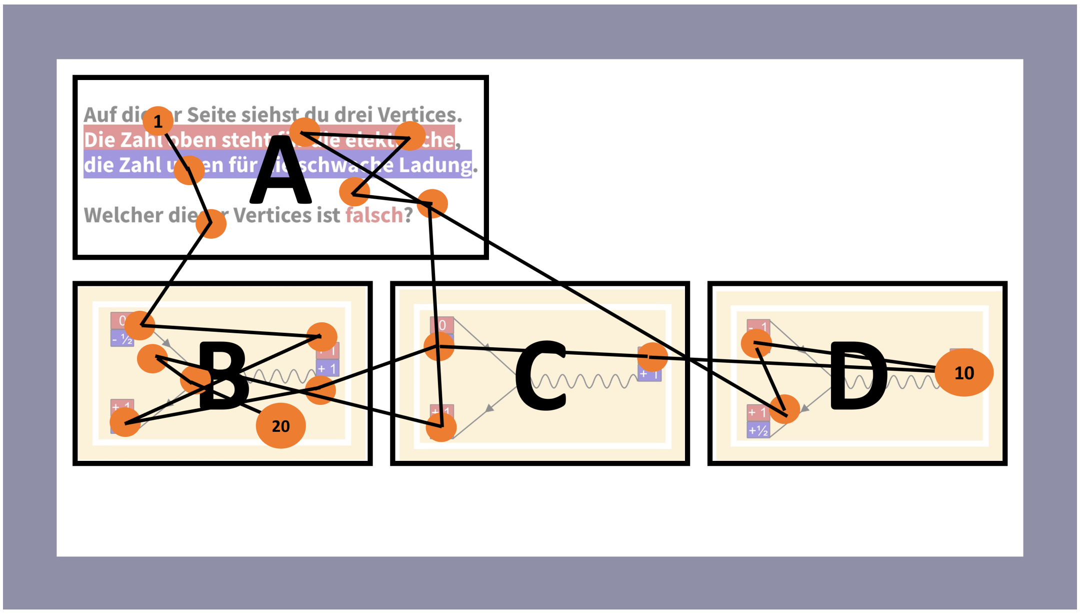


FIGURE 5.5: A mock example of a scan path between AOIs denoted A, B, C, and D. The first, tenth, and twentieth (last) fixation in the example are labelled.

of the former is the fixation/transition ratio FTR, and an example of the latter is the transition entropy H_t introduced in Section 4.4.3.

Fixation/transition ratio

The calculation of the FTR is described by Rodemer et al. (2020) as follows: first, the complete scan path sequence is calculated, i.e., assigning each fixation which happened within an AOI to its AOI. The collapsed sequence is calculated from this, collapsing consecutive fixations within one AOI to one fixation. As an illustrative example, Figure 5.5 shows a mock example of a scan path with twenty fixations spread over four AOIs. The scan path starts at the task description, AOI A, continues across the answer options, denoted as B, C, and D, jumps back to the task description A, and continues through options C and B. The resulting complete scan path sequence would be "AAABBBBCCDDDAAAACBBB", whereas the collapsed sequence would be "ABCDACB". The fixation/transition ratio FTR is then calculated as

$$\mathrm{FTR} = \frac{L_{\mathrm{complete}}}{L_{\mathrm{collapsed}} - 1} \tag{5.1}$$

where L_{complete} is the length of the complete sequence and $L_{\mathrm{collapsed}}$ is the length of the collapsed sequence. In the example, it would be $\frac{20}{6} \approx 3.3$. This value, however, can only be interpreted when compared with others, taken either on the same stimulus by another group of participants or by the same group of participants on a different stimulus.

Transition entropy

The transition entropy is also calculated based on the scan path sequence. Here, the first step is to create a $n \times n$-Matrix $\mathbb{P}$ (with n being the number of AOIs) where in the i-th row and j-th column, the probability of a transition from AOI i to j is notated. The respective probability is calculated as the ratio between the number of transitions from AOI i to j and all transitions starting at AOI i. Also, the stationary probabilities $\vec{P}$, i.e., the probabilities of a fixation happening in AOI i, must be calculated. In the example, these would be:

$$\mathbb{P} = \begin{pmatrix} 0.71 & 0.14 & 0.14 & 0 \\ 0 & 0.83 & 0.17 & 0 \\ 0 & 0.33 & 0.33 & 0.33 \\ 0.33 & 0 & 0 & 0.67 \end{pmatrix} \text{ and } \vec{P} = \begin{pmatrix} 0.35 \\ 0.35 \\ 0.15 \\ 0.15 \end{pmatrix} \tag{5.2}$$

The transition entropy H_t is then calculated according to Equation 4.1; in the example, it would be $H_t \approx 1.04$. To compare the entropy between different stimuli, I normalised the entropy to the maximum entropy which could theoretically be reached on the stimulus. This would be an even distribution over all AOIs, and therefore $P_i = P_{ij} = \frac{1}{N}$ with N being the number of AOIs and therefore $H_{\text{max}} = \log_2 N$ (Krejtz et al., 2015). In the example, this maximum entropy would therefore be $\log_2(4) = 2$ and the normalised entropy $H_{t,\text{norm}} \approx 0.52$.

Since the frequencies of transitions are not the probabilities directly but rather their maximum likelihood estimators, they have errors which are calculated as

$$\sigma_P = \sqrt{\frac{P(1-P)}{N}}. \tag{5.3}$$

with N being the number of fixations based on which the respective P is calculated. Using Gaussian error propagation, the standard error of the transition entropy is thus calculated as

$$\sigma_{H_t} = \sqrt{\sum_i \left(\left(\sigma_{P_i} \sum_j (P_{ij} \log_2(P_{ij})) \right)^2 + \left(\sigma_{P_{ij}} P_i \sum_j \frac{\ln(P_{ij}) + 1}{\ln(2)} \right)^2 \right)}. \tag{5.4}$$

Therefore, the standard error of the transition entropy in the example would be estimated as $\sigma_{H_t} \approx 0.25$ and the standard error of the normalised entropy would be $\sigma_{H_{t\text{norm}}} = \frac{\sigma_{H_t}}{H_{\text{max}}} \approx 0.13$. Thus, the entropy estimation is not very precise if the number of fixations is low. As seen from Equation 5.3, the error is inversely proportional to the square root of the number of fixations, and the precision increases with a larger number of fixations.

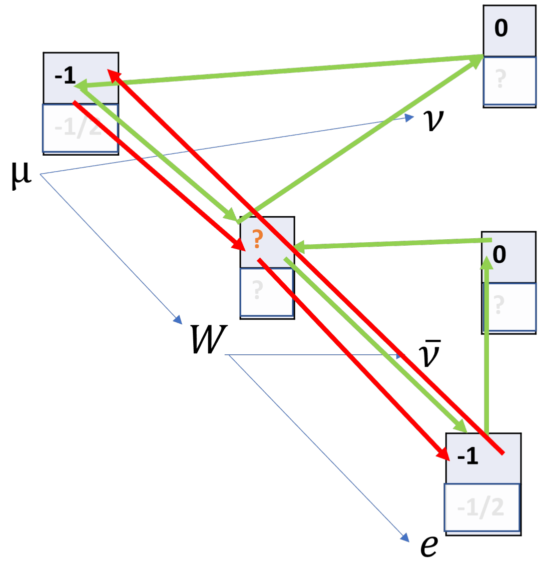


FIGURE 5.6: Example for "meaningful" (green) vs "unmeaningful" (red) triple transitions in the calculation of the triple transition ratio.

Triple transition ratio

In addition to these general measures, I have introduced an ET measure specific to Feynman diagrams, referred to as *triple transition ratio*. This ratio refers to the particular process of examining Feynman diagrams to check for charge conservation. To calculate this ratio, I defined "meaningful triple transitions", i.e. transitions between three charge annotations in which one always has to be the sum of the two others. In the example in Figure 5.6, the green transitions are meaningful as they compare three charges, which need to be compared to check the diagram for charge conservation. In contrast, the red transition shows a path which is not productive in solving the task. A productive viewing pattern should manifest in a high percentage of these meaningful triple transitions compared to all triple transitions (with a "triple transition" being defined as two consecutive transitions between three charges).

5.3 Instrument design

Verbal data and questionnaires accompanied the data collected during the studies. While the collection of verbal data is further described in Section 5.2.3 and the respective sections in the descriptions of the studies (Sections 6.4, 6.5, 6.6), in this section the design of the questionnaires is described.

5.3.1 Prior knowledge

Since one of the most effective predictors of learning outcomes is prior knowledge, assessing what learners already know is essential. For this purpose, research-based assessment instruments (RBAIs) play a similar role in educational research as the “physical” measuring instruments in research experiments. These tests cover a wide range of contents in various areas of physics (mechanics, electricity and magnetism, thermodynamics, optics and waves, quantum mechanics, astronomy, and more). A comprehensive overview of research-based assessment instruments is given by American Association of Physics Teachers (2024) and Madsen et al. (2017), Madsen et al. (2019). Moreover, there is an advanced discussion and widely used methodology for the development and validation of RBAIs in physics and astronomy education (see e.g. Adams and Wieman (2011)).

However, no RBAI has been developed so far, covering a broader range of concepts in elementary particle physics. There are tests about areas with partial overlap, such as radioactivity (Molz, 2016; Prather & Harrington, 2001; Woithe, 2020). There have been two approaches to assessing students' knowledge in particle physics, both mentioned in Section 2.2.1, by Gourlay (2016) and Tuzón and Solbes (2016). While Tuzón and Solbes (2016) used a questionnaire with open-ended answers, Gourlay (2016) used concept-mapping to assess students' conceptions.

Even though concept mapping is a promising tool to assess conceptions about particle physics (cf. Kranjc Horvat (2022)), it was unsuitable for my studies for research-practical reasons. The assessment had to be included in an eye-tracking study, which is time-consuming, and the time frame was about 45-60 minutes per participant. At the same time, a valid concept map would require students to be familiar with the method and to have enough time to create the map (in the study by Gourlay (2016), it took a whole physics lesson of 90 minutes).

In the first student study, the questionnaire to assess prior knowledge consisted of whether students had previously learned about the contents. This was done to save time during the study. Still, it yielded unreliable results, as students often were unsure whether they had learned about it, and even if they had, it would have been unreliable information about their prior knowledge.

Therefore, for the second student study, I built on the work of Tuzón and Solbes (2016) to create a questionnaire that could serve as a basis for developing an assessment instrument. Table B.1 shows which questions I used. I did not use the whole questionnaire because the questions should have been mapped to topics covered in the learning material. As done in the original study by Tuzón and Solbes (2016), the questionnaire was administered as an open one, meaning the students could enter free-text answers. The answers were eventually coded based on the coding used in the original study, which is given in their supplemental information. Students could get one, a half or zero points per answer. The coding scheme used in this study is shown in the last column of Table B.1.

For the third student study, I developed a new questionnaire with multiple and single-choice questions, which was only loosely based on the previous questionnaire. The questions are based on the topics covered in the learning material. In this case, the questions were posed as multiple- or single-choice. The questions, answer options, and their rationale can be found in Table B.2. The first question was particular because it did not ask for knowledge but which particles one "has heard of". The scoring of this question was done in a way that "basic knowledge", i.e., the proclaimed knowledge of only the "basic particles" electron, proton, and neutron, gave 1/3 of a point, "partial knowledge", i.e., the proclaimed knowledge of particles other than the basic particles gave 2/3 of a point, and "expert knowledge", i.e., the proclaimed knowledge of all the mentioned particles, gave one entire point. All other questions gave one point for a correct answer and 0 points for an incorrect answer.

The only question common to both questionnaires was the question "*What is CERN?*", which was used to find and compare different student demographic's conceptions about the nature of the European Organisation of Particle Physics. The comparison between the two groups was particularly interesting because the students in the second student study were physically at CERN, while students in the third study were at their schools in Geneva and Göttingen. However, in the third study, this question was a mere additional question and did not account for the final score of the test instrument.

In the third student study, one of the research goals was to assess conceptions about particle interactions. These were evaluated using open comprehension questions since the conceptions should be assessed without introducing a potential bias by answer options. These questions are presented in Section 6.6.4 and Table 6.7.

5.3.2 Cognitive load

The cognitive load was assessed in all three studies using an instrument presented by Leppink et al. (2013), which considers the three dimensions of cognitive load as given in Section 4.3: extraneous, intrinsic, and germane cognitive load. However, since the questionnaire in that study was developed to assess the cognitive load of university students while learning statistics, it had to be adapted, both in the wording of the respective topic and the exact wording of a question. For example, item 5 in the original questionnaire reads, "The instructions and/or explanations were, in terms of learning, very ineffective.". This item might seem very complicated for a high school student. It was thus simplified to "I did not learn a lot from the explanations." Even though the new formulation might not carry all the nuances of the meaning of the original question, the priority lay on understanding the question. Another difference to Leppink et al. (2013) was that I used a five-point instead of an eleven-point Likert scale, as it was suggested during the pilot study that ten points seemed overwhelming for study participants. For the second study, some items were altered to match the original items better. For example, item 6 in the original version reads, "The instructions and/or explanations were full of unclear language." and was simplified in the first study to "I had to read the explanations more than once to understand them." This phrasing, however, could also relate to intrinsic cognitive load, i.e., the difficulty of the topic, which is why it was phrased back to better match the original phrasing.

Due to its shortcomings, I changed the instrument to assess the cognitive load in the third student study for an instrument developed by Novak et al. (2018) to assess the cognitive complexity of e-textbook learning. This instrument is based on an extension of Leppink et al.'s original questionnaire (Leppink et al., 2014), which I was not aware of. However, Novak et al.'s questionnaire only assessed extraneous and intrinsic load, so I added the items assessing germane cognitive load from the previous questionnaire to the new one. The questions can be found in Table B.4. Furthermore, I changed the items to simplify the precise wording, for example, items starting "I invested a very high mental effort..." were changed to "Ich fand es sehr mühsam...", which instead translates to "I found it very cumbersome to...". Also, occurrences of "e-text" were changed to "material".

5.3.3 Motivational/Affective factors

As stated by the CATLM (cf. Section 4.3), motivational factors play an important role when learning (with multimedia, but also in general). To assess motivational factors, I used an instrument which was developed to "measure reactions to self-directed instructional materials" (Keller, 2010, p. 277) and is based on the ARCS model of motivation. This model was developed by Keller (1987) to find effective ways to influence students' motivation to learn. It describes four factors that must be managed for learners to become and remain motivated: Attention, Relevance, Confidence, and Satisfaction. The first factor, attention, is a prerequisite for learning (as it is a central factor also in using eye tracking, cf. Section 4.4). While it is easy to draw attention to a particular stimulus, it is hard to sustain it. Relevance does not necessarily come from a subject matter itself, and it can also come from how something is taught. Confidence is closely related to what Dweck (2006) coined the term "growth mindset" for. Confident learners – or learners who exhibit a growth mindset – attribute their successes to ability or effort instead of luck, they believe that they can achieve their goals through their actions, while unconfident people have a constant fear of failing. Satisfaction refers to learners wanting to feel good about their accomplishments.

Keller (2010, p.282 f.) presents the "Instructional Material Motivation Scale" which measures motivation in these four subscales. The original survey had 36 items answered on a five-point Likert scale. I used an adapted version of this instrument in the third student study. Since the time frame was minimal, I used a short version of the questionnaire, which was used by Novak et al. (2022) to assess students' motivation with e-textbooks. This adapted instrument had 16 items with four items per dimension (attention, relevance, confidence, and satisfaction). It had to be slightly adapted by substituting "e-text" with "material". The items in their German translation as they were used in the third student study and the original from Novak et al.'s instrument are shown in Table B.5 together with the correspondence to the four dimensions.

5.4 Statistical analysis

5.4.1 Item analysis

Questionnaires were used to assess three categories of data during this project: prior knowledge/learning outcomes, cognitive load, and motivational/affective factors. The validity and reliability of a questionnaire should be assessed when using an instrument. The following chapter evaluates the reliability in all instances by providing a measure for internal consistency through *Cronbach*'s α. As a rule of thumb, the internal consistency should be $\alpha \gtrsim 0.8$ for a reliable test (Bortz & Doering, 2016, p. 198f.). A high internal consistency on the single scale would be desirable, especially when an instrument should assess single constructs, which is the case for the cognitive load questionnaire and the instructional motivational survey.

Another method to assess internal consistency, but for tests with multiple-choice items, i.e., when an item is either "right" or "wrong", is the Kuder-Richardson index (KR-20) which is calculated as (Ding & Beichner, 2009)

$$\text{KR-20} = \frac{K}{K-1}\left(1 - \frac{\sum_i^K p_i(1-p_i}{\sigma^2}\right), \tag{5.5}$$

where K is the number of items in the test, p_i is the average score in the ith item, and σ is the standard deviation of total test scores. Ding and Beichner (2009) suggest that KR-20 should be larger than 0.7 for the test to be considered reliable for a group measurement and larger than 0.8 to be reliable for assessing individuals. Contrary to that, Adams and Wieman (2011) suggest that the internal consistency of a test instrument, measured by α or KR-20, can also be lower, as it does not necessarily claim to measure a single construct.

A test instrument should furthermore have discriminatory power, i.e., it should cover a wide range of scores to discriminate between high- and low-ability students. This discriminatory power is measured by Ferguson's δ, which is given by (Ding & Beichner, 2009)

$$\delta = \frac{N^2 - \sum_i^K f_i^2}{N^2(1-\frac{1}{K+1})}, \tag{5.6}$$

where N is the total number of students and f_i is the frequency with which the ith score appears. This formula is only valid if each item gives either 1 or 0, i.e., the number of possible scores equals the number of items. As seen in this section, some of the items of the test instruments were also scored with half- or third-integer values. In this case, the students' total scores were rounded to the nearest integer value to compute δ.

Besides an analysis of the complete test instrument, there are also different methods for analysing a test instrument on a per-item basis. These methods are comprehensively reviewed, e.g., by Ding and Beichner (2009). Within this work, I will follow an item analysis based on classical test theory, as the purpose of this analysis is merely to evaluate the items and reliability of the instrument. Further study of the instrument would need a more sophisticated theory, like item response theory, which also analyses the latent abilities of the subjects. For this analysis, however, a more extensive data set would be needed than I obtained within the studies presented in this work.

Item analysis is typically done by computing three different coefficients: item difficulty p, discriminatory index D, and point biserial coefficient r_{pbi}. The item difficulty is measured by the average score obtained on an item (if the items are scored with either 1 or 0, this equals the proportion of students who answered correctly). Adams and Wieman (2011) suggest having a mix of easy and difficult items while Ding and Beichner (2009) propose for p to be in a range between 0.3 and 0.9. The discriminatory index D is calculated as the difference in average scores between students with high ability and those with low ability. Usually, the distinction between high and low ability is made by their total score, so it could be the average score of those in the upper and lower quartiles. Ding and Beichner (2009) suggest that the discriminatory index be larger than 0.3 for the item to have some discriminatory power. The third coefficient of item analysis is the point biserial coefficient r_{pbi}, which is similar to the discriminatory index but discriminates low- and high-ability students on the individual item level. It measures the correlation between those who answered correctly to an item and those who answered incorrectly. It is calculated as

$$r_{pbi} = \frac{\overline{X_1} - \overline{X_0}}{\sigma} \sqrt{p(1-p)}, \tag{5.7}$$

where $\overline{X_1}$ is the average total score of those who answered an item correctly and $\overline{X_0}$ that of those who answered incorrectly, σ is the standard deviation of total test scores, and p is the difficulty of the item in question as defined above. If items are scored with partial points, a choice must be made as to whether they count as correctly or incorrectly. This work usually counts partially correctly answered items as incorrectly in the r_{pbi} computation. Ding and Beichner (2009) suggest for r_{pbi} to be larger than 0.2 for an item to be reliable. Adams and Wieman (2011) argue that an item could have a low r_{pbi} if this item considers a concept that students learn much better or worse than the other concepts, which could be desirable for a test instrument.

5.4.2 Significance analysis

To analyse the data quantitatively, I used statistical methods found in any textbook about research methods in social sciences (I used a German standard textbook by

Bortz and Doering (2016)). These are only briefly introduced here.

A **hypothesis test** establishes a formal approach to assess claims about a population. In this approach, a *null hypothesis* H_0 stating no significant association between variables, and an *alternative hypothesis* H_1 proposing the opposite is formulated. Statistical tests evaluate the evidence against the null hypothesis, allowing one to reject it (supporting H_1) or fail to reject it (inconclusive results).

The concept of **statistical significance** quantifies the likelihood that an observed effect is not merely due to chance. A standard measure is the *p-value*, representing the probability of obtaining an outcome as extreme or more extreme than what we observed, assuming the null hypothesis is true. A low p-value, typically below 0.05, indicates a statistically significant result, strengthening the evidence against the null hypothesis.

The hypothesis test usually used within this work is the **t-test**. This statistical method specifically addresses comparing the means of two groups. It utilizes the *t-distribution* to assess the probability of observing a difference between the means as extreme or more extreme than what we observed, assuming no actual difference exists between the populations (null hypothesis). A low p-value from a t-test indicates a statistically significant difference between the group means. A t-test can be performed on an unpaired or a paired sample, the latter typically being the case on a test in a pre-post design where one student's value in the pre-test is compared to the same student's value in the post-test. When comparing two groups with unequal variances, the traditional t-test becomes unreliable. *Welch's correction* addresses this issue by adjusting the t-test statistic to account for unequal variances, providing a more accurate assessment of statistical significance. The significance of a test result is then tested by comparing the t-test statistic with a distribution that depends on the *degrees of freedom* of a test statistic. The values of this distribution can be found in textbooks or are programmed into statistical software packages. Reporting a t-test result is done in the form $t(\text{df}), p$, where df is the number of degrees of freedom, t is the statistic calculated by the t-test, and y is the resulting p-value. In the simple t-test, which assumes equal variances of the mean, df is $n - 1$ where n is the number of observations. In contrast, the t-test with Welch correction uses the Welch-Satterwaithe equation to calculate df (Welch, 1947). This formula is programmed into the standard statistical software packages and, therefore, not reproduced here.

To quantify the magnitude of an effect, the typically used measure is **Cohen's d**. This metric serves as a measure of *effect size*, complementing statistical significance. It expresses the magnitude of the difference between two groups normalised by their standard deviations. Common interpretations for Cohen's d are:

- Negligible effect: $d < 0.2$
- Small effect: $0.2 \leq d < 0.5$

- Moderate effect: $0.5 \leq d < 0.8$
- Large effect: $d \geq 0.8$

If the means of several groups are compared, the standard method is the **Analysis of Variance (ANOVA)**. It partitions the total variance in the data into components attributable to between-group differences and within-group variations. Analyzing this variance distribution, ANOVA helps determine if the observed differences among groups are statistically significant. An ANOVA result is reported as $F(\text{DFn}, \text{DFd}), p, \eta^2$, where DFn is the number of degrees of freedom between the groups and DFd is the number of degrees of freedom within the groups, F is the F-statistic calculated by the ANOVA, p the resulting p-value, which depends on the F-distribution specific to the degrees of freedom, and η^2 the proportion of total variance which is explained by the differences between the groups. The ANOVA is typically accompanied by a post-hoc test to evaluate the significance of differences between the groups.

To quantify the linear relationship between two continuous variables, **Pearson's correlation coefficient** R is used. It ranges from -1 (perfect negative correlation) to +1 (perfect positive correlation), with 0 indicating no linear association. The strength of the association is interpreted based on the absolute value of the coefficient. It is calculated for all scatter plots and displayed as part of the plots. To visualise the correlation, a regression line is drawn on top of the data. Pearson's R is also shown in some boxplot diagrams where two groups are compared. In these cases, however, additional t-tests are performed.

Chapter 6

Outcomes

6.1 Expert study

To determine possible opportunities and challenges Feynman diagrams could pose when using them to teach particle physics to upper secondary students, I conducted interviews with experts in particle physics education. This expert interview study was published as part of a special issue on "Teaching and Learning Quantum Theory and Particle Physics" in the journal MDPI Physics (Dahlkemper et al., 2022) and is reprinted below. The supplemental material, i.e., the interview guides and the coding manual, are presented in the appendix chapter C.

physics

MDPI

Article

Opportunities and Challenges of Using Feynman Diagrams with Upper Secondary Students

Merten Nikolay Dahlkemper [1,2,*], Pascal Klein [2], Andreas Müller [3], Sascha Marc Schmeling [1] and Jeff Wiener [1]

[1] European Organization for Nuclear Research (CERN), 1211 Geneva, Switzerland
[2] Physics Education Research Section, Faculty of Physics, University of Göttingen, 37077 Göttingen, Germany
[3] Faculty of Sciences, Section of Physics and University Institute of Teacher Education (IUFE), University of Geneva, 1211 Geneva, Switzerland
* Correspondence: merten.dahlkemper@cern.ch

Abstract: Particle physics is an exciting subject for high school students, and there have been various approaches on how to introduce the topic in the classroom. Feynman diagrams (FDs) are an often-used form of representation in particle physics and could play an important role in such an introduction. However, their potential educational value has not yet been investigated. To this end, we interviewed four experts in the field of particle physics education on the opportunities and challenges Feynman diagrams could pose for high school students. We analyzed their answers using a thematic analysis framework, categorizing them into five themes. The results of these interviews show that there are two challenges (FDs elicit and perpetuate inadequate conceptions about particle physics, and FDs can only be treated superficially in school) and three opportunities (FDs can link particle physics and other physics topics in high school education, FDs offer an opportunity for different particle physics topics to be taught, and FDs offer a connection to current research). The results of this expert interview study lead to several suggestions on how to design learning environments that incorporate Feynman diagrams.

Keywords: particle physics; Feynman diagrams; secondary education; expert interviews; design-based research

Citation: Dahlkemper, M.N.; Klein, P.; Müller, A.; Schmeling, S.M.; Wiener, J. Opportunities and Challenges of Using Feynman Diagrams with Upper Secondary Students. *Physics* **2022**, *4*, 1331–1347. https://doi.org/10.3390/physics4040085

Received: 24 August 2022
Accepted: 18 October 2022
Published: 28 October 2022

1. Introduction

Particle physics has been introduced in high school physics curricula in many countries. When teaching about particle physics, an often-used representation is that of so-called Feynman diagrams (FDs). However, while being of great use for particle physics, it is unclear whether using FDs in education benefits learners, since possible inadequate conceptions could be connected to them. The current study explores potential opportunities FDs can create for learning particle physics and the challenges that using these diagrams pose. To meet these aims, we interviewed experts in particle physics education.

1.1. Particle Physics in High School Education

The first calls to introduce particle physics in high school physics teaching date back to the 1980s [1], with one of the first conferences on the teaching of modern physics taking place at the European Organization for Nuclear Research (CERN) in 1984 [2]. Students might see particle physics as an exciting topic, which can serve as an example of the nature of physics as being tentative and constantly changing. Moreover, unknown phenomena and the newest scientific discoveries have been proven to be extremely interesting topics for adolescents [3]. Finding answers to the fundamental questions of nature can positively influence students' attitudes toward physics. The call for particle physics in high school education has since been replicated several times (e.g., [4–6]). Despite these calls, particle physics only rarely makes it into the curricula and subsequently into the teacher training courses at university. However, several suggestions have been made on how to incorporate particle physics into high school education (e.g., [7–10]). These suggestions vary greatly in the concepts they include and the approaches they take to introduce these concepts to high school students. They range from an overview of the "particle zoo" [10] to discussions of particle interactions, conservation laws, and symmetries [11,12] to coherent learning material for teaching the Standard Model of particle physics by connecting charges, particle interactions, and elementary particles [9].

Some approaches have been criticized as presenting the topic of particle physics in a superficial and oversimplified way [13,14]. Furthermore, presenting all the elementary particles in a merely enumerative manner might not be very motivating and might lead to rote learning of particle names instead of achieving a conceptual understanding of particle physics [15].

So far, few studies have been conducted to systematically investigate students' understanding of particle physics or the impact of specific interventions in the field [6,16,17].

1.2. Feynman Diagrams

A specific topic within particle physics is that of Feynman diagrams. The idea of FDs was first presented publicly by Richard Feynman at a conference in the spring of 1948 [18] and was subsequently published by Freeman Dyson and Feynman [19,20]. Since then, they have become a commonly used graphical tool in many areas of theoretical physics, but especially in quantum field theory and, thus, particle physics. In fact, they have been invented mainly as a tool for "bookkeeping" and visualization in the perturbation-theoretical calculation of processes in particle physics [18]. However, there is a debate about whether or not FDs are a form of representation of physical processes [21–23]. This debate dates to a disagreement between the inventor of the diagrams, Feynman, and the physicist who connected the diagram to a set of mathematical rules, Dyson. This disagreement is also known as the Feynman–Dyson split [18].

FDs are also part of many introductions to particle physics for high school students (e.g., [7–9]). However, Passon et al. [24] highlight various ideas associated with FDs that are present in educational material, which can be considered to be misleading. For example, any reading in which particles are assigned a trajectory in time and space is, according to the authors, physically nonsensical. Various physicists and physics educators have attempted to give educational introductions to FDs. For instance, Pascolini and Pietroni [25] introduced the diagrams as "accurate metaphors" and attempted to introduce the rules underlying Feynman diagrams to learners using a mechanical model. This approach was empirically investigated in a post-test one month after the intervention, where the learners performed satisfactorily. Hoekzema et al. [11,12] used a reduced form of FDs, which they call "reaction diagrams", to explain conservation laws and symmetries in particle physics. Their proposed teaching activity consists of two 50 min lessons embedded in a larger teaching project about nuclear reactions and elementary particles. The activity had been used for two years as part of a modern physics teaching project. In particular, the approach with the reduced form of the FDs was deemed more comprehensible than a previous text with conventional FDs.

Figure 1 shows a typical example of an FD. Explanations for this type of diagram differ in the extent to which it is explained 'literally'. "Literally" means, in this context, that the diagram is understood as a particle and an antiparticle traveling in straight lines toward each other, meeting and creating a photon, and, at some point, a new particle-antiparticle pair flying away from each other. While some parts of these explanations are compatible with the physical interpretation of the diagrams (for example, particles annihilate each other and new particles are created), others are rather inadequate conceptions (for example, particles do not have determined trajectories). In some texts, FDs are introduced as space–time diagrams [26], while in others, the individual fundamental vertices are explained and linked to mathematical terms in order to stress the inherent mathematical meaning of the diagrams [27]. Several authors try to avoid the problem of a too literal reading by emphasizing the fact that FDs contribute to a probability amplitude [7,28,29].

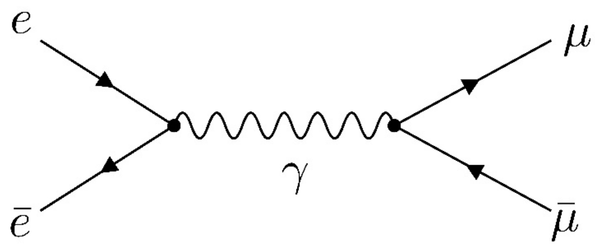


Figure 1. An example of a Feynman diagram of electron–positron ($e\bar{e}$) annihilation. The diagram describes the transformation of an electron–positron pair into a muon–antimuon ($\mu\bar{\mu}$) pair. The interaction is mediated by a photon (γ) which is the interaction particle of the electromagnetic interaction. The direction of arrows determines whether it is a particle or an antiparticle.

A defining property of every FD is that charges are conserved throughout the diagram, meaning that the total amount of any charge (electromagnetic, weak, and strong charge) is the same on the left and on the right side of the diagram. In Figure 1, this can be straightforwardly checked for the electric charge, since the electron and the positron on the left have a total electric charge of zero, which is also true for the muon and the antimuon on the right side.

Different tools have been developed to teach the underlying principles of FDs, most notably the conservation of charges. For example, *FeynGame* is a software that can be used to draw FDs and check them for correctness [30]. Another example is the educational card game *Feynman-Rhombino*, which allows players to create FDs by applying the rules of charge conservation [31].

1.3. Context of the Present Study: Design-Based Research

As illustrated above, several proposals suggest how FDs can be used for educational purposes. However, to our knowledge, no empirical study has investigated the suitability and feasibility of FDs for use in the physics classroom. Therefore, the overarching aim of this research project is to iteratively develop a learning environment that introduces FDs to upper secondary students (16–18-year-olds) and to test the educational use of Feynman diagrams. To do so, we use the framework of design-based research (DBR) [32].

DBR has been used to create learning environments in various fields of modern physics, for example, for the topics of chaotic systems [33], general relativity [34], special relativity [35], and, in particular, particle physics [17].

The DBR framework, also known as design research, design experiments, or educational research design, has been widely used and has resulted in learning interventions that create improved outcomes or student attitudes among the students [36,37]. There are many different approaches to DBR, but a central element is the iterative design and testing of the material. DBR produces both new theories and teaching practices to impact learning [38]. In this way, DBR is a method designed to bridge the gap between research and practice.

A particularly interesting link can be created between DBR and the model of educational reconstruction [39]. In this model, the educational design is informed both by the content structure of the subject matter and by the perspectives of the learners. Analysis of the content structure leads to the formulation of key messages, whereas that of the learners' perspective identifies relevant prior knowledge and potential learning difficulties of the respective subject matter. These two perspectives influence each other in an iterative process. Therefore, the reconstruction of the subject matter does not necessarily follow the content structure of the subject matter. Both the analysis of the content structure and the learners' perspective can be informed by different sources such as experts, literature, and the students themselves.

1.4. Scope of the Study

Building on the existing work about how (not) to incorporate FDs in educational material about particle physics, the current study presents a systematic analysis of the challenges and learning opportunities that are connected to the teaching of FDs. This analysis is the basis for an educational reconstruction of FDs, which is currently undergoing iterative testing with high school students. The research question that guided this work was: What challenges and opportunities do FDs pose for high school students?

2. Materials and Methods

To answer the research question, explorative interviews were conducted with experts in the field of particle physics education [40]. The aim was to determine and summarize the experts' opinions on what purpose FDs could serve in upper secondary school education.

2.1. Selection of Experts

The experts were selected based on their expertise in conveying particle physics to high school students.

We selected three researchers from two German universities who are involved in education programs in particle physics aimed at high school students. Two of them are education researchers, and one is a particle physicist. All three experts have previously published on the educational use of FDs.

After conducting interviews with these three experts and discussing the results among the research team, a discussion point regarding the epistemological role of FDs was raised. Therefore, a science philosopher was interviewed as fourth expert. This expert had expertise in the philosophy of science and conceptual analysis of contemporary physics.

The present study is exploratory in that, for the first time, it gathers data for the challenges and opportunities of integrating important content of contemporary physics into high school physics teaching. While the sample size is small, the results provide evidence for considerable and informative variance on the one hand and increasing overlap between the answer sets on the other hand. This is well in line with the main purpose of the present study, namely an in-depth insight about potential challenges and opportunities to guide further development, and not an exhaustive treatment of the topic (see [41,42]).

2.2. Conducting the Interviews

The first author conducted the interviews in a semi-structured way, that is, the interviewer had an interview guideline with the core questions, but he could diverge from it depending on the answers the interviewees gave.

The interview guideline was constructed based on two guiding questions, which resulted from the research question stated above. These were:

- What challenges are connected to teaching FDs to high school students?
- What opportunities for physics education at the high school level are provided by FDs?

The interview guideline, which is published in the Supplementary Material, contained 12 questions, which were divided into three parts: In the first part, the interviewees described what FDs are. This part served as an entry point to the interview and was not connected to one of the guiding questions. The second part was concerned with possible learning obstacles the experts could imagine or had encountered. This included inadequate conceptions students might already have, or what prior knowledge is needed to read and draw FDs. The third part was connected to the second guiding question and focused on potential learning opportunities offered by FDs. In this part, it was asked, for example, how an appropriate instruction would look like and how students could benefit from learning about Feynman diagrams. In the end, the experts were asked how they would describe a particular Feynman diagram to a high school student.

Special emphasis was put on appropriateness for the target group. For example, concerning the question of what a Feynman diagram is, the experts were specifically asked to describe it for a 17-year-old high school student. Similarly, when asked about potential learning obstacles, the experts were asked about learning obstacles specific to high school students.

The guideline for the interview with the fourth expert was adapted to his expertise, which differed from the expertise of the other interviewees. Specifically, the original guideline was shortened, and the following three introductory segments were added to the interview guideline to better align the results of the fourth interview with the first three interviews: The first segment was intended to clarify the definition of terms such as "model" and "theory". The second segment touched upon the connection between models and reality, since one of the open discussion points concerned the epistemological meaning of FDs. In this part, the expert was asked, for example, to which extent models play a role in science and what role mathematical descriptions play for models. The third segment focused on the role of modern physics in schools, since the second open discussion point after the first three interviews led to the question of whether modern physics—or particle physics in particular—should be taught in high school. In this part, the expert was asked, for example, which aspects of the nature of science he sees as vital for a meaningful discussion of science in a high school classroom.

The interviews were conducted in German via a videoconferencing software and were subsequently transcribed verbatim by the interviewer. The interviews were scheduled to take one hour and lasted for 63, 81, 37, and 81 min. The third expert mentioned during the interview that he did not feel completely well, which led to shorter answers and a faster interview pace, resulting in a shorter interview length.

2.3. Coding Scheme

The interview transcripts were categorized according to the thematic analysis framework [43]. The coding scheme contained two dimensions based on the guiding questions, that is, opportunities and challenges of FDs. These dimensions subsequently served as the highest level of the category system.

The transcripts were scanned for segments referring to these dimensions. Only those segments that answered one of the two guiding questions were categorized. Segments were of different lengths, depending on how much context was given within one segment. The shortest segments consisted of one sentence and the longest of about ten sentences. The first author inductively categorized these segments with preliminary category descriptions. The segments were categorized at the least general level possible, based on their thematic similarities. After categorizing approximately 30% of the transcripts, the categorization was discussed with the co-authors and revised, that is, categories were added, renamed, or merged. Based on the revised categorization, the transcripts were categorized in full. The categories were then subsumed into common themes and subthemes. Some subthemes also had sub-subthemes. Since, depending on the number of segments in a theme, the division into themes and (sub-)subthemes is different, the lowest level is generically called categories in the further description. Further on, categories were dissolved, and the segments were categorized into other existing categories. Through this process, the number of categories was reduced. This categorization was subsequently reviewed while the categories were given instructive descriptions to be incorporated into a coding manual. This coding manual was then discussed with two German-speaking doctoral students in the field of physics education who were not previously involved in the research process. Based on this discussion, the manual was again adapted by redefining category descriptions and merging categories to further reduce their number.

2.4. Validating the Coding of the Data

The revised coding manual was eventually given to the two doctoral students who used this coding manual to code approximately 10% of the coded text segments. This coding manual had 37 categories categorized into five themes. Interrating yielded a Krippendorff's alpha [44] of 0.49. This interrating was subsequently discussed between the first author and the two interraters. Interrater reliability after the discussion was increased to an acceptable level of 0.87. During this interrater discussion, several remarks about the coding manual were made, which led to another revision of both the coding manual and the coding of categories. Disagreements in the coding were caused by giving too little context within segments, which is why some segments were extended to give more context. The final coding manual is shown in the Supplementary Material. It is divided into two dimensions (challenges and opportunities), of which the first dimension contains 12 categories divided into two themes and four subthemes. The second dimension contains 12 categories divided into three themes. A second round of interrating was conducted with a new set of segments and a new interrater who had not been involved in the process before. This interrating showed an initial Krippendorff's alpha of 0.37, and after another interrater discussion and a revised coding, the interrater reliability showed 100% agreement.

3. Results

In total, five themes emerged from the analysis of all four expert interviews, which in turn were assigned to the two dimensions "challenges" and "opportunities". Two themes concerned challenges and three themes concerned opportunities. Of the challenges, the most prominent theme was that FDs elicit and perpetuate inadequate conceptions, which is in line with prior work (see Section 1.2). The second theme was that FDs can only be discussed superficially when discussing them with high school students. In contrast, among the opportunities, the following three themes became apparent: (a) FDs offer the opportunity to discuss perspectives on science, (b) FDs offer a link between particle physics and other high school physics topics, and (c) FDs can be a hook for different particle physics topics, such as the conservation of charges in particle transformation processes.

These five themes are explained further in the following text using subthemes and example segments from the transcripts. The segments presented here are translated from the original. The experts are referred here to as E1, E2, E3, and E4 in the following. E1, E2, and E3 are the first three experts, while E4 is the science philosopher who was later interviewed.

3.1. Challenges

Table 1 shows the two themes and corresponding subthemes present within the dimension of challenges. The subthemes are also further divided into sub-subthemes. Moreover, not all the themes were mentioned by all the experts. Table 1, therefore, shows which of the four experts mentioned which sub-subtheme.

Table 1. Themes and subthemes of the dimension challenges. The order of the (sub-)subthemes within a (sub)theme is according to the frequency of mentions.

Theme	Subtheme	Sub-Subtheme	Mentioned by Experts (E1, E2, etc.)
Feynman diagrams (FDs) elicit and perpetuate inadequate conceptions	Types of inadequate conceptions connected to FDs	Particle processes are embedded in spacetime	E1, E2, E3, E4
		FDs show observable processes	E1, E2, E3, E4
		Particles are small balls	E1, E2, E3
		Focus on the concept of "building blocks" and neglect of the concept of "interaction"	E1, E2
	Potential sources of inadequate conceptions	Use of scientific language is a source of misconceptions	E1, E2, E3
		The axes of FDs are misleading	E2, E3
Particle physics can only be treated superficially	Limitations by educational setting	The time used for particle physics could be used otherwise	E1, E3
		Necessary prior knowledge is missing in school-level physics	E1, E3
		It is a challenge for teachers to teach modern physics	E1
	The disciplinary handling of FD is not taught in school	Calculations might be too difficult	E1, E2, E3, E4
		Drawing FDs is challenging	E2
		Some concepts are too difficult for school-level physics	E2

3.1.1. FDs Evoke and Perpetuate Inadequate Conceptions about Particle Physics

The experts discussed several inadequate conceptions, as well as possible sources of these conceptions, that are connected to FDs.

The conception that was mentioned most often as an inadequate conception connected to FDs was that they represent processes embedded in space–time. E1 pointed out that "this form of illustration in the sense of spatiotemporal embedding" should be avoided. However, E3 has "also noticed with teachers" that they "try to read off information from [the] geometry about the movement . . . of the particles involved". E2 also clarifies that the interpretation of an FD as a space–time diagram is "only a crutch here" because it cannot "be a real space-time diagram at all", since there is "no concept of path in quantum mechanics". E4 describes it as a "graphic idealization of a process from which one thinks that it actually does not happen that way, but which seems to be the correct technique for bookkeeping".

Close to that idea, the experts said that it is a common misconception that FDs show observable processes. E2 explains that an FD is "not at all an image of a process in that sense" but "a quantum mechanical amplitude" and that "a process is not only described by one FD but by infinitely many". E3 emphasized that "an FD is not a literally readable description of an actually happening process". E1 went on to say that "a physical process, that must be something that I can observe in the world. The rock that falls. And FDs just do not play in this league". He also pointed out that FDs "ostensibly tell stories about what is happening" and expands on this point by emphasizing that quantum field theory, that is the theory underlying elementary particle physics, is "no less strange" than quantum mechanics and that the "form of clarity" suggested by FDs is "not [offered] by our modern theories".

Another inadequate conception closely connected to FDs is the conception of particles as being small balls which physically collide with each other. E2 explained in the context of Compton scattering that even university students still have the conception that "an electron- and a photon-ball meet each other and then there is an elastic collision, and then there is somehow energy- and momentum-transition and then the electron and the photon fly away from each other with a different momentum". E1 compared this to the transport conception of electric current which was "as wrong as only few things" but "still this naïve particle conception could be measured . . . so strongly". He also made a point about interaction particles stating that "the idea that a particle mediates something through an exchange" is "pure use of a metaphor" in his eyes. He emphasized that "nothing is exchanged. It is not here before and there afterwards".

Not directly related to FDs, E1 mentioned the challenge that particle physics is often thought of as "searching the fundamental building blocks" while, in fact, "the part-whole relation of modern physics" would be in fact "superposition and mixing". E2 also mentioned this sub-subtheme but was more practical in the sense that he mentioned that he "is not a friend of first introducing all the particles because that only leads to learning them by heart". They shared the concern of conveying a reductionist image of particle physics as the theory of "the fundamental building blocks".

The experts discussed some sources of misconceptions stemming from the FDs. One important source for the origin of these misconceptions is the use of FDs in the everyday life of scientists and hence the use of scientific jargon which is closely related to their use. E1 described this circumstance as "the use of FDs is always embedded in a practice" and "talking about FDs . . . is simply jargon". For students, however, "these diagrams . . . are simply pure poison" because "there is the danger that school only conveys the jargon but not the practice of course". He stated that "experts talk like topsy-turvy about them" and would use "the most naïve metaphors such as 'comes in', 'comes out', 'get scattered'", but this would all be "embedded into practice which then clears the misconceptions". Besides these general considerations about scientific language, the experts mentioned several specific terms which might be a source of misconceptions, in particular, the term "virtual particle", which E2 calls "very misleading" because "it leads one to say they actually weren't there, as in virtual reality, and that's not the case". A term which was explicitly mentioned as a term used by experts by E2 and E3 but not suitable for students was the term "propagator".

Another source of misconceptions that was mentioned was the drawing of space- and time-axes in FDs. E2 explained that "in the region where the interactions take place, there is no linear order of time anymore . . . but only the initial and final state, where the two lines come in and where the two lines go out, there it can be imagined in good approximation as trajectories". He admitted that he was not sure whether "the axes should be omitted in the first place or only the *t*[time]-axis" and suggested to investigate *x*-*t*[space-time]-axes, whether they help or lead to even more misconceptions". E3, on the other hand, suggested that "one should not draw a space axis".

3.1.2. Feynman Diagrams Can Only Be Treated Superficially with High School Students

The second theme in the dimension of challenges concerned the treatment of FDs in in school physics classes. Two subthemes belong to this theme. The first subtheme is related to the framework conditions that might inhibit the appropriate treatment of FDs in the classroom. The second subtheme concerns the lack of practice when handling FDs.

One challenge that was pointed out by E1 was that the discussion of particle physics in school lessons takes time away from the discussion of other curricular topics, since particle physics is often not marked as a mandatory topic in high school physics. On the question of whether FDs or particle physics should be treated in school lessons, he explained that this was also connected with the question of what would be lost as a result. This time "should not be taken away from quantum mechanics, for example". E3 pointed out that "if the connection to theory and to quantum mechanics cannot be made due to lack of time and so forth, then it [the treatment of FDs] shouldn't be done".

E3 also pointed out that for an appropriate introduction of FDs in school lessons, a certain level of "prior knowledge is necessary in order not to be subject to a too literal reading of the diagrams", that is, "prior knowledge of quantum objects in general". E1 also raised the concern that the topic of particle physics might be taught "by teachers who never learned about that topic in their education", since particle physics is often not part of the mandatory courses for teacher training students.

The lack of practice is most prominent in calculations that are too difficult for school. E1 pointed out that particle physics in schools can only convey "an overview knowledge. Even more than in other fields of school physics, students will not learn any calculations, they won't solve any concrete problems". E2 described one experience with high school students when it was too difficult "to explain in one day" to draw a certain conclusion from a calculation. E4 points out that "one cannot hope that one can gain a complete understanding of these theories without diving deep into the mathematics".

E2 also mentioned the experience that "it seems to be more difficult for [university] students to draw FDs than to interpret presented FDs", but he admitted that "this was for reasons [he] doesn't understand".

Moreover, E2 mentioned two concepts which are explicitly too complicated for students, namely the transformation to the center-of-mass inertial system for arguing that certain decays are kinetically forbidden and the mixing via the Cabibbo–Kobayashi–Maskawa matrix.

3.2. Opportunities

Table 2 shows the three themes and their associated subthemes that were attributed to the dimension of 'opportunities'. In contrast to the first dimension, the subthemes are not further divided into sub-subthemes.

Table 2. Themes and subthemes for the dimension opportunities. The order of the subthemes within a theme is according to the frequency of overall mentions.

Theme	Subtheme	Mentioned by
FDs offer a link between particle physics and high school topics	FDs are suited to teach conservation laws	E1, E2, E3
	FDs link particle physics and quantum mechanics	E1, E2, E4
	FDs offer an insight into the use of structurally equivalent representations	E1, E2, E3
	FDs offer an analogy to resonance phenomena in classical oscillations	E2, E3
FDs offer an opportunity to teach different particle physics topics	Outer and inner lines/virtual particles	E1, E2, E3
	Introduction of interaction particles	E1, E2, E3
	Suggestions for educational uses of FDs	E1, E2, E3
	Particle types	E2, E3, E4
	Introduction of pair production and annihilation	E1
FDs offer a connection to current research	FDs help scientists to discuss particle processes	E1, E2, E3, E4
	Particle physics is a showcase for modern science	E1, E2, E3, E4
	Students can find FDs in popular scientific representations	E1, E3

3.2.1. FDs Offer a Link between Particle Physics and Other High School Physics Topics

The experts mentioned in the interviews that FDs are linked in many ways to concepts that occur in school lessons. Most prominently, the concept of charge conservation was judged to be very connectable. E2 described FDs as “a wonderful means to check for conservation laws” and even estimated that “charge conservation … could definitely be done with younger [than upper secondary] students”. Furthermore, E1 uses conservation laws when he is asked to explain a specific FD.

Almost as often as the first subtheme, the experts mentioned that a link between quantum mechanics and particle physics could be established using FDs. E1 emphasized that in the teaching of elementary particle physics, “a certain unity of modern physics” should be presented and that “one should see where one can deepen what one has already learned in quantum mechanics, that this is regarded as [a] unity”. E2 pointed out that “one should refer to the other world of quantum physics as early as possible”. Moreover, E4 pointed out that “it is important to see that the Standard Model is a quantum theory, that the quantum effects which one has maybe discussed more generally are built into it”. More specifically regarding FDs, E1 stated that he would “elucidate the meaning of an FD analogous to the single amplitudes in the double slit, … that there is one contribution of the wave function, which expresses that the electron goes through one slit”. But this would not be a process “which occurs in nature. But the square of the sum of both of these parts corresponds to the observation”. E2 summarized this by stating that “FDs don’t describe the truth singularly but only the coherent sum describes the truth”.

The experts also described the analogy of FDs offering an insight to structural equivalence in analogy to the teaching of electricity. E1 explained that FDs have structural similarities with drawings of electrical circuits. They both have in common that it is important to see “where something is only symbolically noted and how literally a symbolic depiction can be read … This kind of structural knowledge can be used … not to fall into the trap to think so vividly”. The experts called this notion of invariance concerning different drawings ‘topological equivalence’. E3 explained electrical circuits in this sense as “an example where one does not need to focus on the exact geometry in order to analyze what is depicted”. E2 expanded on this thought by making the analogy that “the conservation law is the junction rule of currents for example. Current just doesn’t get lost. … And these charge currents are also in the FDs”. E1 also made another point connected to the teaching of electricity by stating that particle physics could already be introduced together with electricity “if we want to make the point in particle physics that actually there are no particles, then this could be the conclusion which already informs the teaching of electricity in lower secondary schools”.

Another analogy that was highlighted by the experts was between the concept of virtual particles, which are "off mass-shell", and the concept of forced oscillations in mechanics. E2 explained this concept as: "in classical mechanics it is not special at all to have an oscillating system which can be excited with all frequencies and when it is excited with the eigenfrequency I get a high amplitude and it's just like that with virtual particles when I excite them with their eigenfrequency, then the process is very likely because I have a high amplitude". He also connected this concept to the nuclear fusion in the sun, which is happening so slowly because "there all the W-particles ... are highly virtual". This concept might be "fascinating for high school students". E3 also mentioned this analogy and estimated that "it might work well for teachers", while for students it might be more difficult since "forced oscillations ... do not play a role everywhere or only in selected courses".

3.2.2. FDs Offer an Opportunity for Different Particle Physics Topics to Be Taught

While the previous theme was about potential links between particle physics and other high school physics topics, the experts mentioned several aspects of particle physics that could be worthwhile to teach in high school physics. In all of these aspects, FDs play an important role.

One very prominent topic was the distinction between initial, intermediate, and final states. E1 stressed that "the initial state has a meaning, and the final state has a meaning. And when the diagrams do not differ in the initial and final state, then they aren't distinguishable in the measurement". Moreover, then one would be "interested in all possible intermediate processes which can connect them [the initial and final state] according to the rules". This distinction, in his view, has to be "brought home centrally". Further, E2 said that he would "clearly define particles in the initial state and particles in the final state". E3 further explained the distinction between outer lines and inner lines of FDs as for an inner line the "mass-energy-momentum relation ... does not have to be fulfilled" while it has to be fulfilled for an outer line. He also made the connection that "virtual particles are inner lines in FDs". This concept, on the other hand, connects to the analogy of forced oscillations in mechanics mentioned in the previous theme. Connected to the topic of the inner lines is also the topic of "topological equivalence", which was already mentioned in connection with electrical circuits. E3 stressed this topic by explaining that "the temporal order of vertices is not fixed. One can shift vertices and transfer diagrams in other topologically equivalent diagrams, which are then the same diagrams".

Another promising opportunity mentioned is to introduce and further explain the notion of interaction particles (also known as exchange- or messenger-particles or force carriers). E1 even stated that he believes that "the only reason why one should talk about FDs is that one can introduce exchange particles or because exchange particles obtain their metaphorical meaning from this graphical symbolic language". E3 stated the contrasting view that "one needs the concept of exchange- or messenger-particles for that [FDs]" which, in turn, "only make sense ... in this view of fundamental interactions". E2 expands this concept by stating that "when the concept of exchange- or messenger-particles is introduced, then one has to say that these exchange- or messenger-particles couple to charge, that is, they see—in quotation marks—only other objects which carry or own the charge of this interaction".

Two of the experts also mentioned a subtheme on how FDs are connected to different particle types. E2 stated that "it suffices to first only work with the electron, photon, up and down quark, ... electron neutrino, ... W and Z", that is, he only wants to introduce a selection of particles. E3 suggested that "messenger particles could be drawn differently from matter particles to emphasize the distinction". Another important point concerns the antiparticles, which in FDs are depicted with an arrow against the direction of time. E3 acknowledged the inadequate conception of a particle moving backward in time but also stressed that "in his experience, teachers accept [that it's just a symbolic convention]".

One other aspect raised by E1 was the introduction of pair production and particle annihilation, which would be "the main point of relativistic quantum field theory". He explains that "the type of question on which FDs could be an answer to" is about how a detector can detect many particles although only two particles interact with each other.

The experts also made several suggestions on how to introduce FDs in high school classrooms. For example, both E2 and E3 mentioned *Feynman-Rhombino*, a card game which was invented to practice the rules of creating FDs [31]. These rules were described by E2 as "chaining up [basic] vertices to build scattering processes, decay processes, annihilation- and creation processes". E2 also suggested that "in school, one can explain some basic calculation principles", and he described FDs as "a wonderful tool . . . that I can bring something from the right to the left side, from initial to final state" and by that one can make connections to reaction equations in chemistry or even mathematical equations.

3.2.3. FDs Offer the Opportunity to Discuss Insightful Perspectives on Science

The experts mentioned several aspects where FDs can offer perspectives on science in general and modern physics in particular.

E4 distinguished the roles of FDs "between a pedagogical role to convey the theory and a scientific role to work with the theory as a scientist". The latter was described by E2 from the point of view of a particle physicist. Specifically, for him, FDs are "[a] very clear representation, [a] very helpful, but also very impressively powerful representation". E4 raised the point that "FDs can serve to . . . also [show] the methodology of particle physics". According to him, physics classes in upper secondary schools should convey "the way of thinking, the methodology which lies behind [the theories and physics]" and "FDs [can be] an extremely good illustration [for this]".

Apart from the connection to the daily work, the experts also mentioned that FDs in particular, and particle physics in general, are showcases for modern science and modern physics. E2 emphasized that "physics has come so far that it has managed to predict what may happen, . . . what may not happen and even with what probability this or that will happen". This "predictive power" is something that "should definitely be emphasized at school". E4 also described this theory as "the most successful theory that humankind has ever come up with". E3 suggested that students can learn "something about the procedures of and mathematization in theoretical physics" using FDs. E1 suggested using FDs as an opportunity to "[convey] the big message of modern physics . . . We have theories which work perfectly in a functional sense, but actually we do not know what the objects refer to, at least they are not spatially and temporally embedded". He also raised another point on how particle physics might be useful for students, namely by highlighting "the science-sociological and nature of science aspect that science is a collaborative endeavor".

Another aspect related to the engagement with science is the critical examination of scientific representations, such as in textbooks or popular science texts. E3 explains that "of course, it should also be a learning goal" that "the students can classify popular science literature or sources . . . " E1 even notes that "the main benefit in the educational context" is that "one can point out widespread misunderstandings" and "send the students on a journey to find misleading or false representations in popular, but also subject-related representations". This could be "a contribution to an experience of autonomy".

4. Discussion

The results from the expert interviews show the potential, but also the difficulties, associated with the introduction of FDs in high school classrooms. The experts had different opinions on whether and how FDs should be used with high school students. Indeed, all experts presented both opportunities and challenges on this matter. Based on the results, four learning objectives for potential learning material with FDs were developed. These are presented and motivated here.

4.1. Learning Objectives

4.1.1. Charge Conservation

The first learning objective is: “Students will be able to use the concept of conservation of charge to determine whether a particle process is possible or not”. The subtheme “FDs are a tool to teach about conservation laws” was the most often mentioned opportunity, and all three original experts mentioned it (the fourth expert did not mention this opportunity since the scope of that interview was to clarify discussion points connected to more advanced topics such as the connection to quantum physics). Since conservation laws are arguably one of the most important concepts in physics [45], this learning objective is well suited to also connect FDs to different physics topics such as mechanics or electricity. This connection has also been made by Hoekzema et al. [12], who used a modified version of FDs to teach about symmetries and conservation laws. Additionally, Pascolini and Pietroni [25] used FDs explicitly without their original meaning in the mathematical sense to teach concepts such as conservation laws. Furthermore, there have been different approaches to certain educational activities that help students to use conservation laws with FDs in a playful way [30,31].

We argue that conservation laws are the inevitable entry point to any learning unit dealing with FDs. If the deep connection between conservation laws and FDs is not understood, FDs cannot be used. This first learning objective therefore serves as a prerequisite to all of the following.

4.1.2. Interaction Particles

The second learning objective is: “Students will be able to explain the role of interaction particles and motivate their existence in the Standard Model of particle physics”. The experts mentioned interaction particles as one of the most prominent subthemes of the opportunities that FDs can provide for the teaching of particle physics concepts. Lindenau and Kobel [15] argue that the concept of interaction particles is one of the core concepts in the Standard Model of particle physics. A common way to introduce interaction particles is through everyday analogies, such as people on a boat throwing balls and boomerangs at each other, but these tend to promote inadequate conceptions rather than explain physical concepts. We argue that FDs provide a motivation for the concept of interaction particles without going too much into the mathematical details of the Standard Model of particle physics on the one hand and avoiding simple heuristics on the other hand. A more sophisticated way is to connect interactions with fields using the concept of interaction particles [7,28,46]. While this is arguably an adequate way to think about interaction particles—or particles in general—it is also more abstract and thus more difficult to comprehend. However, the concept of interaction particles could be an opportunity to link this idea of how interactions are thought about in particle physics to the concept of the electromagnetic field, which might already be known from earlier school physics.

4.1.3. Superposition

The third learning objective should directly link particle physics to already known quantum physics concepts. It consists of two parts to make it more accessible and relevant to high school education. The first part is: “Students will be able to apply the superposition principle from quantum mechanics to particle processes”. Here, students should realize that a process is not only represented by a single diagram, but that several diagrams always contribute to a process. This is in analogy to the double and multiple slit experiment, where several slits contribute equally to the interference pattern. Extending this principle, the second part of this learning objective is: “Students will recognize particle processes as a superposition of infinitely many contributions”. Here, students learn that for any FD there can be another FD drawn that represents the same particle process.

Also in the literature, Passon et al. [29] mention the above mentioned analogy to the double slit experiment. Moreover, Allday [7] showcased how this principle is important for learning about the nature of interactions in particle physics. Passon et al. [29] made the connection to the concept of topological equivalence and pointed out structural similarities between FDs and electrical circuits.

This concept is closely connected to a common misunderstanding about FDs, that of representation. As stated above, the debate on whether FDs represent physical processes is almost as old as FDs themselves [21,22]. For certain processes, some diagrams are much more dominant in the calculation; therefore, in these cases, one diagram is a good approximation of what is happening. However, to distinguish between such processes where a single diagram might be a good approximation and processes where several diagrams are needed for a reasonable approximation, a good working knowledge of the corresponding mechanisms, the parameters, and the energies and momenta is needed. Therefore, for high school students, we argue that learning about the general principles might already be challenging enough. For advanced students who already know about the series expansion of a mathematical function, this learning objective could also be linked with an introduction to calculations, but in principle, the learning objective is achievable without addressing this mathematical concept.

4.1.4. Work of Particle Physicists

The experts emphasized that FDs can give an insight into how work is conducted in (theoretical) particle physics. To take advantage of this, the fourth learning objective is: “Students will be able to relate the method of FDs to the work of particle physicists”. This learning objective accounts for the fact that a whole theme of opportunities was dedicated to the perspectives on science, and within that, a prominent subtheme was the helpfulness of FDs to scientists. Nevertheless, the helpfulness for scientists alone would not be an argument to use FDs in schools. Some prior knowledge is necessary to use the diagrams in a way scientists use them. We argue that, after achieving the first three learning objectives, students might have acquired enough knowledge to get a glimpse into why and how scientists use FDs. For instance, to achieve this learning objective, one can imagine presenting students with an example of a measurement that does not agree with a theoretical prediction. In this case, the students would have to realize that this could be due, among other things, to the fact that not enough FDs were considered. In this way, the use of FDs can be illustrated in practice.

However, even if all the learning goals were achieved, a full mathematical account would be too ambitious for high school students, as the underlying mathematics requires knowledge of quantum field theory. Nonetheless, as illustrated above, some of the basic underlying calculation principles might be suitable for advanced high school students.

Furthermore, even without going into the details of calculations, a simulation tool could help students understand why the calculation of different diagrams shapes the theoretical predictions for particle physics experiments.

This learning objective is particularly beneficial for learners from a nature of science perspective in the sense that students get to experience first-hand that science is theory-laden by exploring theoretical predictions and getting a glimpse of the tentativeness of science when open questions of modern physics can be examined by FDs (e.g., the high-precision measurement of the magnetic moment anomaly of the muon [47]).

4.2. Challenges to Address When Teaching with FDs

While all these learning objectives pose good opportunities for the learning of and with FDs, the inherent challenges that are connected also need to be addressed. These challenges might translate into domain-specific design principles when designing a learning environment that incorporates these learning objectives.

The most prominent theme mentioned within the challenges dimension was that FDs might evoke and perpetuate inadequate conceptions, most notably one of a space–time embedding of particle processes. One discussed source for this inadequate conception was that of the drawing of an FD as a space–time diagram. Among the experts, there was a tendency towards the opinion that while a time axis is useful to denote what is the initial and what is the final state, a space axis does not make any physical sense. Therefore, we argue that FDs should be clearly distinguished from space–time diagrams, or, at least, drawing a space axis should be avoided. Closely connected to this is the conception that particles are small balls that are physically colliding at the interaction points. Both conceptions might also be perpetuated by the depicting particles using lines with arrows. We therefore suggest exploring the possibility to omit the arrows when using FDs in educational contexts. The only function of the arrows in an FD is to distinguish particles from antiparticles. This distinction can be made differently, for example, by using colors or by writing the corresponding symbol. Conceptualizing antiparticles as particles that move backward in time might be a mathematically allowed description when also attributing a negative energy to antiparticles, but it does not represent the physical reality and might evoke misconceptions about the nature of antiparticles.

Another very prominent inadequate conception was that FDs show observable processes. Avoiding this conception is very challenging since part of the popularity of the diagrams is that they seem to "tell a story of what happens" (E1). We argue that pronouncing the third learning objective (particle processes are superpositions of FDs) addresses this inadequate conception.

The most prominent single subtheme in the expert interviews was the challenge that the use of scientific language is a source of inadequate conceptions. This issue might be addressed by a careful use of language in educational contexts [17,48], for example, when speaking of "decay" by clearly explaining that it is in fact a transformation of particles and speaking of "electrical charge" instead of just "charge" to emphasize the fact that there are different charges. However, more research is needed into which conceptions are connected to these different terms and in which way they are beneficial or hindering for learning.

Other challenges touch on the topic of too little time or missing prior knowledge for the meaningful use of FDs in high school physics classes. We propose to address this challenge by linking FDs to concepts that are already known, such as charges, conservation laws, and, for more advanced students, the superposition principle in quantum mechanics or even series expansion in mathematics. If the latter is not yet known by the students, calculations can be omitted in educational settings or substituted by simulation tools instead.

4.3. Outcomes and Limitations

The scope of the study was to explore possible challenges and opportunities of using FDs in high school education. In doing so, this study was the first one to explore the educational use of FDs empirically. Our goal was to interview experts whose opinions are representative of the field. For example, in the interviews we could find the above-mentioned "Feynman-Dyson split" [18,21] to some extent. While E1 stressed the purely mathematical nature of FDs and strongly opposed the notion of a too literal reading of the diagrams, E2 was very open to using the diagrams as a pictorial representation also in a qualitative sense, that is, without immediately referring to the mathematical expression that certain FDs represent. Nevertheless, we have only covered a narrow perspective on the topic, since all four experts came from the German-speaking region (the first three from Germany, the fourth from Switzerland). This might have led to a biased view. However, the small number of experts allowed an in-depth analysis in line with the explorative purpose of the current study. This study in its present form already gives valuable insight for the development of educational material. Moreover, it serves as a basis for future studies with a larger number of experts that can be envisaged to build upon our initial results by conducting a Delphi study [49] and gathering additional perspectives on the educational value of FDs.

4.4. Outlook

The learning objectives and design suggestions developed in this study are used to create learning material for introducing FDs to 14–18-year-olds within a massive open online course (MOOC) on particle physics. The development of this learning material follows the iterative design principles of the design-based research framework [32]. Specifically, design principles based on multimedia learning theories are developed for these learning materials and are combined with more domain-specific design principles resulting from this study. The developed learning material is tested in teaching experiments using eye-tracking data with high school students from the target group. The learning materials will be further developed based on the results of these teaching experiments.

5. Conclusions

The presented study systematically analyzed the risk and potential that lies in teaching about particle physics to high school students using FDs. We have found two general challenges and three major opportunities that are connected to this peculiar form of representation within particle physics. We argue that the opportunities we have found are calls for using FDs in teaching particle physics, even beyond their original function as mere calculation tools. Indeed, FDs can serve as a tool to connect particle physics to already known concepts, such as charges and conservation laws on the one hand and to other advanced topics such as quantum mechanics on the other hand. Since FDs are used by particle physicists every day, their educational use can offer high school students a window into the daily work of particle physicists. The challenges we have found can be addressed by domain-specific design principles for developing a learning environment for and with FDs.

We are confident that it is useful to introduce high school students to FDs, not least because FDs are part of popular representations of particle physics. As one of the experts who took part in this study stated, one can "send students on a journey to find misleading or wrong depictions in popular or even specialized media ... that can be a contribution to an experience of autonomy". This could be a major strength of using this form of representation in physics classrooms.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/physics4040085/s1, Table S1: Coding Manual; Document S2: Interview guideline E1, E2, E3; Document S3: Interview guideline E4.

Author Contributions: Conceptualization, M.N.D., J.W., P.K. and A.M.; methodology, M.N.D. and J.W.; validation, J.W.; investigation, M.N.D.; writing—original draft preparation, M.N.D.; writing—review and editing, M.N.D., J.W., P.K. and A.M.; supervision, P.K., A.M., S.M.S. and J.W.; project administration, S.M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work has been sponsored by the Wolfgang Gentner Programme of the German Federal Ministry of Education and Research (grant no. 13E18CHA).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data presented in this study, namely the interview segments and, if justified, the full interview transcripts, are available on request from the corresponding author. The data are not publicly available due to the personality rights of the experts.

Acknowledgments: Thank you to the four experts who contributed to this study. Many thanks to Larissa Hahn and Simon Lahme (both University of Göttingen) for a thorough discussion of the coding scheme and for serving as interraters. Moreover, thanks to Marlene Breither (University of Göttingen) for serving as interrater.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Swinbank, E. Particle Physics: A new course for schools and colleges. *Phys. Educ.* **1992**, *27*, 87–91. [CrossRef]
2. Aubrecht, G.J. Report on the conference on the teaching of modern physics. *Phys. Teach.* **1986**, *24*, 540–547. [CrossRef]
3. Sjøberg, S.; Schreiner, C. Results and perspectives from the Rose project. In *Science Education Research and Practice in Europe: Retrospective and Prospective*; Jorde, D., Dillon, J., Eds.; Sense Publishers: Rotterdam, The Netherlands, 2012; pp. 203–236. [CrossRef]

4. Barlow, R. Particle Physics: From School to University. *Phys. Educ.* **1992**, *27*, 92–95. [CrossRef]
5. Kobel, M. High school students' exposure to modern particle physics. *Europhys. News* **2003**, *34*, 108–110. [CrossRef]
6. Tuzón, P.; Solbes, J. Particle physics in high school: A diagnose study. *PLoS ONE* **2016**, *11*, e0156526. [CrossRef]
7. Allday, J. The nature of force in particle physics. *Phys. Educ.* **1997**, *32*, 327. [CrossRef]
8. Zügge, T.; Passon, O. Das Wuppertaler Curriculum der Elementarteilchenphysik. In *Kohärenz im Unterricht der Elementarteilchenphysik*; Passon, O., Zügge, T., Grebe-Ellis, J., Eds.; Springer Spektrum: Berlin/Heidelberg, Germany, 2020; pp. 121–141. [CrossRef]
9. Kobel, M.; Bilow, U.; Lindenau, P.; Schorn, B. *Teilchenphysik: Ladungen, Wechselwirkungen Und Teilchen, 3. Auflage.*; Joachim Herz Stiftung: Hamburg, Germany, 2018. [CrossRef]
10. Polen, C. Particle physics: An essential and engaging part of the high school physics program. *Phys. Teach.* **2019**, *57*, 320–322. [CrossRef]
11. Van de Berg, E.; Hoekzema, D. Teaching conservation laws, symmetries and elementary particles with fast feedback. *Phys. Educ.* **2006**, *41*, 47–56. [CrossRef]
12. Hoekzema, D.; Schooten, G.; van den Berg, E.; Lijnse, P. Conservation laws, symmetries, and elementary particles. *Phys. Teach.* **2005**, *43*, 266–271. [CrossRef]
13. Michelini, M. Innovation of curriculum and frontiers of fundamental physics in secondary school: Research-based proposals. In *Fundamental Physics and Physics Education Research*; Sidharth, B.G., Murillo, J.C., Michelini, M., Perea, C., Eds.; Springer International Publishing: Cham, Switzerland, 2021; pp. 101–116. [CrossRef]
14. Passon, O. Mystifizierung der Quantenmechanik und Trivialisierung der Teilchenphysik. In *Kohärenz im Unterricht der Elementarteilchenphysik: Tagungsband des Symposiums zur Didaktik der Teilchenphysik, Wuppertal 2018*; Passon, O., Zügge, T., Grebe-Ellis, J., Eds.; Springer Spectrum: Berlin/Heidelberg, Germany, 2020; pp. 79–90. [CrossRef]
15. Lindenau, P.; Kobel, M. Introducing an innovative approach of teaching the Standard Model of particle physics at high school. *J. Phys. Conf. Ser.* **2019**, *1287*, 012045. [CrossRef]
16. Gourlay, H. Learning about A level physics students' understandings of particle physics using concept mapping. *Phys. Educ.* **2016**, *52*, 014001. [CrossRef]
17. Wiener, G.J.; Schmeling, S.M.; Hopf, M. Can grade-6 students understand quarks? Probing Acceptance of the subatomic structure of matter with 12-year-olds. *Eur. J. Sci. Math. Educ.* **2015**, *3*, 313–322. [CrossRef]
18. Kaiser, D. Drawing Theories Apart: The Dispersion of Feynman Diagrams in Postwar Physics; University of Chicago Press. 2005. Available online: https://academic.oup.com/chicago-scholarship-online/book/23068 (accessed on 10 October 2022).
19. Dyson, F.J. The radiation theories of Tomonaga, Schwinger, and Feynman. *Phys. Rev.* **1949**, *75*, 486–502. [CrossRef]
20. Feynman, R.P. Space-time approach to quantum electrodynamics. *Phys. Rev.* **1949**, *76*, 769–789. [CrossRef]
21. Passon, O. On the interpretation of Feynman diagrams, or, did the LHC experiments observe $H \to \Gamma\gamma$? *Eur. J. Philos. Sci.* **2019**, *9*, 20. [CrossRef]
22. Meynell, L. Why Feynman diagrams represent. *Int. Stud. Philos. Sci.* **2008**, *22*, 39–59. [CrossRef]
23. Dorato, M.; Rossanese, E. The nature of representation in Feynman diagrams. *Perspect. Sci.* **2018**, *26*, 443–458. [CrossRef]
24. Passon, O.; Zügge, T.; Grebe-Ellis, J. Pitfalls in the teaching of elementary particle physics. *Phys. Educ.* **2018**, *54*, 015014. [CrossRef]
25. Pascolini, A.; Pietroni, M. Feynman diagrams as metaphors: Borrowing the particle physicist's imagery for science communication purposes. *Phys. Educ.* **2002**, *37*, 324. [CrossRef]
26. Jones, G.T. The Uncertainty Principle, virtual particles and real forces. *Phys. Educ.* **2002**, *37*, 223–233. [CrossRef]
27. Woithe, J.; Wiener, G.J.; der Veken, F.F.V. Let's have a coffee with the Standard Model of particle physics! *Phys. Educ.* **2017**, *52*, 034001. [CrossRef]
28. Lambourne, R. Predicting the physics of particles. *Phys. Educ.* **1992**, *27*, 71. [CrossRef]
29. Passon, O.; Lindenau, P.; Kobel, M. Von Feynman-Diagrammen und Stromkreisen. *Unterr. Phys.* **2020**, *180*. Available online: https://www.friedrich-verlag.de/physik/quantenphysik/von-feynman-diagrammen-und-stromkreisen-7480 (accessed on 10 October 2022).
30. Harlander, R.V.; Klein, S.Y.; Lipp, M. FeynGame. *Comput. Phys. Commun.* **2020**, *256*, 107465. [CrossRef]
31. Lindenau, P.; Winkler, O. Feynman-Rhombino. *Unterr. Phys.* **2020**, *180*. Available online: https://www.friedrich-verlag.de/physik/quantenphysik/feynman-rhombino-7486 (accessed on 10 October 2022).
32. Baumgartner, E. et al. [The Design-Based Research Collective] Design-based research: An emerging paradigm for educational inquiry. *Educ. Res.* **2003**, *32*, 5–8. [CrossRef]
33. Duit, R.; Komorek, M. Understanding the basic ideas of chaos-theory in a study of limited predictability. *Int. J. Sci. Educ.* **1997**, *19*, 247–264. [CrossRef]
34. Kersting, M.; Henriksen, E.K.; Bøe, M.V.; Angell, C. General relativity in upper secondary school: Design and evaluation of an online learning environment using the model of educational reconstruction. *Phys. Rev. Phys. Educ. Res.* **2018**, *14*, 010130. [CrossRef]
35. Kamphorst, F.; Vollebregt, M.J.; Savelsbergh, E.R.; van Joolingen, W.R. An educational reconstruction of special relativity theory for secondary education. *Sci. Educ.* **2021**. [CrossRef]
36. Anderson, T.; Shattuck, J. Design-based research: A decade of progress in education research? *Educ. Res.* **2012**, *41*, 16–25. [CrossRef]

37. Haagen-Schützenhöfer, C.; Hopf, M. Design-based research as a model for systematic curriculum development: The example of a curriculum for introductory optics. *Phys. Rev. Phys. Educ. Res.* **2020**, *16*, 020152. [CrossRef]
38. Barab, S.; Squire, K. Design-based research: Putting a stake in the ground. *J. Learn. Sci.* **2004**, *31*, 1–14. [CrossRef]
39. Duit, R.; Gropengießer, H.; Kattmann, U.; Komorek, M.; Parchmann, I. The model of educational reconstruction—A framework for improving teaching and learning science. In *Science Education Research and Practice in Europe: Retrospective and Prospective*; Jorde, D., Dillon, J., Eds.; SensePublishers: Rotterdam, The Netherlands, 2012; pp. 13–37. [CrossRef]
40. Bogner, A.; Littig, B.; Menz, W. *Interviews mit Experten: Eine praxisorientierte Einführung*; Springer: Wiesbaden, Germany, 2014. [CrossRef]
41. Wiener, G.J.; Schmeling, S.M.; Hopf, M. The technique of probing acceptance as a tool for teachers' professional development: A PCK study. *J. Res. Sci. Teach.* **2018**, *55*, 849–875. [CrossRef]
42. Burde, J.-P.; Wilhelm, T. The electron gas model as an introduction to electricity in middle school science. In *Electronic Proceedings of the ESERA 2015 Conference: Science education research: Engaging Learners for a Sustainable Future*; Lavonen, J., Juuti, K., Lampiselkä, J., Uitto, A., Hahl, K., Eds.; University of Helsinki: Helsinki, Finland, 2016; pp. 26–36. Available online: https://www.dropbox.com/s/wrb22pgeociw8dr/eBook2015_Part_1_links.pdf (accessed on 10 October 2022).
43. Braun, V.; Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [CrossRef]
44. Hayes, A.F.; Krippendorff, K. Answering the call for a standard reliability measure for coding data. *Commun. Meth. Meas.* **2007**, *1*, 77–89. [CrossRef]
45. Hill, C.T.; Lederman, L.M. Teaching symmetry in the introductory physics curriculum. *Phys. Teach.* **2000**, *38*, 348–353. [CrossRef]
46. Daniel, M. Particles, Feynman diagrams and all that. *Phys. Educ.* **2006**, *41*, 119–129. [CrossRef]
47. Wolchover, N. 'Last Hope' Experiment Finds Evidence for Unknown Particles. Quanta Magazine. 7 April 2021. Available online: https://www.quantamagazine.org/last-hope-experiment-finds-evidence-for-unknown-particles-20210407/ (accessed on 10 October 2022).
48. Wiener, G.J.; Schmeling, S.M.; Hopf, M. Introducing 12 year-olds to elementary particles. *Phys. Educ.* **2017**, *52*, 044001. [CrossRef]
49. Osborne, J.; Collins, S.; Ratcliffe, M.; Millar, R.; Duschl, R. What "ideas-about-science" should be taught in school science? A Delphi study of the expert community. *J. Res. Sci. Teach.* **2003**, *40*, 692–720. [CrossRef]

6.2 Learning Goals

As described in the previous chapter's discussion, four learning goals could be derived from the experts' statements. This section describes and motivates these learning goals. The first two have been tested in different ways, as described below. It was beyond the scope of the current PhD project to also implement and test the third and fourth learning goals, but I am motivated by why these are valuable goals and how they would be implemented. For each learning goal, key messages have been formulated, which are presented and explained in this section.

6.2.1 Learning Goal 1: Charge Conservation

The first learning goal is about charges and conservation laws and is stated as

> The students are able to check whether a particle process is possible using Feynman diagrams and the law of charge conservation.

Feynman diagrams are drawn according to specific rules. In scientific practice, these rules are formulated in a very mathematical language, assigning each diagram element to a mathematical expression (cf. Section 2.1.2). The rules that govern nature, i.e., the conservation laws, are inscribed in these mathematical rules. These mathematical rules are relatively inaccessible for students, but the underlying principles aren't. Conservation laws play an essential role in all fields of physics, and upper secondary students who are the target group of the learning material developed within this project probably already had some physics lessons about classical mechanics in which they have heard about energy and momentum conservation. Hence, the concept of conservation laws should already be known to students. The concept is now applied to a lesser-known concept of charges. Students probably know about electric charges and that opposite charges repel each other, and the same charges attract each other. However, electric charges are usually not seen as a property of particles but rather as entities that "flow within a wire," for example. In particle physics, several different charges must exist that are closely connected to the different interactions. The electric charge, for example, is closely connected to the electromagnetic interaction, whereas the weak interaction has an analogue called the isospin. Throughout this work, I follow Kobel et al. (2022) who suggest introducing the third component of the weak isospin, which can take the values of $+1/2$ or $-1/2$. In the following and within the learning material, this quantity is referred to as *weak charge*. These charges have to be conserved quantities for the interactions to exist. Therefore, any particle process is only possible if the charges of the involved particles are conserved. This can be checked even before it's known what precisely a particle process is, while at the same time, it is crucial to understand further how particle processes work.

Following from these considerations, the key messages on the path to this learning goal are:

- Particles are described by several properties, among them the so-called charges.
- Different charges exist: The "electric" and the "weak" charges.
- All charges must be conserved separately in any particle process (meaning, the total charge must be equal at the beginning and the end).

6.2.2 Learning Goal 2: Interaction Particles

The second learning goal goes one step further into the understanding of particle processes and is stated as

Students can explain the role and motivate the existence of interaction particles in the standard model by using Feynman diagrams.

In the particle picture, a particle interaction can be described by the exchange of an interaction particle. This description goes back to Yukawa (1935) in which he describes the interaction between neutrons and protons as mediated by so-called mesons. This was well before Feynman diagrams were introduced, in which interaction particles appear as wavy or curly lines to represent interaction terms in perturbative calculations to calculate probabilities of particle processes. Even though the precise mathematical meaning of these terms is not accessible to students, it is relevant for understanding the nature of these particle processes in which different interactions exist, which differ in the interaction particles which mediate this interaction. By introducing several simple particle processes governed by different interactions and their lowest-order Feynman diagrams, it can be understood why there are different types of interactions. These examples are restricted to the electromagnetic and the weak interactions because the mechanism of the strong (colour) charge is more complicated than that of the weak and the electric charge – at least on the level which is accessible to upper secondary students – and would add another layer of complexity which is not needed within the goals of this learning materials. For this learning goal, the concept of different charges from the previous learning goal is connected to the concept that these charges stand for different interactions. The role of the interaction particles is then to determine which interaction governs the respective process. These considerations have led to the following key messages which are incorporated into the learning material:

- Having a charge means a particle participates in a certain interaction.
- Different interactions mediate different processes.
- In the diagram, we draw the interaction particles to show that the process occurs via this interaction.

6.2.3 Learning Goal 3: Superposition of Diagrams

The third learning goal concerns the mathematical nature of particle physics. Since it is a complicated learning goal, it consists of two parts. The first goal is stated as

Students can apply the superposition principle from quantum mechanics to depicting a particle process.

Ideally, the students working with this learning material have already heard about quantum mechanics and the double-slit experiment. This experiment explains a fundamental principle of quantum mechanics, the superposition of different events. If the light is assumed as made out of particles, the interference pattern behind the double slit can only be explained if a photon does not "decide" through which slit it goes, but it takes both ways and "interferes with itself". This is called the superposition principle of quantum mechanics. This can be applied to Feynman diagrams when it comes to explaining the mechanism of how they are used in calculations. A certain process can be drawn in multiple different ways. While it does not matter what the angles and lengths of the lines are, diagrams representing the same process can differ in the type and the number of vertices, i.e., the points where the lines meet. In the experiment, which diagram contributed to a particular process is indistinguishable. Contrary, when calculating the probability of a given process, all contributing diagrams need to be considered. The final result is the square of the sum of all the contributions; therefore, there are interference terms. A very simple analogy which is accessible to students is the binomial formula $(a + b)^2 = a^2 + b^2 + \mathbf{2ab}$ with the latter term introduced as "interference term" between the quantities a and b.[1] With these considerations, the key messages for this part of the learning goal are

- The angles and lengths of lines don't matter for the meaning of an FD.
- A combination of initial and final state can happen differently. All these ways happen simultaneously and must be considered in the computation.

The second part of the learning goal takes the concept a step further and is stated as

Students can understand any particle process as a superposition of infinitely many contributions.

For this learning goal, the idea of a vertex factor has to be introduced which assigns a number dependent on the type of the interaction to each vertex of the diagram. The strength of the interaction determines this vertex factor. To compute the order of magnitude of the contribution of a given diagram, all its vertex factors must be multiplied. To compute the probability of a given process, all the contributions from the diagrams representing this process must be added. This process is complicated by the fact that for any given diagram, another diagram can be drawn which represents the same process, i.e., having the same initial and final state, but which has

[1] In reality, the probability amplitude $\mathcal{A}$ is a complex number, and the quantity to compute is the square of the absolute value. However, this intricacy is omitted to avoid additional confusion.

more vertices, e.g. by adding a photon which is emitted after the initial state and absorbed again before the final state. This addition does not alter the initial and final state, but it changes the contribution of the diagram by adding two vertices. This process could be continued ad infinitum; hence, an exact calculation of a given process's probability would involve computing and adding up infinitely many diagrams. Therefore, the Feynman diagram method can only work if the vertex factors are lower than one. In this case, each diagram consisting of more vertices adds an ever more negligible contribution. If students have advanced maths classes, they know this concept from calculating infinite series, which only converge under certain conditions (a necessary but insufficient condition for the sequence of elements constituting the series to converge towards 0). Particularly interesting for physics is Taylor's series expansion, often used to estimate an arbitrary complicated function. In this estimation, only the first few elements often have to be calculated to get a reasonable good estimation of the desired value. In this manner, a connection can be drawn from maths concepts in school to the standard practices of physicists. These considerations led to the following key messages:

- For a given diagram, we can always draw more diagrams contributing to the same process.
- The more vertices a diagram has, the smaller its contribution to the computation.
- The more diagrams we draw, the more precise the computation is, but the computation gets even more complex.

6.2.4 Learning Goal 4: Particle Discoveries

With the last point of the previous learning goal in mind, the final learning goal draws a connection from the purely abstract concepts to the actual work of physicists:

Students can transfer the approach of Feynman diagrams to modern particle physics.

This can be achieved by drawing the connection between specific Feynman diagrams and how they were used in particle discoveries. An example which might be particularly well suited is the story of the discovery of the Z boson. This is one of the interaction particles of the weak interactions. Its existence was predicted in the mid-1960s, while the first evidence could be found in 1973 in bubble chamber pictures of the bubble chamber *Gargamelle* (Hasert et al., 1973). However, this evidence was only indirect. The first discovery of a Z boson was not made until 1983 in the proton-antiproton collider $Sp\overline{p}S$ (Rubbia, 1985). Later, in the electron-positron collider *LEP*, the so-called Z-resonance was further investigated, so the properties of the Z boson and through that of the weak interaction could be further investigated (Grünewald, 2005). This story can be told along different diagrams representing the processes investigated in the various steps of this story. It is further discussed in Section 8.2 how Feynman diagrams can be used to interpret other disciplinary-specific representations like bubble chamber pictures, event displays, or energy-cross section diagrams.

These considerations are condensed into the following key messages:

- Feynman diagrams help to make theoretical predictions for particle processes.
- The particle processes have to be investigated with particle detectors.
- There are many different ways in which particle processes show up in particle detectors.

This learning goal is closely related to the Nature of Science (NoS). In the following, I give a short, incomplete list of NoS aspects (cf. Section 2.2.3) which can be found in this story. The method of Feynman diagrams is an excellent example of the theory-laden NoS since it showcases how discoveries in particle physics can only be made when scientists understand what they can expect, both qualitatively and quantitatively. This also shows the difference between the empirical and the inferential NoS. While it is crucial to make empirical observations to test theories, results can only be inferred from these observations. A bubble chamber picture like that shown in Figure 8.1 on page 159, for example, cannot be understood without a) knowing how a bubble chamber works and b) knowing that there are "invisible" particles called neutrinos. Also, the story showcases the tentative NoS: Even though the first evidence of the Z boson was found in 1973, it took ten years to confirm its existence, and even then, many of its properties were still to be investigated, which is why also the experiments at the next particle accelerator invested much time and effort into its investigation. To this day, there are open questions regarding the precise value of its mass, which could be a possible path into a completely new field of physics (ATLAS collaboration, 2023; CDF Collaboration et al., 2022). Furthermore, scientists had to be very creative to find ways to measure the particle processes. The observations made with *Gargamelle* were only possible due to the invention of a magnetic horn (Fenkart, 2023). The same physicist who invented the magnetic horn, Simon van der Meer, also invented the so-called stochastic cooling, a crucial invention to make the proton-antiproton collider possible. This idea was pushed by Carlo Rubbia, both receiving the Nobel prize for their contributions to the discovery of the W and Z bosons (Nobel Prize Outreach AB, 2024b).

6.3 Pilot Study: Eye Tracking and Feynman diagrams

The eye-tracking method was tested with Feynman diagrams in the first pilot study. The aim was to check whether some elements within Feynman diagrams draw specific attention from experts and novices. Because the study was conducted during a time when, due to the COVID-19 pandemic, there were still significant contact restrictions, it was conducted with members of my working groups at CERN and Göttingen.

The data were only evaluated qualitatively using heat maps (cf. Section 5.2.4). The results showed that experts' and novices' gaze patterns differed in the amount of attention dedicated to certain areas of the diagrams. For example, when asked for the number of vertices of a diagram, experts directly allocated their attention to the vertices. The effect of spread-out attention will be further investigated in the third student study (Section 6.6.7). At the same time, novices also inspected other areas of the diagram, like the arrows. Experts also interpreted the tasks further and tried to "mentally rotate" vertices, even though it was not asked in the task.

The pilot study showed that Feynman diagrams can be complicated objects for novices. It, therefore, guided the following design process by making the diagrams more accessible in the way presented in Section 5.1.2. It also gave insights into the design of possible tasks, as the tasks used in this study were too complex, and even experts misinterpreted them. The focus of the following studies was then on charge conservation.

The study and its results were published in a conference proceeding to the DPG spring meeting 2021 in German (Dahlkemper et al., 2021). The material used for this study was not designed according to design principles or learning goals but instead served as a testing ground on which kinds of patterns could be expected when different groups look at Feynman diagrams.

Untersuchung der Wahrnehmung von Feynman-Diagrammen mittels Eye Tracking

Merten Dahlkemper*+, Jeff Wiener*, Andreas Müller°, Sascha Schmeling*, Pascal Klein+

* CERN, Esplanade de Particules 1, CH-1211 Genève 23, Schweiz,
+ Fakultät für Physik, Abteilung Didaktik der Physik, Universität Göttingen, Friedrich-Hund-Platz 1, 37077 Göttingen,
° Faculty of Science/Physics Section IUFE, Universität Genf, Pavillon d'Uni Mail, Boulevard du Pont d'Arve 40, CH-1211 Genève, Schweiz

merten.dahlkemper@cern.ch

Kurzfassung

Seit mehr als 70 Jahren sind Feynman-Diagramme (FD) aus der Elementarteilchenphysik kaum mehr wegzudenken, da sie komplexe Rechnungen auf eine kompakte Weise veranschaulichen. Dies wird vielfach auch für die Behandlung im Schulunterricht genutzt.Aus fachdidaktischer und lerntheoretischer Sicht wird der Nutzen verschiedener Repräsentationsformen zum Problemlösen und Lernen als zentral erachtet. Dennoch wird im Rahmen der Vermittlung von Teilchenphysik kontrovers darüber diskutiert, ob und in welcher Form FD im Unterricht der Schule vorkommen sollen, da umstritten ist, inwiefern der Nutzen dieser Darstellungen ihre potentiellen Nachteile durch resultierende Missverständnisse und Fehlvorstellungen übersteigt. Diese Schwierigkeiten und das Fehlen empirischer Untersuchungen zum visuellen Umgang mit FD weisen auf den Forschungsbedarf zu diesem Thema hin. Das Lernen mit graphischen Repräsentationen wie Feynman-Diagrammen beinhaltet visuelle Prozesse. Um diese Prozesse genauer zu untersuchen, verwenden wir Eye Tracking als eine inzwischen auch in den Fachdidaktiken zunehmend verbreitete Methode zur Messung der visuellen Aufmerksamkeit, die uns Einblick in die kognitiven Prozesse geben kann. Die Stichprobe unserer Eye Tracking-Studie setzt sich zum einen aus Studierenden zusammen, die mit der Repräsentationsform nicht vertraut sind, und zum anderen aus Forschenden in dem Gebiet der Elementarteilchenphysik. Das mittelfristige Ziel der Arbeit ist, aus den Studienergebnissen forschungsbasierte Instruktionen zum Betrachten, Zeichnen und Anwenden von FD im Rahmen eines Onlinekurses zur Teilchenphysik zu entwerfen.

1. Motivation

Die Teilchenphysik ist eines der Forschungsfelder, welches sich in der Öffentlichkeit seit vielen Jahren und zunehmend auch im Schulunterricht großer Beliebtheit erfreut. Dies erfordert, dass angemessene Lehr- und Lernmaterialien (weiter)entwickelt werden, mit denen das Thema Teilchenphysik Schüler:innen nahe gebracht werden kann.

Eine Darstellung, die im Zusammenhang mit didaktischen und populärwissenschaftlichen Beiträgen immer wieder gezeigt wird, sind die Feynman-Diagramme (FD; Kaiser, 2005). Nicht zuletzt sorgte auch der Erfinder Richard Feynman (1918-1989) selbst für die Popularisierung „seiner“ Diagramme (vgl. Feynman, 2014).

Jedoch werden die Diagramme in der theoretischen Physik auf eine andere Weise verwendet als es eine eher „buchstäbliche Lesart“ nahelegen würde. Dies ist ein Hinweis darauf, dass es eine Diskrepanz zwischen der *disziplinspezifischen* und *pädagogischen Affordanz* gibt. Das heißt, die Bedeutung, die der Repräsentationsform innerhalb der Disziplin zugemessen wird, ist nicht ohne weiteres auf seinen Nutzen zum Lehren von physikalischen Inhalten übertragbar (vgl. Airey & Eriksson, 2019; Airey & Linder, 2017).

In der physikdidaktischen Forschung gibt es bisher keine empirischen Untersuchungen dazu, wie mit Feynman-Diagrammen gelernt wird.

Aus diesem Grund untersuchen wir eben diese Wahrnehmung von Feynman-Diagrammen bei Personen unterschiedlicher Expertise. Als Methode zur Untersuchung von Lernprozessen bei visuellen Repräsentationen hat sich Eye Tracking als eine vielversprechende Methode etabliert, mit der z.B. Informationsverarbeitungsprozesse ebenso wie die Effektivität von Instruktionsdesigns erforscht werden können (z.B. Lai et al., 2013). In diesem Beitrag werden die Möglichkeiten von Eye Tracking für die Untersuchung des Themas Feynman-Diagramme diskutiert.

2. Hintergrund

2.1. Lernen mit multiplen Repräsentationen

In der Physik ist das Lernen mit multiplen Repräsentationen (also z.B. einer Text-Bild-Kombination) allgegenwärtig. Dabei haben multiple Repräsentationen beim Lernen drei Schlüsselfunktionen: Sie können sich ergänzen, indem sie komplementär zueinander sind, sie können sich ergänzen, indem sie sich jeweils einschränken, oder sie können ein tieferes Verständnis konstruieren (Ainsworth, 1999, 2006; Opfermann et al., 2017).

Im Falle von Feynman-Diagrammen kann beispielsweise ihre Interpretation anhand von bekannteren Repräsentationen wie Stromkreisen oder des Doppelspaltversuchs eingeschränkt werden (Passon et al., 2020). Ebenso können Feynman-Diagramme zu einem tieferen Verständnis von Teilchenphysik beitragen, indem sie schon vorhandenes Wissen (etwa über die Existenz von Elementarteilchen) erweitern (z.B. indem anhand der Feynman-Diagramme die Idee von virtuellen Teilchen erläutert wird (Jones, 2002)).

Mit dem Lernen sind jedoch verschiedene kognitive Aufgaben verbunden. Zum einen müssen die Lernenden verstehen, wie Informationen in die Repräsentationsform kodiert sind. Außerdem müssen sie verstehen, wie die Repräsentation mit der jeweiligen Domäne in Beziehung steht; sie müssen wissen, wie die richtige Repräsentation ausgewählt und wie sie konstruiert wird (Ainsworth, 2006).

Hier macht die Cognitive Theory of Multimedia Learning (CTML; Mayer, 2002), basierend auf Erkenntnissen zur Informationsverarbeitung im Arbeitsgedächtnis, Vorschläge, wie die Repräsentationen präsentiert werden können, um das Lernen zu unterstützen. So wird z.B. empfohlen, textuelle und bildliche Darstellungen gemeinsam und in räumlicher Nähe zueinander darzustellen (Kontiguitätsprinzip). Im Design der vorliegenden Studie wurde versucht, die Empfehlungen bestmöglich zu berücksichtigen.

2.2. **Feynman-Diagramme im Physikunterricht**

Feynman-Diagramme wurden das erste Mal im Frühjahr 1948 durch Richard Feynman im Rahmen einer Konferenz öffentlich vorgestellt (Kaiser, 2005, S. 43). Seitdem haben sie sich zu einem weit verbreiteten graphischen Werkzeug in vielen Bereichen der theoretischen Physik, vor allem aber in der Quantenfeldtheorie und damit der Teilchenphysik entwickelt. Hauptssächlich werden sie als „Werkzeug zur Buchführung“ bei der störungstheoretischen Berechnung von Prozessen in der Teilchenphysik gebraucht (Kaiser, 2005). Innerhalb der theoretischen Physik (genauer gesagt der Quantenfeldtheorie) besitzen Feynman-Diagramme als semiotische Resource eine hohe disziplinspezifische Affordanz. Diese ist definiert als die vereinbarten bedeutungsstiftenden Funktionen, die eine semiotische Ressource für eine bestimmte Community (die in der Disziplin Tätigen) erfüllt (Airey, 2015; Airey & Eriksson, 2019). Die Diagramme sind dabei eine *disziplinspezifische Kurzschrift* (Airey & Linder, 2017, S. 101 f.) für einen komplizierten mathematischen Ausdruck.

Der disziplinspezifischen Affordanz steht die pädagogische Affordanz gegenüber, die definiert ist als die Eignung, einer Resource, bestimmte Inhalte zu vermitteln (Airey, 2015).

Laut Airey ist jedoch eine hohe disziplinäre Affordanz, wie sie bei Feynman-Diagrammen vorliegt, in der Regel mit einer verminderten pädagogischen Affordanz verbunden. Dies ist insofern bedeutsam, da die Diagramme heute nicht nur in wissenschaftlichen, sondern auch in populärwissenschaftlichen Publikationen und zum Teil auch in Lehr-Lernmaterielien für den Schulunterricht zu finden sind (z.B. Kobel et al., 2018). Hier werden sie in der Regel als (vermeintlich) anschauliche Darstellung teilchenphysikalischer Prozesse verwendet. So zeigen etwa Passon et al., (2018, 2020) verschiedene mit Feynman-Diagrammen verbundene Vorstellungen auf, die so nicht haltbar sind. Etwa ist jede Lesart, in der den Teilchen eine Trajektorie in Zeit und Raum zugeordnet wird, physikalisch unsinnig. Es wurden von verschiedenen Physiker:innen und Physikdidaktiker:innen Versuche unternommen, didaktische Einführungen zu Feynman-Diagrammen zu geben. Pascolini & Pietroni (2002) führten die Diagramme als „akkurate Metaphern“ ein und versuchten, die den Feynman-Diagrammen zugrunde liegenden Regeln anhand eines mechanischen Modells den Lernenden nahe zu bringen. Dieser Ansatz wurde in einem reinen Post-Test-Design empirisch untersucht, wobei ein positiver Einfluss auf das Lernen festgestellt wurde.

Generell unterscheiden sich die Erklärungen dadurch, wie stark auf eine „wörtliche Lesart“ eingegangen wird. In einigen Texten werden die Feynman-Diagramme als Raum-Zeit-Diagramme eingeführt ((Jones, 2002), während in anderen die einzelnen fundamentalen Vertices erläutert und mit mathematischen Termen verknüpft werden (Woithe et al., 2017). Wieder andere Texte gehen explizit auf die Vorstellung einer zu wörtlichen Lesart ein und beschreiben Feynman-Diagramme als einen Beitrag zu einer Wahrscheinlichkeitsamplitude (Allday, 1997; Lambourne, 1992; Passon et al., 2020).

Einen anderen Ansatz wählen Hoekzema et al. (2005). Diese arbeiten in ihrem Text mit einer reduzierten Form der Feynman-Diagramme, welche sie „reaction diagrams“ nennen, um damit Erhaltungssätze und Symmetrien in der Teilchenphysik zu erklären. Der Text wurde an Schulen eingesetzt und erhielt positives Feedback von Lehrpersonen. Insbesondere wurde der Ansatz mit der reduzierten Form der Feynman-Diagramme als verständlicher als ein vorhergehender Text mit herkömmlichen Feynman-Diagrammen beurteilt.

2.3. **Eye Tracking in der physikdidaktischen Forschung**

Eye Tracking ist eine Technik, mit der die Augenbewegungen nachverfolgt werden können. Ihr Einsatz in der Bildungsforschung beruht auf der sog. „Eye-Mind-Hypothese“ (Just & Carpenter, 1976), welche besagt, dass die gegenwärtige Blickposition

Auskunft darüber gibt, wo jeweils die Aufmerksamkeit liegt.

In der Physikdidaktik wird Eye Tracking etwa seit zehn Jahren verwendet. Die Anwendungsfelder reichen hierbei von der Untersuchung von Lernprozessen beim Lernen mit Simulationen (Hoyer & Girwidz, 2020) über die Untersuchung von Assessment-Szenarien mit multiplen Repräsentationen (Rosengrant et al., 2009), den Einfluss aktiver Manipulationen oder instruktiver Hilfestellungen (Klein et al., 2019; Madsen et al., 2013) bis hin zur Nutzung von Eye Tracking-Metriken als Prädiktoren für die Antwortkorrektheit beim Lösen von Aufgaben (Küchemann et al., 2021). Mit Eye Tracking gelang es beispielsweise, postulierte Aufgabenanforderungen beim Arbeiten mit Kinematik-Diagrammen empirisch nachzuweisen (Klein et al., 2018). Auch wurden bekannte Lernschwierigkeiten beim Umgang mit Kinematik-Diagrammen, wie die Punkt-Intervall-Konfusion, durch die Analyse von Blickdaten bestätigt.

Neben diesem konfirmatorischen Charakter zeigt sich das Potential von Eye Tracking auch darin, den visuellen Umgang mit Repräsentationen erstmalig zu erforschen. Im Kontext von Vektorfeldern wurden beispielsweise visuelle Strategien identifiziert, die Expert:innen zur Beurteilung der Divergenz anwenden; sie betrachten die Diagramme überwiegend systematisch mit vertikalen und horizontalen Sakkaden (Klein et al., 2021). Insbesondere im letzteren Fall ist es wichtig, die Methode des Eye Tracking mit anderen Indikatoren für das Verständnis der vorliegenden Repräsentationen zu kombinieren, etwa mit Verbaldaten (van Gog et al., 2005).

Darüber hinaus wird Eye Tracking im Design von Lernmaterialien verwendet, um bestehende Theorien, wie etwa die CTML zu testen und zu erweitern (Alemdag & Cagiltay, 2018; Jarodzka et al., 2017; Mayer, 2010).

Im vorliegenden Fall soll Eye Tracking verwendet werden, um einerseits Hypothesen über Lernschwierigkeiten bei Feynman-Diagrammen zu überprüfen und andererseits Strategien zu identifizieren, die Expert:innen beim Betrachten von Feynman-Diagrammen verwenden.

3. Forschungsziele

Das kurzfristige Ziel der Arbeit ist es, herauszufinden, ob es beim Lernen mit Feynman-Diagrammen spezifische Elemente innerhalb von schriftlichen und graphischen Instruktionen gibt, die lernförderlich oder -hinderlich sind und welche Informationen Eye Tracking hierüber liefern kann.

Außerdem soll herausgefunden werden, wie sich die visuellen Strategien beim Betrachten von Feynman-Diagrammen zwischen Personen mit und Personen ohne Vorwissen und Expertise in der Anwendung von Feynman-Diagrammen unterscheidet.

Mittelfristig soll im Sinne der fachdidaktischen Entwicklungsforschung (*Design-based research;* The Design-Based Research Collective, 2003) mit Hilfe der Erkenntnisse der vorliegenden Studie eine Lerneinheit zum Thema Feynman-Diagramme im Rahmen eines Online-Kurses zu Teilchenphysik für Oberstufenschüler:innen am CERN entwickelt werden.

4. Pilot-Studie

Es wurde eine explorative Studie durchgeführt, anhand derer untersucht werden soll, welche Erkenntnisse bei der Anwendung von Eye Tracking auf Aufgaben mit Feynman-Diagrammen zu erwarten sind.

4.1. Methoden

Zunächst wurden einführende Erklärtexte (Instruktionen) und Aufgaben in englischer und deutscher Sprache zum Thema entworfen. Die Instruktionen umfassten 7 Seiten und waren text- und bildbasiert. Ein Beispiel für eine solche Instruktion ist in Abb. 1 gezeigt.

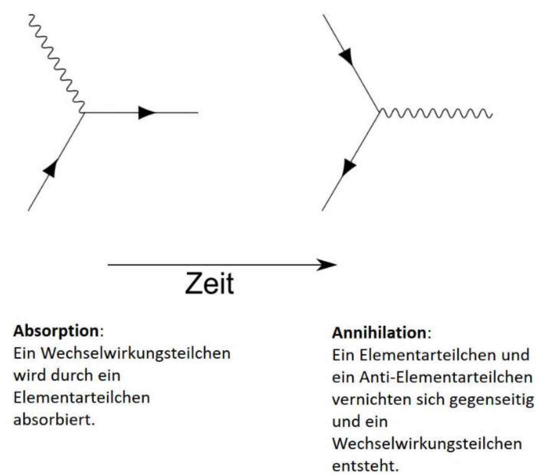


Abb. 1: Ausschnitt aus einer der sieben Instruktionsfolien. Gezeigt sind zwei der vier fundamentalen Vertices. Im Text darunter ist die jeweilige physikalische Interpretation beschrieben (*Absorption*, *Annihilation*). Neben den Vertices und dem Erklärtext ist als zusätzlicher Hinweis der Zeitpfeil als „Leserichtung“ für die Diagramme abgebildet.

Als Basis für die Erstellung der Instruktionen und Aufgaben dienten gängige Textbücher der Elementarteilchenphysik sowie Vorüberlegungen zu potentiellen Lernschwierigkeiten auf Basis von publizierten Lernressourcen (s.o.; z.B. Kobel et al., 2018; Woithe et al., 2017). Besonderer Wert wurde in den Instruktionen auf die Einführung der fundamentalen Vertices sowie die inhärente

Beinhaltung der Erhaltungssätze in den Feynman-Diagrammen gelegt.

Die insgesamt neun zu lösenden Aufgaben wurden in vier Aufgabentypen unterteilt: Im ersten Aufgabentyp sollten die Teilnehmenden beurteilen, ob ein gegebenes Diagramm gemäß der zuvor in den Instruktionen beschriebenen Regeln einen gültigen Prozess darstellt, im zweiten Typ sollte die Anzahl der Vertices angegeben werden, aus denen ein dargestelltes Feynman-Diagramm besteht, und im dritten und vierten Typ sollte ein gegebenes Diagramm mit anderen Diagrammen verglichen werden.

An der Studie nahmen insgesamt 16 Mitglieder der Didaktik-Forschungsgruppen am CERN sowie der Universität Göttingen teil. Der Grad an Expertise reichte von Schulwissen in Physik bis zur Promotion in Teilchenphysik. Während die Teilnehmenden die Instruktionen lasen und die Aufgaben lösten, wurden ihre Blickbewegungen aufgenommen. Die Aufnahme der Blickbewegungen erfolgte mit einem Remote Eye Tracker der Firma Tobii Pro (Fusion, Version 120 Hz). Die Teilnehmenden saßen etwa 80 – 100 cm vom Bildschirm entfernt und konnten sich frei bewegen.

Die 9-Punkt-Kalibration ergab eine Genauigkeit der Blickdaten zwischen 0,31° und 1,33° bei einer Präzision zwischen 0,17° und 0,7° auf einem 24“-Bildschirm mit einer Auflösung von 2560 x 1440 Px. Im besten Fall konnten die Blickbewegungen also auf etwa 5 mm oder 20 Pixel genau aufgelöst werden, im schlechtesten Fall auf 14 mm oder 70 Pixel. Dies korrespondiert in etwa mit den o.g. Materialien, bei denen die kleinsten Einzelstrukturen etwa 5 mm groß waren. Innerhalb der Instruktionen und Aufgaben konnte nicht gesprungen werden und nicht die Richtung gewechselt werden. Die Teilnehmenden hatten jedoch für jede einzelne Seite so viel Zeit wie sie brauchten.

Um zusätzliche Informationen über die Überlegungen der Teilnehmenden zu erhalten, wurden im Anschluss an die Lösung der Aufgaben den Teilnehmenden die Instruktionen und Aufgaben zusammen mit einem Video ihrer Augenbewegungen gezeigt und sie wurden befragt, wie sie mental beim Lösen der Aufgaben vorgegangen sind. Im folgenden wird dieser Teil als Retrospective Think Aloud (RTA) bezeichnet.

4.2. **Erste Ergebnisse**

Anhand der Erklärmuster im RTA wurden die Teilnehmenden in drei Gruppen eingeteilt: „Expert:innen“, „Noviz:innen“ und „Intermediates“. Die Expert:innen zeichneten sich dadurch aus, dass sie beim Erklären auf die physikalischen Begriffe (z.B. *Teilchen*, *Anti-Teilchen*, *Emission*, *Absorption*) referierten, während die Noviz:innen diese Begriffe nicht verwendeten. Bei den Intermediates waren Ansätze von Expertendenken, aber kein kohärentes Muster erkennbar.

Die Eye Tracking-Daten wurden anschließend anhand dieser Einteilung in die verschiedenen Expertise-Gruppen aufgeteilt. Die kumulierten Blickdaten wurden anschließend pro Gruppe gegenüber gestellt und auf qualitative Unterschiede hin miteinander verglichen.

Von den vier Aufgabentypen eigneten sich drei für eine Analyse der Eye Tracking-Daten. Der erste Aufgabentyp, in dem beurteilt werden sollte, ob ein gegebener fundamentaler Vertex (also ein fundamentaler Baustein eines Feynman-Diagramms) den vorher gezeigten Regeln entspricht, war hingegen ein guter Indikator für die oben beschriebene Einteilung.

Im zweiten Aufgabentyp wurde nach der Anzahl der Vertices gefragt, aus denen das gezeigte Diagramm besteht. Für diese Aufgabe waren dementsprechend lediglich die Vertices relevant, also die Bereiche im Diagramm, an dem sich drei Linien treffen. In Abb. 2 sind die kumulierten Blickdaten für die drei Gruppen aufgetragen. Es zeigt sich hier, dass nur die Expert:innen den Großteil ihrer Aufmerksamkeit auf den relevanten Teil der Abbildung verwenden. Zwischen Intermediates und Noviz:innen gibt es in dieser Hinsicht keinen qualitativen Unterschied.

Im dritten Aufgabentyp sollte beurteilt werden, aus welchen der vier auf der linken Seite gezeigten Vertices das auf der rechten Seite gezeigte Diagramm zusammengesetzt ist. Die Blickmuster bei einer Aufgabe dieses Typs ist in Abb. 3 gezeigt. Hier fällt auf, dasss die Expert:innen im Vergleich zu den Intermediates und Noviz:innen relativ viel Aufmerksamkeit auf dem im Sinne der Aufgabe „falschen“ Vertex B verwenden. Rein qualitativ sind auch hier die Unterschiede zwischen Intermediates und Noviz:innen kleiner als die zwischen den Expert:innen und den anderen Gruppen.

Im vierten Aufgabentyp sollte beurteilt werden, ob ein gegebenes Feynman-Diagramm aus den beiden zusätzlich gegebenen Vertices zusammengesetzt sein kann. Die kummulierten Blickdaten für eine Aufgabe dieses Typs sind in Abb. 4 gezeigt. Hier ist auffällig, dass sich die Blickdaten zwischen den einzelnen Gruppen im Gegensatz zu den anderen gezeigten Beispielen qualitativ kaum unterscheiden.

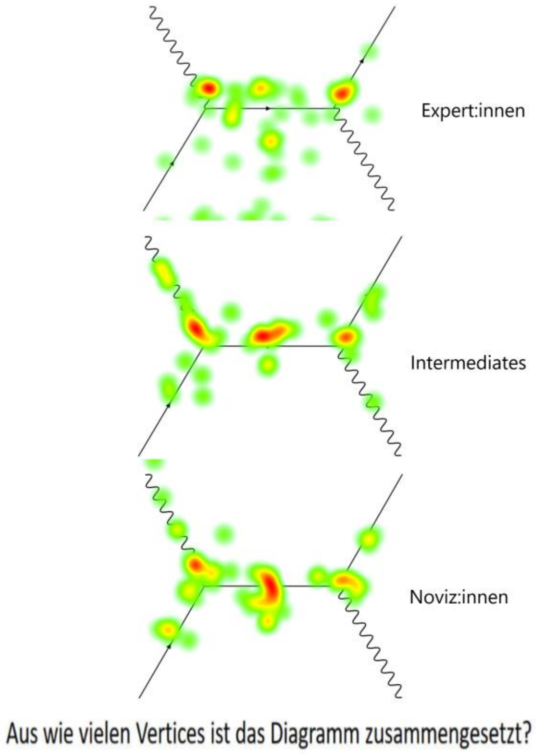


Abb. 2: Kumulierte Blickdaten für Expert:innen (oben), Intermediates (Mitte) und Noviz:innen (unten) bei einer Aufgabe des Typs 2 („Aus wie vielen Vertices besteht das Diagramm?“)

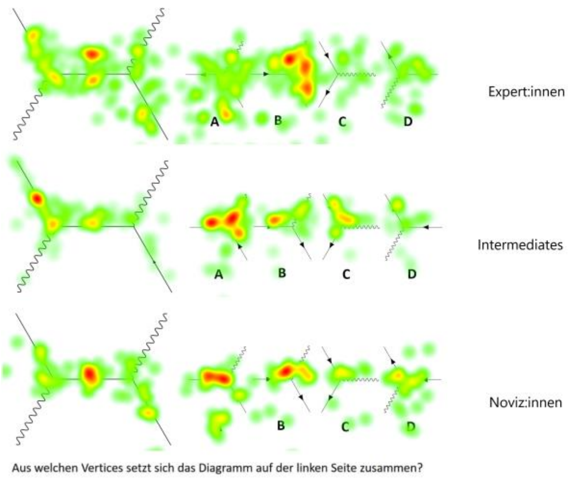


Abb. 3: Kummulierte Blickdaten für Expert:innen (oben), Intermediates (Mitte) und Noviz:innen (unten) bei einer Aufgabe des Typs 3 („Aus welchen Vertices setzt sich das Diagramm auf der linken Seite zusammen?“). Die richtige Antwort ist in diesem Fall D + A.

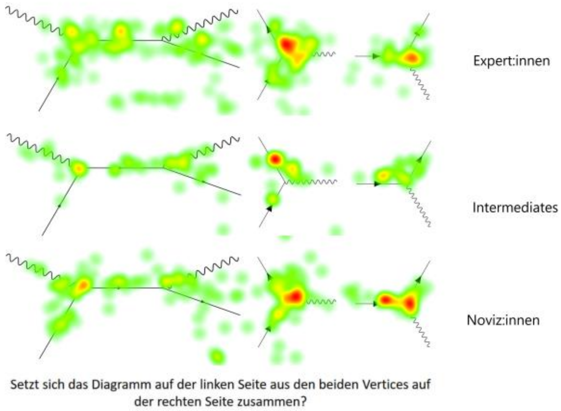


Abb. 4: Kummulierte Blickdaten für Expert:innen (oben), Intermediates (Mitte) und Noviz:innen (unten) bei einer Aufgabe des Typs 4 („Setzt sich das Diagramm auf der linken Seite aus den beiden Vertices auf der rechten Seite zusammen?“). Die Antwort ist in diesem Fall „Nein“.

5. Diskussion

In der vorliegenden Pilotstudie wurde erstmals die Wahrnehmung von Lernmaterialien zum Thema Feynman-Diagramme empirisch mit Hilfe von Eye Tracking untersucht. Die Blickmuster in Abb. 2 lassen darauf schließen, dass bei Nennung bestimmter Begriffe in einer Aufgabe bei Personen mit Vorwissen der Blick auf bestimmte Stellen in der Repräsentation gelenkt wird. Dies ist im Einklang mit bisherigen Eye Tracking-Studien, in denen ebenfalls gezeigt werden konnte, dass Expert:innen relativ viel visuelle Aufmerksamkeit auf relevante als auf irrelevante Bereiche der Aufgabe verwenden und insgesamt weniger Zeit zum Lösen der Aufgaben brauchen als Noviz:innen (vgl. Klein, et al., 2019; A. M. Madsen et al., 2012).

Wie die Abbildungen 3 und 4 zeigen, ist dieses Ergebnis in unserer Studie jedoch nicht persistent über alle Aufgabentypen. So kam es auch vor, dass Expert:innen relativ gesehen mehr Zeit auf einer falschen Antwortoption verbrachten als Intermediates und Noviz:innen, oder es konnten keine Unterschiede zwischen den Gruppen beobachtet werden. Dies kann unterschiedliche Gründe haben. So erwähnten mehrere Expert:innen in den RTA-Daten, dass sie versucht hatten, die einzelnen Vertices mental zu rotieren, um sicher zu gehen, dass der Vertex nicht doch eine Option sein könnte. Noviz:innen erwähnten dies nicht.

Eine Einschränkung dieser Pilotstudie ist ihre kleine Stichprobe sowie die fehlende Repräsentativität in Bezug auf die Zielgruppe der Materialien (Oberstufenschüler:innen). Aus diesem Grund haben wir uns für eine qualitative und heuristische Form der Auswertung entschieden. Die vorliegenden Ergebnisse können lediglich ein Hinweis auf weitere Forschung sein.

6. Ausblick

Im Anschluss an die vorgestellte Pilotstudie soll eine Hauptstudie durchgeführt werden, in der eine größere Stichprobe von Schüler:innen, Studienanfänger:innen sowie Physiker:innen untersucht wird. Die Instruktionen und Aufgaben dieser Studie werden dabei im Sinne einer didaktischen Rekonstruktion auf Basis von Literatur, den Erkenntnissen aus der Pilotstudie sowie Interviews mit Expert:innen erstellt. Die Erkenntnisse aus dieser Studie sollen anschließend genutzt werden, um Lehr-Lernmaterialien zu erstellen. Hier könnten auch interaktive Aufgaben, wie etwa das Zeichnen von Diagrammen (vgl. Ainsworth et al., 2011) oder sog. *Eye-Movement-Modeling-Examples* (vgl. van Gog et al., 2009) zum Einsatz kommen.

7. Literatur

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*(2), 131–152. https://doi.org/10.1016/S0360-1315(99)00029-9

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations. *Learning and Instruction*, *16*(3), 183–198. https://doi.org/10.1016/j.learninstruc.2006.03.001

Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to Learn in Science. *Science*, *333*(6046), 1096–1097. https://doi.org/10.1126/science.1204153

Airey, J. (2015). *Social Semiotics in Higher Education: Examples from teaching and learning in undergraduate physics*. 103. http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-266049

Airey, J., & Eriksson, U. (2019). Unpacking the Hertzsprung-Russell Diagram: A Social Semiotic Analysis of the Disciplinary and Pedagogical Affordances of a Central Resource in Astronomy. *Designs for Learning*, *11*(1), 99–107.

Airey, J., & Linder, C. (2017). Social Semiotics in University Physics Education. In D. F. Treagust, R. Duit, & H. E. Fischer (Hrsg.), *Multiple Representations in Physics Education* (S. 95–122). Springer International Publishing. https://doi.org/10.1007/978-3-319-58914-5_5

Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education*, *125*, 413–428. https://doi.org/10.1016/j.compedu.2018.06.023

Allday, J. (1997). The nature of force in particle physics. *Physics Education*, *32*(5), 327. https://doi.org/10.1088/0031-9120/32/5/016

Feynman, R. P. (2014). QED. In *QED*. Princeton University Press. https://www.degruyter.com/document/doi/10.1515/9781400847464/html

Hoekzema, D., Schooten, G., van den Berg, E., & Lijnse, P. (2005). Conservation Laws, Symmetries, and Elementary Particles. *The Physics Teacher*, *43*(5), 266–271. https://doi.org/10.1119/1.1903808

Hoyer, C., & Girwidz, R. (2020). Animation and interactivity in computer-based physics experiments to support the documentation of measured vector quantities in diagrams: An eye tracking study. *Physical Review Physics Education Research*, *16*(2), 020124. https://doi.org/10.1103/PhysRevPhysEducRes.16.020124

Jarodzka, H., Gruber, H., & Holmqvist, K. (2017). *Eye tracking in Educational Science: Theoretical frameworks and research agendas*. https://doi.org/10.16910/jemr.10.1.3

Jones, G. T. (2002). The uncertainty principle, virtual particles and real forces. *Physics Education*, *37*(3), 223–233. https://doi.org/10.1088/0031-9120/37/3/306

Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*(4), 441–480. https://doi.org/10.1016/0010-0285(76)90015-3

Kaiser, D. (2005). *Drawing Theories Apart: The Dispersion of Feynman Diagrams in Postwar Physics*. University of Chicago Press.

Klein, P., Becker, S., Küchemann, S., & Kuhn, J. (2021). Test of understanding graphs in kinematics: Item objectives confirmed by clustering eye movement transitions. *Physical Review Physics Education Research*, *17*(1), 013102. https://doi.org/10.1103/PhysRevPhysEducRes.17.013102

Klein, P., Küchemann, S., Brückner, S., Zlatkin-Troitschanskaia, O., & Kuhn, J. (2019). Student understanding of graph slope and area under a curve: A replication study comparing first-year physics and economics students. *Physical Review Physics Education Research*, *15*(2), 020116. https://doi.org/10.1103/PhysRevPhysEducRes.15.020116

Klein, P., Viiri, J., & Kuhn, J. (2019). Visual cues improve students’ understanding of divergence and curl: Evidence from eye movements during reading and problem solving. *Physical Review Physics Education Research*, *15*(1), 010126.

https://doi.org/10.1103/PhysRevPhysEducRes.15.010126

Klein, P., Viiri, J., Mozaffari, S., Dengel, A., & Kuhn, J. (2018). Instruction-based clinical eye-tracking study on the visual interpretation of divergence: How do students look at vector field plots? *Physical Review Physics Education Research*, *14*(1), 010116. https://doi.org/10.1103/PhysRevPhysEducRes.14.010116

Kobel, M., Bilow, U., Lindenau, P., & Schorn, B. (2018). *Teilchenphysik: Ladungen, Wechselwirkungen und Teilchen—Unterrichtsmaterial ab Klasse 10*. Joachim Herz Stiftung. https://www.teilchenwelt.de/fileadmin/user_upload/Redaktion/Netzwerk_Teilchenwelt/material/Band1_UnterrichtsmaterialTP_2018.pdf

Küchemann, S., Klein, P., Becker, S., Kumari, N., & Kuhn, J. (2021). *Classification of Students' Conceptual Understanding in STEM Education using Their Visual Attention Distributions: A Comparison of Three Machine-Learning Approaches*. 36–46. https://www.scitepress.org/PublicationsDetail.aspx?ID=RR+w7e5T9Tc=&t=1

Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., Lee, M.-H., Chiou, G.-L., Liang, J.-C., & Tsai, C.-C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*, 90–115. https://doi.org/10.1016/j.edurev.2013.10.001

Lambourne, R. (1992). Predicting the physics of particles. *Physics Education*, *27*(2), 71. https://doi.org/10.1088/0031-9120/27/2/003

Madsen, A. M., Larson, A. M., Loschky, L. C., & Rebello, N. S. (2012). Differences in visual attention between those who correctly and incorrectly answer physics problems. *Physical Review Special Topics - Physics Education Research*, *8*(1), 010122. https://doi.org/10.1103/PhysRevSTPER.8.010122

Madsen, A., Rouinfar, A., Larson, A. M., Loschky, L. C., & Rebello, N. S. (2013). Can short duration visual cues influence students' reasoning and eye movements in physics problems? *Physical Review Special Topics - Physics Education Research*, *9*(2), 020104. https://doi.org/10.1103/PhysRevSTPER.9.020104

Mayer, R. E. (2002). Multimedia learning. In *Psychology of Learning and Motivation* (Bd. 41, S. 85–139). Academic Press. https://doi.org/10.1016/S0079-7421(02)80005-6

Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and Instruction*, *20*(2), 167–171. https://doi.org/10.1016/j.learninstruc.2009.02.012

Opfermann, M., Schmeck, A., & Fischer, H. E. (2017). Multiple Representations in Physics and Science Education – Why Should We Use Them? In D. F. Treagust, R. Duit, & H. E. Fischer (Hrsg.), *Multiple Representations in Physics Education* (S. 1–22). Springer International Publishing. https://doi.org/10.1007/978-3-319-58914-5_1

Pascolini, A., & Pietroni, M. (2002). Feynman diagrams as metaphors: Borrowing the particle physicist's imagery for science communication purposes. *Physics Education*, *37*(4), 324. https://doi.org/10.1088/0031-9120/37/4/306

Passon, O., Lindenau, P., & Kobel, M. (2020). Von Feynman-Diagrammen und Stromkreisen. *Unterricht Physik*, *180*. https://www.friedrich-verlag.de/physik/quantenphysik/von-feynman-diagrammen-und-stromkreisen-7480

Passon, O., Zügge, T., & Grebe-Ellis, J. (2018). Pitfalls in the teaching of elementary particle physics. *Physics Education*, *54*(1), 015014. https://doi.org/10.1088/1361-6552/aadbc7

Rosengrant, D., Thomson, C., & Mzoughi, T. (2009). Comparing Experts and Novices in Solving Electrical Circuit Problems with the Help of Eye-Tracking. *AIP Conference Proceedings*, *1179*(1), 249–252. https://doi.org/10.1063/1.3266728

The Design-Based Research Collective. (2003). Design-Based Research: An Emerging Paradigm for Educational Inquiry. *Educational Researcher*, *32*(1), 5–8. https://doi.org/10.3102/0013189X032001005

van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model's eye movements. *Computers in Human Behavior*, *25*(3), 785–791. https://doi.org/10.1016/j.chb.2009.02.007

van Gog, T., Paas, F., van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the Problem-Solving Process: Cued Retrospective Reporting Versus Concurrent and Retrospective Reporting. *Journal of Experimental Psychology: Applied*, *11*(4), 237–244. https://doi.org/10.1037/1076-898X.11.4.237

Woithe, J., Wiener, G. J., & Veken, F. F. V. der. (2017). Let's have a coffee with the Standard Model of particle physics! *Physics Education*, *52*(3), 034001. https://doi.org/10.1088/1361-6552/aa5b25

Danksagung

Wir danken Julia Woithe für hilfreiche Hinweise zum Materialdesign sowie Stefanie Peter für ihre Hilfe bei der Durchführung der Studie. Diese Arbeit wurde gefördert durch das Wolfgang-Gentner-Programm des Bundesministeriums für Bildung und Forschung
(BMBF-Forschungsvorhaben 05E18CHA).

6.4 First student study

Following the pilot study with members from the working group who, in part, had a background in (particle) physics, I conducted an eye-tracking study with high school students. The nature of this study was, again, rather exploratory. The research question of this study was: *What are some gaze patterns high-school students use when confronted with Feynman diagram-like representations?*. The term "Feynman diagram-like" is used here because the diagrams used in the learning material draft of this study were educational reconstructions of Feynman diagrams of the form presented in Figure 5.2 on the left side where the arrows pointed only in one direction or no arrows were used at all.

6.4.1 Description of the student sample

These students were at CERN as part of the *High-School Student Internship Program* (HSSIP)[2]. Within this programme, 24 students per country could come to CERN for two weeks to visit the facilities, have lectures, and do workshops. The students participating in this programme were generally highly motivated by particle physics. Of these students, 51 participated in the study (12 from Belgium, seven from Denmark, 17 from Greece, 10 from Romania, and five from Switzerland). The students were between 16 and 20 and, on average, 17 years old (standard deviation: 1 year). Of these students, 33 said that they had heard about particle physics in school, and 35 said that they had learned about particle physics out of school, e.g., from books or online videos, and all of them had lectures about particle physics during their stay at CERN.

6.4.2 Design of the learning material

This study tested a first draft of the learning material. The draft only incorporated the first learning goal. The design approach was to start very easily and then successively adjust to students' needs. The downside of this approach was that the material was altered several times during the study, so the data taken with the first students could ultimately not be compared to those taken with the last students. The design of the learning material followed the design principles presented in Section 5.1.2. I used graphics, mainly simplified Feynman diagrams and explanatory text, to convey the key messages in Section 6.2. The simplified Feynman diagrams differed from conventional Feynman diagrams by having only one type of line and no arrows. Also, I used charge annotations so the students do not have to remember any charge numbers but can directly read them from the diagram. The principle of the learning material was still very similar between the first and the last draft used in that study. It was designed using *Microsoft PowerPoint* in a way that it could be

[2]The programme has been discontinued, but an overview is still given on its website (CERN, 2023).

easily used within the eye tracking software (cf. Section 5.2). Therefore, the learning material did not include any videos, simulations, or interactive elements, which would depend on the participant's input.

The introduction of the final version of the material used in this study comprised the following elements. The single steps are illustrated in Figure D.1 in the appendix.

- A motivation of the topic by stating that in particle accelerators, particles collide and that we can draw diagrams of this process.
- A first diagram of a scattering process, in which the elements of a diagram are described, namely the names of particles using letters, like "*e*" for "electron", charge annotations, and arrows (during this study arrows were exclusively used in one direction, or lines with no arrows at all) which signify interactions between particles.
- It is shown what the sum of the electric charge in the initial state is, and from that, it is motivated what the charges of the product of that interaction have to be.
- The introduction of another charge besides the already known electric charge, the so-called weak charge.

Building on this introduction, students were shown a series of tasks, one example shown in Figure 6.1 on the top. In these tasks, students were shown a scattering-type diagram but with "imaginary particles", i.e., particle names and charges without any correspondence in physics, which they had to check for charge conservation, i.e., choose for electric and weak charge separately whether it is conserved. After they clicked the solution and verbally gave a reason for their solution, they were shown the solution. The number of tasks varied throughout the study, as it was quickly noticed that the tasks were all answered correctly by the students and seemed to be no challenge for students, so their number was reduced.

After these scattering-type diagrams, a diagram of the decay type was introduced, this time with a physical example, namely the leading-order diagram for the muon transformation (i.e., the process $\mu^- \rightarrow \nu_\mu \overline{\nu_e} e^-$ mediated by a W^- boson). After the participating particles in this process (muon, neutrino, antineutrino, electron, W boson) were introduced and for some of them the electric and weak charge was mentioned (electric charge of moun, electron, neutrino, and anti-neutrino), they were asked what the electric charge of the W boson was in the diagram. Once they answered the question and gave a reason for this verbally, they were shown the solution to this question and the weak charges of the muon, electron, and neutrino were introduced, from which they were asked to deduce the weak charge of the W boson. Once this question was solved, they were asked to deduce the weak charge of the anti-neutrino. Figure 6.1 at the bottom shows the first of these questions, while the other tasks of this ty[e are shown in Figure D.3 in the appendix. Students took on

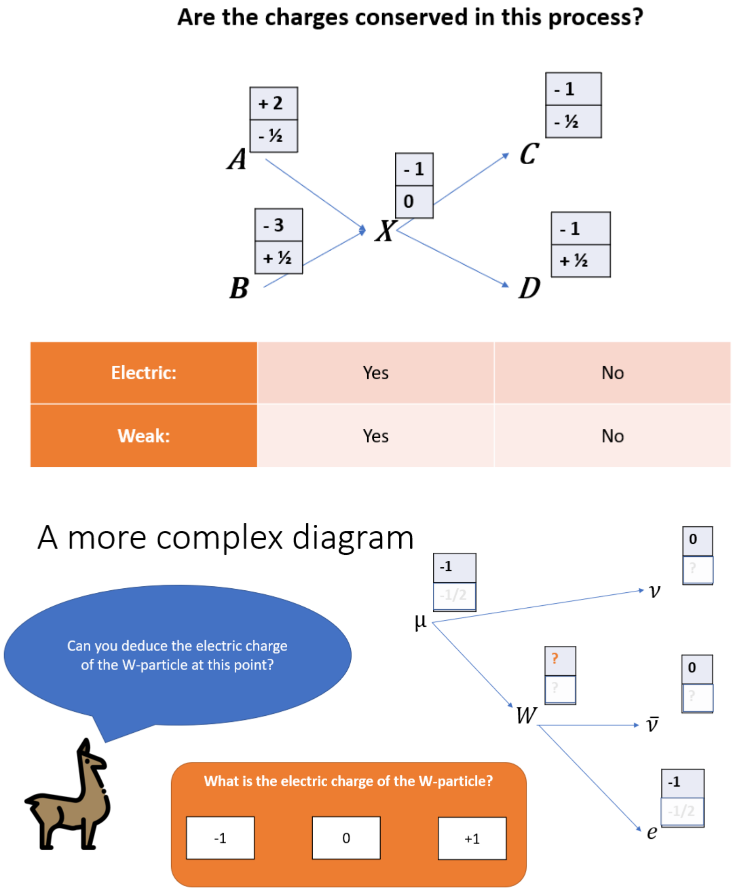


FIGURE 6.1: Example tasks from the first student study. **Top:** Task about checking scattering-type diagrams for charge conservation. **Bottom:** Task about completion of a decay-type diagram. The question is, "What is the electric charge of the W-particle?" The question refers to the diagram displayed below. The students are expected to answer with the correct answer of "-1". They are then asked to give a brief explanation of their answers orally.

average 11 minutes (standard deviation: 3 minutes) to work through this learning material. After they worked through the learning environment, they were shown their own eye movements (cf. Section 5.2.3). These retrospective think aloud interviews took on average 17 minutes (standard deviation: 5 minutes).

6.4.3 Students' cognitive load

After completing the learning environment, students completed the cognitive load questionnaire described in Section 5.3.2 and presented in Table B.3. The questionnaire assessed three types of cognitive load: intrinsic, extraneous, and germane load. Table 6.1 shows the means and standard errors of the three scales alongside the values for internal consistency, measured by Cronbach's α, of the three scales (cf. Section 5.4.1.

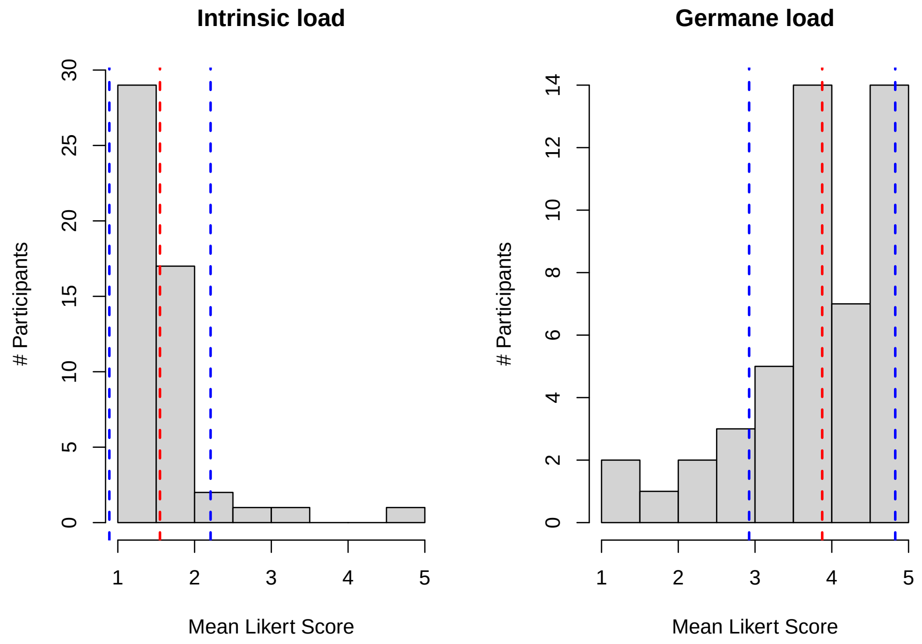


FIGURE 6.2: Histograms of two dimensions of the Cognitive load questionnaire. The red vertical line is the mean of the scale, while the blue lines depict one standard deviation around the mean. As seen in Table 6.1, the internal consistency of the extraneous load scale is very low, which is why its distribution is not shown.

Scale	Mean	α
IL	1.55 ± 0.10	.78
EL	1.86 ± 0.09	.26
GL	3.88 ± 0.14	.81

TABLE 6.1: Table with a summary of the three scales of the Cognitive Load questionnaire: intrinsic load (IL), extraneous load (EL), and germane load (GL). The scores on each scale reach from 1 to 5. The mean and Cronbach's α for the extraneous load scale were calculated by omitting item five of the questionnaire since α would have been negative otherwise.

Particular attention should be drawn to the extraneous load scale. For this scale, the initial α was -0.05, which means that at least one item was even slightly negatively correlated with the others. This item was number five ("I did not learn a lot from the explanations."). A possible explanation would be that some students read the question incorrectly (e.g., by omitting the "not"). Even when omitting this item, the internal consistency was still very low. Therefore, the extraneous load scale is deemed unreliable within this study.

For the other two scales, the distribution of scores is shown in Figure 6.2. It is visible that the intrinsic load, i.e., the perceived difficulty of the material's content, peaks at the very bottom of the scale. This means that students perceived this version of the learning material as very easy. Compared to that, the germane load distribution was in the upper half of the scale, so students still felt that they could learn something from the learning material.

6.4.4 Students' perceptions of Feynman diagrams

The study was the first to analyse high school students' visual perceptions of Feynman diagrams. No systematic data analysis was performed because the learning material was altered often. Instead, the study was used as an exploratory pre-study to inform the design process for later studies. It was already apparent in the students' answers to the questions in the learning environment that ceiling effects were reached, as most, if not all, of the students answered all the questions correctly (cf. also the intrinsic load scale in Figure 6.2). Therefore, a detailed item analysis of these questions is omitted.

Even though the data gathered in this first student study were not systematically analysed, some results can be presented qualitatively using heat maps (cf. Section 5.2.4 to explain this type of data visualisation). The heatmaps in Figure 6.3 show the aggregated gaze distribution of 27 students on three different tasks based on the same decay-type diagram. The visualisation can only be shown for this subset of students because this slide was only presented to this subset of participants. The patterns in these heatmaps allow for some tentative conclusions about students' gaze patterns on Feynman diagrams. For the first task (for which the stimulus without a heatmap is shown in Figure 6.1 on the bottom right), for which the task was to find the electric charge of the W particle, which is in the centre, the hotspots of the attention are on that said electric charge, which is a question mark in that stimulus, on the electric charge of the muon (on the top left), and on the electric charge of the neutrino (on the top right). This is reflected by the fact that students used the latter two to deduce the electric charge of the W particle. Smaller hotspots are on the electric charge of the electron and on the anti-neutrino, which is reflected by the fact that few students used these two charges to deduce the electric charge (more on these strategies in the second student study, Section 6.5.2). However, the hotspot on the anti-neutrino is less pronounced than that on the electron, which wouldn't be the case if both charges were used equally to deduce the electric charge of the W particle. This could hint that the electron was also used with the muon, as depicted in one line with the W particle.

An analogue explanation can be given for the other two heat maps in Figure 6.3. The middle one shows the task where students were asked what the weak charge of the W particle was. There, the hotspots are clearly on the weak charges of the muon and the neutrino, the only two weak charges that can be used directly to deduce the weak charge of the W particle. Less pronounced spots can also be found on the weak charges of the anti-neutrino and the electron. Even though the weak charge of the anti-neutrino is not yet known, some students deduced it from the weak charge of the neutrino on top of the fact that the charges of a particle are always the opposite of the charge of its anti-particle. A similar pattern can be seen on the heatmap of the third task, where the question was to deduce the weak charge of the anti-neutrino. Here, the predominant way was to deduce it from the charges of the electron and the

W particle, but a small fraction also deduced it from the fact that the anti-neutrino should have the opposite charge of the neutrino.

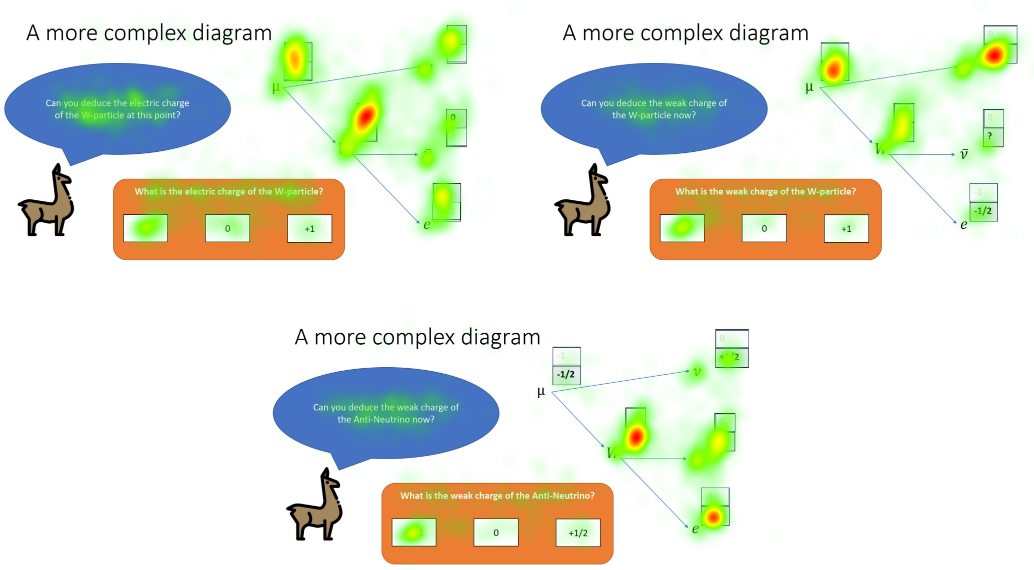


FIGURE 6.3: Heatmaps of the aggregated gaze patterns of 27 students on the three questions about the decay-type diagrams.

In summary, this inspection of heat maps already revealed some interesting facts about how students perceived the diagrams in the learning environment. These strategies were more closely investigated in a second student study, where the learning material was slightly adapted.

The full version of the learning material at the end of this study, the answers to the cognitive load questionnaire, as well as analysis scripts and additional heat maps are publicly available. The links can be found in appendix D.

6.5 Second student study

The second student study used an updated version of the learning material and had a more explicit focus than the first student study, which was somewhat exploratory. The research question of this study was: *Which visual strategies do students use when examining Feynman diagrams to learn about charge conservation?* It was part of the second research goal stated in Section 3.2 in the way that from the strategies students used, practices to better present the Feynman diagrams in a meaningful way could be determined.

This study's version of the learning material also incorporated the second learning goal by including explanations about interaction particles. The learning material also included questions to summarise the taught concepts. However, the objective of this study was to analyse the visual strategies; therefore, a systematic analysis of these summaries was not the main focus and was eventually omitted.

6.5.1 Materials and Methods

Design of Learning Materials

The learning material used in this study was an updated version of the material used in the first student study as described in Section 6.4.2 and shown in the appendix chapter D. The learning material focused, in particular, on the decay-type diagram already discussed above, a diagram where one particle transforms into two particles, of which one or more particles transform into other particles. This diagram was already investigated in the previous study but gained particular attention because it is beneficial for students to understand the mechanics of this type of diagram to combine weak interaction and charge conservation. Also, it was for the first time that the learning material incorporated the second learning goal. The explanation of interaction particles mainly was along the explanation of Allday (1997) (cf. Section 2.2.2)where he explains muon transformation mediated by the W-particle through interaction with the weak field. The steps are illustrated in the appendix in Figure D.4. As in the previous study, the diagrams were explained step by step, and questions were incorporated to collect data about the students' understanding of the diagrams and enhance their learning experience. The students answered the questions orally during the intervention to reflect on their learning and gather data.

Instruments

The students answered a pre-questionnaire to gather demographic information, to assess their English language skills roughly, and their prior knowledge of particle physics (the instrument is described in Section 5.3.1 and presented in the appendix table B.1). After the intervention, they answered a questionnaire about their cognitive load, which is described in Section 5.3.2 and presented in the appendix table B.3) and had the opportunity to give feedback about the learning material.

Learning environment questions. In the learning material, participants were confronted with questions similar to those in the first student study (cf. Section 6.4), which can be classified into two question types: a) in the first question type, the students should check if a given Feynman diagram fulfils charge conservation by comparing the charges on the left and the right side of the diagram. The students should simultaneously check for the conservation of electric and weak charge. b) For a given Feynman diagram that the students haven't seen before, they should give the correct weak or electric charge for certain particles, assuming that the respective charge is conserved. They should also give a reason for their answer. One example of a question of the second type is given in Figure 6.1. The students answered all questions orally, and their answers were audiotaped.

Description of the Study

The students who participated in the study were recruited voluntarily. After the procedure was explained to them and they gave informed consent, they were seated in front of a computer screen with a Tobii Fusion eye tracker (see Section 5.2.1). Their voice and eye movements on the screen were recorded during the process. At first, they answered the pre-questionnaire described above. After that, they were shown the learning material, which they could click to move forward. When questions came up, they answered them verbally and were asked by the investigator to reason their answers. After finishing the learning material, they were redirected to the post-questionnaire, which included the cognitive load questionnaire and the possibility of giving feedback.

Description of the Sample of Students

The study was conducted at the European Organization for Nuclear Research (CERN) with high school students within the age range of 16-19 (average age 17.8 years, standard deviation 0.8 years) who came to CERN within two different programs: a) Part of the students were at CERN in the framework of a 2-week internship program during which they worked on projects under the supervision of CERN scientists and had lectures about different particle physics related topics[3]. b) Another part of the students was at CERN as part of a one-day visit together with their school class to participate in two different physics workshops[4].

The students' participation in the study was voluntary. Overall, 21 students from five countries participated in the study: 14 as part of the internship program from Serbia and Italy, six as part of a one-day visit to CERN from the Netherlands and Portugal, and one as an individual intern from Germany. The students answered the question in English, except for the German student, who answered in German since their native language was German. Data about the gender of 19 students is available; of these, 10 identified as male, eight identified as female, and one identified as non-binary. The students assessed their English language skills in reading, listening, writing, and speaking on a 5-point scale between "very poor" and "very good". In all four categories, the students only rated themselves as "OK" (point 3) or better. The descriptive results of this assessment are presented in Table 6.2. Thus, the students' overall English self-assessment is very good. It took the students on average 28 minutes to go through the learning material (standard deviation: 8 minutes). The significantly longer time than in the

Skill	Mean	SD
Reading	4.8	0.4
Listening	4.6	0.6
Writing	4.6	0.6
Speaking	4.5	0.7

TABLE 6.2: Self-assessed English skills of $N = 21$ students in the second student study

[3]the HSSIP program also described in Section 6.4, see footnote 2 on page 93.
[4]The visits took place in the framework of the S"Cool Lab PLUS+ program.

first study (cf. page 95) can be explained by incorporating the second learning goal into the draft of the learning material.

6.5.2 Results

Prior Knowledge

Students' prior knowledge was assessed using the instrument presented in Section 5.3.1 in Table B.1. The average score achieved by the students was 5.00 (out of a maximum of 9), with a standard deviation of 2.19. The scores were then placed into three categories, s.t. each category would account for about 1/3 of the participants. Scores of 3.5 or lower were called "Low", scores of 4 to 6.5 were called "Medium", and scores higher than 6.5 were called "High". With that, six students achieved a high score, eight a medium score, and seven a low one. The internal consistency of the questions was with $\alpha = 0.81$ and KR-20 = 0.70 in an acceptable range, even though they are at the lower edge of what is typically deemed acceptable. Adams and Wieman (2011) state that a high value in a reliability index would not necessarily be desirable as it would be an indication of redundant items (cf. Section 5.4.1). The discriminatory power, measured by Ferguson's δ, is 0.92 and therefore also in an acceptable range (Ding & Beichner, 2009). Table 6.3 shows a per-item summary of the nine items. It shows that the item difficulty varied widely, which is desirable according to Adams and Wieman (2011). The item discrimination also varied but was mostly within the acceptable range apart from item 8, which was also the most difficult item ("Which path does a particle take between two points in space?"). The point biserial coefficient r_{pbi} was acceptable for all items.

Item	Correct	Partially correct	Incorrect	p	D	r_{pbi}
1	12	8	1	0.76	0.35	0.40
2	12	3	6	0.64	0.79	0.76
3	10	0	11	0.48	1	0.83
4	8	2	11	0.43	0.76	0.75
5	8	11	2	0.64	0.32	0.35
6	12	5	4	0.69	0.57	0.61
7	6	6	9	0.43	0.62	0.45
8	2	0	19	0.10*	0.17*	0.41
9	14	7	0	0.83	0.43	0.49
Total	84	42	63	α	KR-20	δ
				0.81	0.7	0.92

TABLE 6.3: Summary of the prior knowledge questionnaire. The items and scoring scheme are presented in Table B.1. p denotes the average score obtained on an item, D denotes the discriminatory index, and r_{pbi} denotes the point-biserial coefficient. α, KR-20, and δ denote Cronbach's α, the Kuder-Richardson index, and Ferguson's δ. Values which are outside the acceptable range according to Ding and Beichner (2009) are noted with an asterisk.

The test was only administered before the learning material because it primarily assessed declarative knowledge. An administration of the test after the learning material would have mostly tested the mere retention of facts, which would not be the learning goal I wanted to achieve with the material.

Answers to the Questions in the Learning Environment

For the first learning objective, the students mainly gave the correct answer. In the first question of the first question type ("Are the charges conserved?"), 20 of 21 gave the correct answer. The one student who did not answer correctly corrected themselves after being asked to check and gave the correct reasoning. For the second question type ("What is the electric/weak charge of the W-particle/neutrino?"), the picture was more mixed. Here, 15 students initially gave the correct answer to the first question. All six who didn't answer correctly corrected themselves when asked to check their answer and gave the correct reasoning. No item analysis using psychometric indicators is performed for these questions, because students saw the solution to one question before they answered the next one, therefore an item analysis would be of very limited use in this case. Of the six students who answered the first question of the second question type incorrectly, two had achieved a low score in the prior knowledge test, two a medium score, and two a high score. For the second question of this question type, only two students did not initially give the correct answer, one of whom also did not initially give the correct answer in the first question (both had achieved a low score in the prior knowledge questionnaire). For the third question, all students initially gave the correct answer. There were two different strategies for answering the first question of the second question type (Bottom slide of Figure 6.1): One used the vertex formed by the particles on the top of the diagram (muon, neutrino, and W-particle, further referred to as *production vertex*) and completed the equation -1 = 0 + x, while the other used the vertex formed by the particles on the bottom right of the diagram (W-particle, electron, and anti-neutrino, further referred to as *decay vertex*[5]) and completed the equation x = 0 -1. In their verbalized reasoning, students used both of these strategies. I categorised the students' reasoning for their answers to this question into three categories: whether they mentioned only the particles in the production vertex, only the particles in the decay vertex, or a mix of both. Overall, seven students mentioned exclusively the production vertex in their answers, while six mentioned exclusively the decay vertex. Another eight students mentioned the production and decay vertex, either because they used a different strategy than the ones described above or because they used both strategies after each other. Four students who answered incorrectly belonged to this "mixed" group. Partly, they corrected themselves using the other particles in their correction, and partly, they noticed a mistake in their original reasoning.

[5]The term "decay" is used here purely for descriptive purposes within this work to distinguish the two types of vertices. See also footnote 1 about the "decay diagram" on page 44.

Cognitive Load

Like in the first student study, students filled out a cognitive load questionnaire at the end of the session. The questionnaire is presented in Section 5.3.2 and shown in Table B.3 in comparison with the questionnaire used in the first study. As seen in Section 6.4.3, the internal consistency of the extraneous load scale was extremely low. A similar problem occurred with the refined version of the questionnaire. Cronbach's α, in this case, was 0.55, which is higher than the corrected α in the previous case, but still a very low value. Therefore, item 5, which wasn't changed from the first to the second study, was omitted again, and the corrected α was calculated. This time, it was in an acceptable range. The values for the means and standard errors of the three scales, alongside the values for Cronbach's α, are shown in Table 6.4. The corresponding distributions of the scales are shown in Figure 6.4. It is visible that the intrinsic load is, compared to the distribution in the first student study, presented in Figure 6.2, much broader distributed and generally higher. Therefore, students perceived this refined learning material as more difficult.

Scale	Mean	α
IL	2.51 ± 0.24	.82
EL	1.36 ± 0.13	.77
GL	3.86 ± 0.29	.95

TABLE 6.4: Table with a summary of the three scales of the Cognitive Load questionnaire: intrinsic load (IL), extraneous load (EL), and germane load (GL). The scores on each scale reach from 1 to 5. The mean and Cronbach's α for the extraneous load scale were calculated by omitting item five of the questionnaire.

Differences between populations

The intrinsic load scale in Figure 6.4 shows a bimodal distribution, which can be explained by the two populations participating in the study. Figure 6.5 (left) shows the distribution of intrinsic load scores discriminated by population. The "intern" students (the students from the HSSIP program and the individual German intern) perceived the learning material as more effortless. A Welch Two Sample t-test reveals that the means in intrinsic load of the groups ($M = 3.2, SD = 0.5$ and $M = 2.2, SD = 1.1$ for the S'Cool Lab+ and HSSIP group, respectively) are statistically significant ($t(18.5) = 2.83, p = 0.01$ with an effect size of $d = 1.15$). In the right histogram in Figure 6.5, this split is even more apparent as it shows the distribution of the prior knowledge scores discriminated by the populations. A t-test also reveals a significant difference between the HSSIP group ($M = 5.9, SD = 1.8$) and the S'CoolLab+ group ($M = 2.7, SD = 1.4$), $t(11.5) = 4.4, p = .001$, with a large effect size of $d = 2.01$.

Since the intrinsic load on some learning material measures perceived difficulty, it should vary among the participants with their prior knowledge. It would thus be a test for the validity of the questionnaire to test for a correlation between the intrinsic load measured by the cognitive load questionnaire and the test score measured by the prior knowledge questionnaire. The correlation should be negative, as a low intrinsic load indicates perceived ease, which should be correlated with high prior

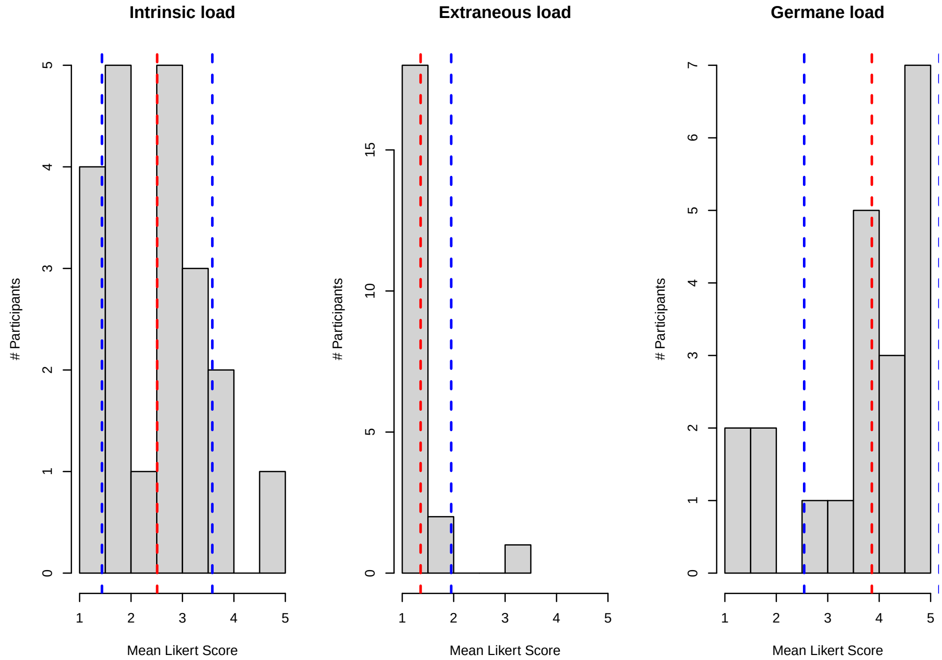


FIGURE 6.4: Histograms of the three dimensions of the Cognitive load questionnaire. The red vertical line is the mean of the scale, while the blue lines depict one standard deviation around the mean.

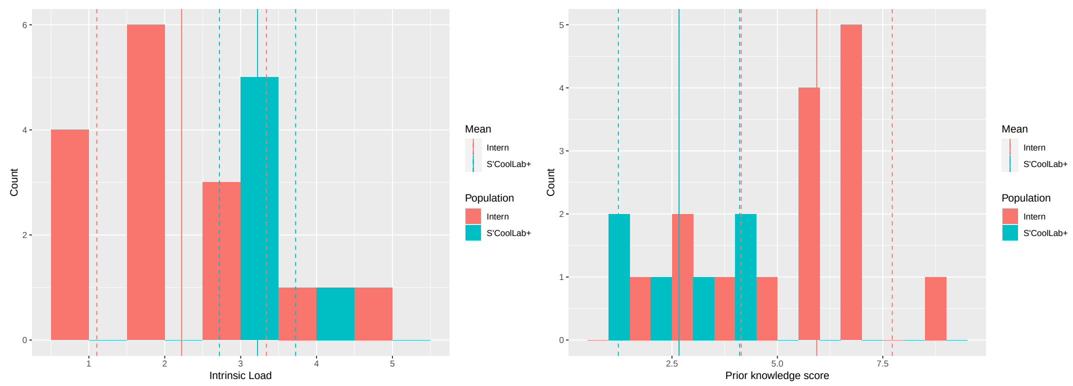


FIGURE 6.5: The distributions of the intrinsic load (left) and the prior knowledge score (right) measured within the two populations taking part in the second student study. The red distributions show the distributions of the students who were at CERN as part of an internship program, while the turquoise distributions show the distribution of students who were at CERN as part of a one-day visit. The solid lines denote the means of the respective distributions, while the dashed lines denote the standard deviations.

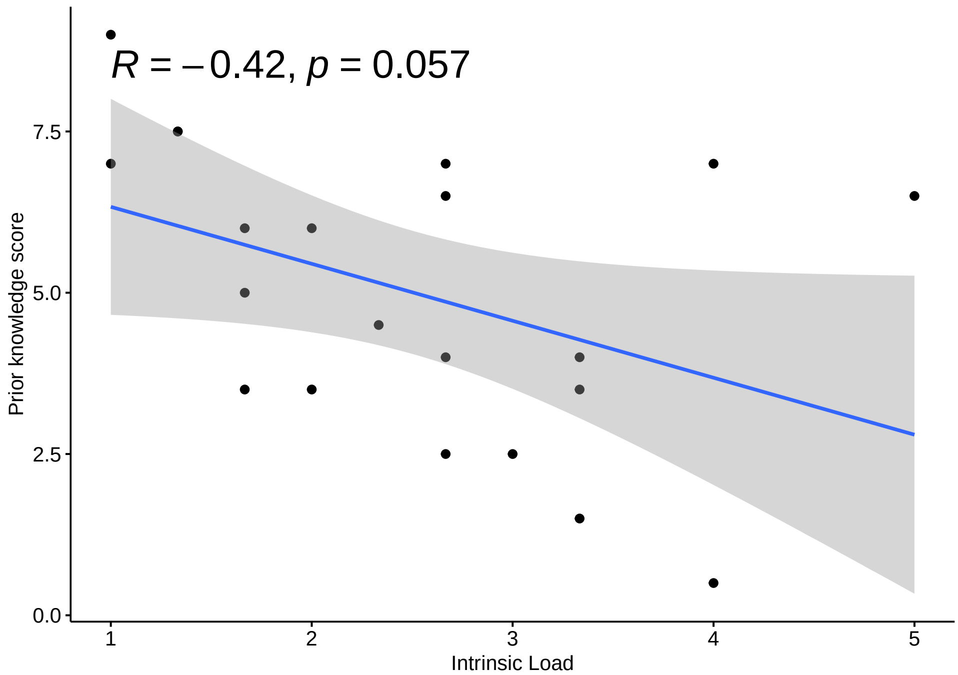


FIGURE 6.6: Scatter plot intrinsic load against the prior knowledge test score. The correlation is not significant as denoted by p.

knowledge. This correlation is shown in Figure 6.6. It is visible that the correlation between the two quantities points in the right direction, even though it is not significant. Nevertheless, it is a promising sign that the questionnaire measures knowledge relevant to the activity.

Eye Tracking data

Data about eye movements can give us insights into the students' thought processes. However, they are most valuable when combined with verbal data from students. Therefore, to triangulate the eye tracking data with the verbal data, the students were split along their reasoning as described above into a "decay vertex reasoning group", a "production vertex reasoning group", and a "mixed reasoning group". We then calculated the time the students spent on the particles belonging to the two vertices (production vertex, meaning the muon on the left and the neutrino on the top right, and decay vertex, meaning the anti-neutrino and the electron on the bottom left), normalized by the whole time they spent on the diagram. Figure 6.7 shows two so-called heat maps that visualise the students' attention distribution on the first question of the second question type (as shown on the bottom of Figure 6.1). The heat maps show that the students' reasoning is reflected in how they examine the diagram: The attention distribution for the production vertex reasoning group is mostly limited to the production vertex. In contrast, the attention distribution of the decay vertex reasoning group is focused on the decay vertex but also more spread out. I performed a one-way ANOVA to see if there are differences between the three reasoning groups in the relative time spent on

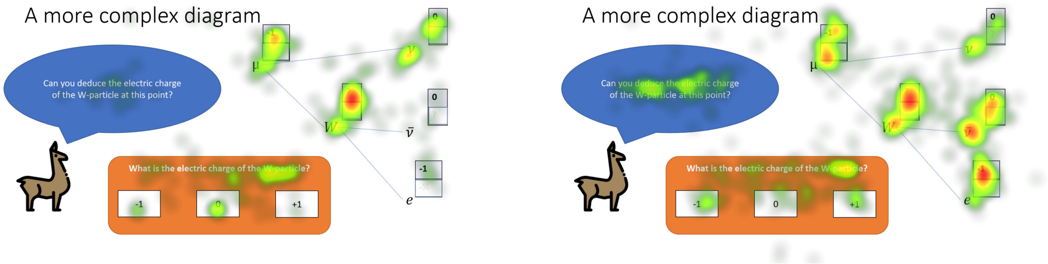


FIGURE 6.7: Heat maps of the attention distribution spent on parts of the question, which is pictured at the bottom of Figure 6.1. The left heat map shows the attention distribution of students who reasoned their answer using the production vertex, while the right heat map shows that of students who exclusively used the decay vertex. Note that the attention distribution of the production vertex reasoning group is almost entirely focused on that vertex. In contrast, the attention distribution of the decay vertex reasoning group is more spread out. See the text for details.

Reasoning	Number	$t_{\text{pro}}/t_{\text{tot}}$	SD	$t_{\text{dec}}/t_{\text{tot}}$	SD
Production	6	0.43	0.10	0.04	0.04
Decay	6	0.22	0.12	0.35	0.11
Mixed	8	0.29	0.06	0.2	0.06

TABLE 6.5: Relative times the three reasoning groups spent on the production vertex ($t_{\text{pro}}/t_{\text{tot}}$) and the decay vertex ($t_{\text{dec}}/t_{\text{tot}}$).

the production and the decay vertex. This ratio's average and standard deviation can be found in Table 6.5. For both vertices, the relative time spent was significantly different (production vertex: $F(2,17) = 7.0, p = .006, \eta^2 = 0.45$, decay vertex: $F(2,17) = 21.6, p = 2 \times 10^{-5}, \eta^2 = 0.72$). A posthoc analysis revealed that for the decay vertex, all group differences were significant, with the difference between decay and production vertex reasoning group being the most pronounced (difference of 0.32 with a 95 % confidence interval between 0.19 and 0.44, $p = 10^{-5}$). In contrast, the difference between the decay vertex reasoning group and the mixed reasoning group was not significant for the production vertex. Still, the other differences were, again, with the difference between the production vertex reasoning group and decay vertex reasoning group the most pronounced, while being less pronounced than in the decay vertex (difference of 0.21 with 95 % confidence interval between 0.06 and 0.36, $p = .005$).

6.5.3 Discussion

The eye-tracking data suggest that the students' reasoning when examining the diagrams is reflected in their visual attention. We can see from the heat maps in Figure 6.7 and the data described in Section 6.5.2 that, when students examine the diagram in a more complete way, i.e. distribute their attention across the whole diagram, they favour one reasoning strategy, namely the "decay vertex reasoning", above the other. This is a strong suggestion as to which strategy is perceived easier by the students. A possible reason for this is the overall structure of the diagram: The decay vertex looks similar to the vertices the students have seen earlier in the learning material as part of other diagrams. The production vertex, however, might not be perceived as such, possibly because the neutrino on the top right is on the same level as the final particles of the decay vertex. This could lead to confusion, so the “decay vertex strategy” might be favoured by those who thoroughly examine the diagram. The implications of this result are further discussed in Section 7.1.2.

One limitation of the study was the lack of a classic pre-post design, making it challenging to check for the achievement of learning goals systematically. Also, the current study showed a ceiling effect in the answers to the questions in the learning environment, i.e., not enough students answered incorrectly, so no inferences could be made about differences between those who answered correctly and those who answered incorrectly. A follow-up study with an updated version of the learning material, using an updated version of a test instrument for particle physics and a pre-post design, addresses these limitations. Also, to avoid ceiling effects, the examples to learn from are modelled on diagrams different from those used as test questions.

This study gave insights into students’ perceptions of the *decay-type diagram*. The students’ verbal data showed that they used different strategies when examining the diagram, which were reflected in their gaze patterns. Regardless of their strategy, most students could answer the questions. However, it was visible from the eye-tracking data that some students did not examine the diagram entirely, and mainly, these students chose the slightly more complicated strategy. This suggests that students need support to examine the decay-type diagram in the sense that the different vertices need to be highlighted, for example, by using visual cues or guiding the explanatory texts. These suggestions will be further discussed in Section 7.2.

The full version of the learning material, the answers to the questionnaires, eye tracking data as well as analysis scripts are publicly available. The links can be found in appendix D.

6.6 Third student study

The third study used a wholly refurbished version of the learning material and incorporated multiple-choice tasks with parallel tasks of different diagrams. It also differed from the first two because it didn’t use a think-aloud protocol, as the design foresaw doing the study with multiple students in parallel.

This study took the previous study’s findings into account in the way that it used examples from the learning materials and compared the decay-type diagram to other diagrams with respect to the first learning goal. It also went further since it systematically tried to test the second learning goal with a questionnaire.

6.6.1 Research Questions

This study aimed at both research goals stated in Chapter 3. It consisted of two parts with two independent research questions

1. Which cognitive processes are triggered when learning with Feynman diagrams about charge conservation?
2. Which conceptual challenges do students exhibit after engaging with instruction about Feynman diagrams about the interactions of particles?

The first part of this study aimed to study students' information processing of Feynman diagrams. In particular, it was to find out which processes students predominantly use when solving specific tasks regarding charge conservation. The results of that part should help to gauge the effectiveness and inform the further design of the learning material with Feynman diagrams.

The second part aimed to create a catalogue of students' conceptions of Feynman diagrams and interaction particles to inform future learning material design using Feynman diagrams.

Together, these two research questions contribute directly to the second research question of the first research goal by identifying conceptual difficulties of Feynman diagrams and indirectly to the second research question by inferring suggestions for teaching with Feynman diagrams.

6.6.2 Description of the study

This study aimed to investigate the impact of learning with Feynman diagrams on students' understanding of charge conservation and to identify the conceptual challenges students face after receiving instruction about the interactions of particles. The study involved 27 students from eleventh and 12th grade from a German high school in Göttingen, with an additional eight students from 10th grade from a German high school in Geneva. The participants' ages ranged from 15 to 19 years, with a gender distribution of 22 males and 11 females. The data collection involved the following steps:

- Pre-Questionnaire: Participants completed a pre-questionnaire assessing their prior knowledge in particle physics before instruction. This instrument is further described in Section 5.3.1 and presented in the appendix table B.2. Its purpose was to assess the level of students, not necessarily to test for the achievement of the learning goals. There were learning environment questions and comprehension questions after the learning environment for the latter. Therefore, the prior knowledge questionnaire was only used in the beginning.

- Learning Material Presentation: Participants worked through learning material about Feynman diagrams on a computer screen while their eye movements were recorded. The material was overhauled entirely based on the material used in the previous studies. The study consisted of two groups: one with a pre-post and another with a post-only design. This splitting of groups is further described in Section 6.6.4 and illustrated in Figure 6.11.
- Post-Questionnaire: Following the instructional session, participants completed a post-questionnaire that included comprehension questions related to the material they had learned.
- Cognitive Load Questionnaire: The study also utilized the Cognitive Load Questionnaire, described in Section 5.3.2 and presented in the appendix table B.4, to assess the cognitive load experienced by the participants during the instructional session.
- Instructional material motivation survey: Motivation levels were assessed using the instructional material motivation survey, described in Section 5.3.3 and presented in the appendix table B.5.

6.6.3 Material description

The students got the following text as an introduction to charge conservation[6]:

> *Alle Teilchen haben bestimmte unveränderliche Eigenschaften, die sogenannten Ladungen. In diesem Kurs schauen wir uns zwei verschiedene davon an, nämlich die elektrische und die schwache Ladung. Es ist ein grundlegendes Gesetz der Physik, dass die gesamte Ladung eines Systems, genau wie dessen Energie, sich über die Zeit nicht ändert. Man sagt dazu auch: Die Ladungen sind erhalten. In jedem Vertex eines Diagramms sind sowohl die elektrische als auch die schwache Ladung unabhängig voneinander erhalten. Das heißt, dass sowohl die Summe der schwachen Ladung als auch die Summe der elektrischen Ladung auf der linken Seite genau die gleiche ist wie auf der rechten Seite.*

The English translation of this text is as follows:

> All particles have specific invariable properties called charges. In this course, we examine two different ones: the electric charge and the weak charge. It is a fundamental law of physics that a system's total charge, like its energy, does not change over time. This is also called "the charges are conserved." In each vertex of a diagram, the electric charge and the weak charge are conserved independently of each other. This means that the sum of the weak charge and the sum of the electric charge on the left-hand side is exactly the same as on the right-hand side.

[6]In this material, the focus is on the weak and the electromagnetic interaction, while the strong interaction is only mentioned as a third interaction governing the world of particles. A discussion about this choice is given in Section 5.1.2

The students were given examples of valid and invalid Feynman diagrams (see Figure E.1 in the appendix) to showcase how the charges are displayed in the diagram and how to distinguish diagrams with and without conserved charges.

The students got seven different tasks in which they had to apply their knowledge about charge conservation. The naming of the tasks was done according to the following scheme:

- *Completion*: Choose the correct particle/vertex among three options to complete an unfinished diagram (called C1, C2, C3), with C1 given in Figure 6.8
- *Vertex*: Choose the wrong Vertex among three options (called V1, V2), with V2 given in Figure 6.9.
- *Diagram*: Choose the wrong diagram among three options (called D1, D2), with both tasks given in Figure 6.10

An more detailed overview of the tasks is given in Table 6.6. The tasks fall into two different categories of cognitive processes according to Bloom's revised taxonomy (Anderson & Krathwohl, 2001; Mayer, 2002b) which are *attributing* as part of the *analyzing* category (Completion category) and *checking* as part of the *evaluating* category (Vertex and Diagram categories). All tasks are given in the appendix in Figure E.2.

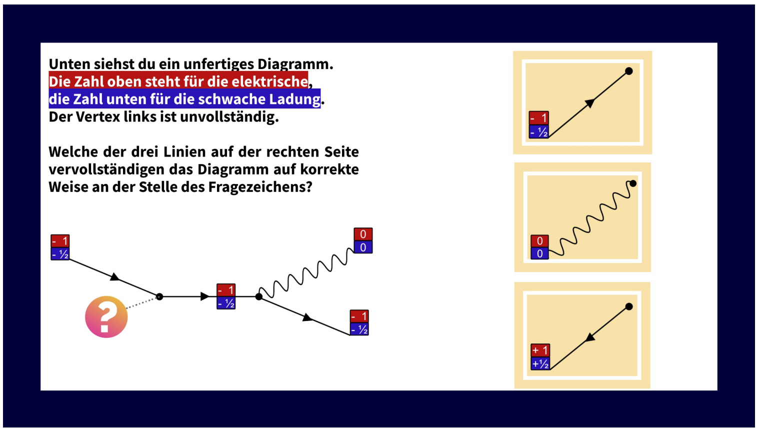


FIGURE 6.8: First task of the Completion category (C1). The text reads (in German): *Below is an unfinished diagram. The number on top signifies the electric, and the bottom signifies the weak charge. Which of the three lines on the right side completes the diagram correctly at the place of the question mark?*. The correct answer is the second answer option.

The second part of the material was interactively designed, meaning that students could jump back and forth in the material and choose their preferred examples for explanations. The explanation steps are as follows. Screenshots of the learning material in which these steps are realised are given in the appendix in Figure E.3:

- Brief explanation of the different parts of a diagram: Students see a diagram with annihilation and a pair production without annotations as one example of

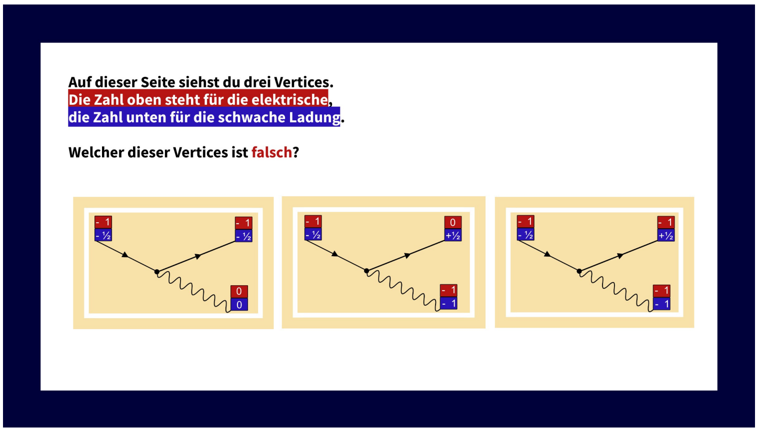


FIGURE 6.9: Second task of the Vertex category (V2). The text reads (in German): *On this page you see three vertices. The number on top signifies the electric, and the bottom signifies the weak charge. Which one of these vertices is wrong?*. The correct answer is the third (due to non-conservation of electric charge) answer option.

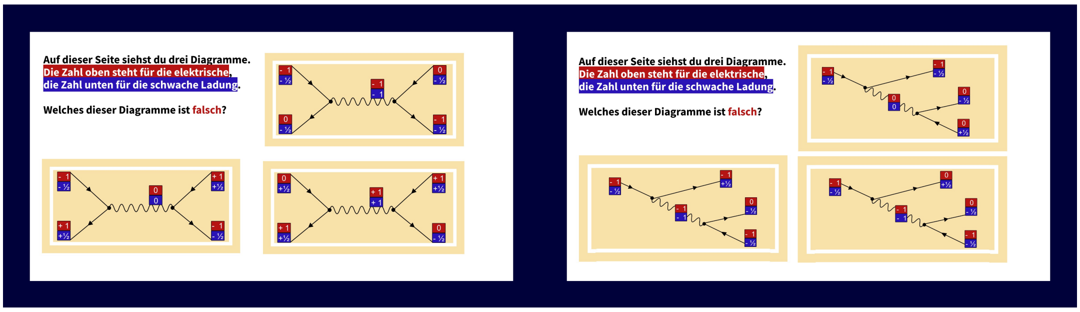


FIGURE 6.10: Tasks of the stimuli of the Diagram category: D1 (left), D2 (right). The text reads (in German): *On this page you see three diagrams. The number on top signifies the electric, and the bottom signifies the weak charge. Which one of these diagrams is wrong?*. The correct answers are the bottom right (D1) and the bottom left (D2) answer option.

a Feynman diagram. They can click on any of the straight lines with an arrow to the right ("matter particles"), any of the straight lines with an arrow to the left ("anti-matter particles"), and the wavy line ("interaction particles") and get a brief explanation in a pop-up window.

- An example of a process governed by the weak and the electromagnetic interaction, including its explanation. The examples are the muon transformation for the weak and electron-electron scattering for the electromagnetic interaction. The examples are chosen for their accessibility and relation to real-world examples (muon transformation: cosmic particles; electron-electron scattering: physical touch[7]).

- Four more examples with real-world connections for particle processes (fluorescence, PET scanning, cosmic muon creation, and hydrogen fusion in the

[7]Though it is not the main effect of how electrons influence each other while touching, it is an example widely used in popular depictions of science (e.g., Veritasium (2013))

	Completion	Mistake-finding
Vertex	C1 can be counted in this category, even though a whole diagram is depicted (see Figure 6.8).	V1 and V2 show different types of vertices. While V1 shows a symmetric vertex, where two particles transform into one, V2 shows a slightly asymmetric one where it is potentially not directly visible which particles transform into each other (see Figure 6.9).
Diagram - Scattering-type	C2 shows a diagram where the intermediate particle has to be found to complete two vertices of a scattering-type diagram.	D1 shows three scattering-type diagrams where the charge conservation is not correct in one of them (see Figure 6.10, left).
Diagram - Decay-type	C3 shows a decay-type diagram where the lower "decay vertex" has to be completed.	D2 shows three decay-type diagrams where the one with incorrect charges has to be found (see Figure 6.10, right).

TABLE 6.6: Overview of task types in the third student study. The row/column combination shows the respective diagram/task combination of the different tasks.

solar interior). Students can choose which and how many of the examples they look at. For each example, they get a textual explanation of the process and are asked to choose the Feynman diagram that describes this process out of four diagrams. When they choose correctly, they are provided with a more detailed explanation of the process and are further asked which of the two aforementioned fundamental interactions is involved.

- At the end of the chapter, the students are provided with a summary of the key elements of the explanations.

6.6.4 Data collection

Students were seated before a screen while their gaze was recorded with an eye tracker (see Section 5.2.1). For technical reasons, eye-tracking data could not be obtained from all participants. From the 35 students who participated in the study, usable eye-tracking data could be obtained from 28. After all the students had seen an introduction slide to Feynman diagrams and the introduction text about charge conservation, one group (pre-post group) was immediately asked to solve the seven tasks before being shown the above examples (pretest). The students answered the question by clicking on the answer using the computer mouse. They could not go

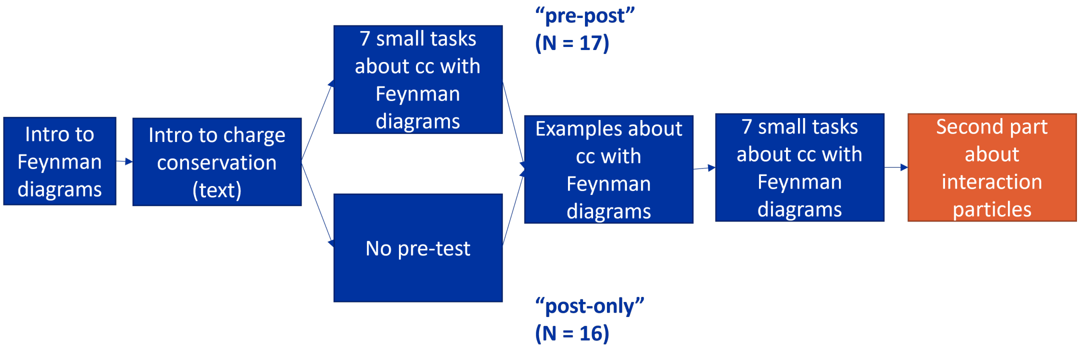


FIGURE 6.11: Scheme of the procedure of the first part of the third student study as described in Section 6.6.2. The students have been divided into two groups.

back to a previous slide during this part. The other students (post-only group) were shown the examples first. After the students were shown the examples, they all solved the same seven tasks the first group had already solved (posttest). The procedure is demonstrated in Figure 6.11. This splitting of the groups was done to investigate the effects of the examples shown. However, it was not a simple pre-post design to control for the effects of repeated testing. The latter is called the "testing effect" and is even prevalent when not giving feedback (for a review of this effect, see Roediger and Butler (2011)). Therefore, three comparisons between answers in group/test combinations are performed: One between the answers of the pre-post group in the pretest and their answers in the posttest to investigate effects of the examples and testing effect combined, one between the answers of the pre-post group in the pretest and the answers of the post-group to investigate the effects of the examples alone, and one between the answers of the pre-post group in the posttest and the post-only group, to explore the testing effect.

After the students went through the charge conservation tasks, they were shown an interactive learning environment where they could go back and forth at their own pace. This learning environment covered the topic of interaction particles as described in section 6.2.2. The learning environment consisted of explanations in text form, graphics in the form of Feynman diagrams and illustrative pictures. After they went through the learning material they answered the post-questionnaire and the two surveys described above. The post-questionnaire contained five prompts, two of them being comprehension questions. The prompts are given in Table 6.7 in the original German version and their English translation. The questions from the surveys are given in the appendix tables B.4 and B.5. The free-text questions were answered in written form using a computer keyboard. While all 35 students filled out the prior knowledge questionnaire in the beginning and the comprehension questions, the cognitive load questionnaire and the instructional motivational survey were only completed by 29 students (only one in Geneva and all 28 in Göttingen). The time it took students to complete the different parts of the learning

Prompt (original phrasing)	**English translation**
Bitte schreibe ganz kurz (maximal zwei Sätze), was du in diesem Lernmaterial gelernt hast.	Please write briefly (2 sentences maximum) what you have learned in this learning material.
Bitte schreibe, welche Fragen du noch zum Inhalt des Lernmaterial hast. Das können weiterführende Fragen sein oder solche, die sich aufgrund von unklaren Erklärungen ergeben.	Please write which questions you still have with regards to the content of the learning material. These can be further questions or questions due to unclear explanations.
Erkläre bitte in eigenen Worten, wie ein Feynman-Diagramm die Wechselwirkungen zwischen Teilchen beschreibt. Sei so detailliert wie möglich (mindestens 4-5 Sätze).	**Please explain in your own words how a Feynman diagram describes interactions between particles. Be as detailed as possible (at least 4-5 sentences).**
Erläutere in eigenen Worten die Rolle eines Wechselwirkungsteilchens in einem Teilchenprozess.	**Explain in your own words the role of an interaction particle in a particle process.**
Hast du Verbesserungsvorschläge für das Material?	Do you have suggestions for improvement of the material?

TABLE 6.7: Prompts in the post-questionnaire in the original German version and their English translation. The third and fourth prompts (written in bold) are furthermore referred to as comprehension questions.

material is summarised in Table 6.8.

6.6.5 Students' cognitive load and motivation

The instructional motivation survey, presented in Section 5.3.3 in Table B.5, consisted of four scales – attention, relevance, confidence, and satisfaction. Table 6.9 summarizes each scale's means, standard errors and Cronbach's α as well as the average of the scales, denoted as motivation. The table shows that the internal consistency lies within an acceptable range for three of the four dimensions, while the relevance scale lacks internal consistency. The results for this scale, therefore, have to be interpreted cautiously. The distributions of the scores can be seen in Figure 6.12.

Part	M	SD
Introduction and charge conservation examples	5:03	1:15
Charge conservation questions in pretest (Pre-Post)	5:14	1:22
Charge conservation questions in posttest (Pre-Post)	4:09	1:30
Charge conservation questions in posttest (Post-only)	5:01	1:38
Interaction particles	14:25	5:26

TABLE 6.8: Time in minutes and seconds spent on different parts of the learning material. For the quiz questions, the times are distinguished by the three group/test combinations.

Scale	Mean	α
A	3.75 ± 0.15	.84
R	3.69 ± 0.14	.54
C	3.70 ± 0.15	.80
S	4.02 ± 0.15	.83
M	3.80 ± 0.12	.78

TABLE 6.9: Table with a summary for the answers of the instructional material motivational survey, with means, standard errors, and Cronbach's α in the four scales (A)ttenttion, (R)elevance, (C)onfidence, and (S)atisfaction, as well as their combination score for (M)otivation. The scores on each scale reach from 1 to 5.

The means of all the scales are at the upper end of the scale. Especially the satisfaction scale has a very high value, with very few students answering the questions belonging to this scale in the lower half of the Likert scale (cf. Figure 6.12, third panel).

The cognitive load questionnaire, described in Section 5.3.2 alongside the questionnaires used in the first and second studies consisted of three scales: intrinsic, extraneous, and germane load. The three scales' means and α values are summarised in Table 6.10 while the means distributions are shown in Figure 6.13. While the scales for intrinsic and germane load have high internal consistency, the extraneous load scale has a similar problem as in the previous two studies, even though the questionnaire was adapted majorly; for example, it consisted of five instead of three items and the items that mainly caused the low consistency in the previous versions of the questionnaire, were removed or adapted. In this case, it cannot be traced to one item; therefore, the scale remains unchanged, but it must be examined cautiously.

Scale	Mean	α
IL	3.00 ± 0.16	.80
EL	1.94 ± 0.13	.53
GL	4.01 ± 0.13	.82

TABLE 6.10: Table with the means and standard errors in the four scales of the Cognitive Load questionnaire: intrinsic (IL), extraneous (EL), and germane load (GL). The Likert scores on each scale reach from 1 to 5.

The distributions of the three scales shown in Figure 6.13 show that while the germane load, which is connected to the perceived learning outcomes of the students, was at the high end of the scale, the extraneous load, which is connected to features of the learning environment, was relatively low. The intrinsic load connected to the contents' perceived difficulty was normally distributed around the centre of the Likert scale. The latter result is in contrast to the previous studies when the intrinsic load rather peaked at the lower end of the scale. It must be noted, however, that the student demographics were different. While the students in the first and second student study were mostly interns who came to CERN, the students in this study were upper secondary students in their school.

As described in the section about the CATLM (page 30), germane load (or essential processing) is dependent on motivational factors. Therefore, a correlation between the germane load measured by the cognitive load questionnaire and the different motivational scales measured by the instructional motivational survey should be

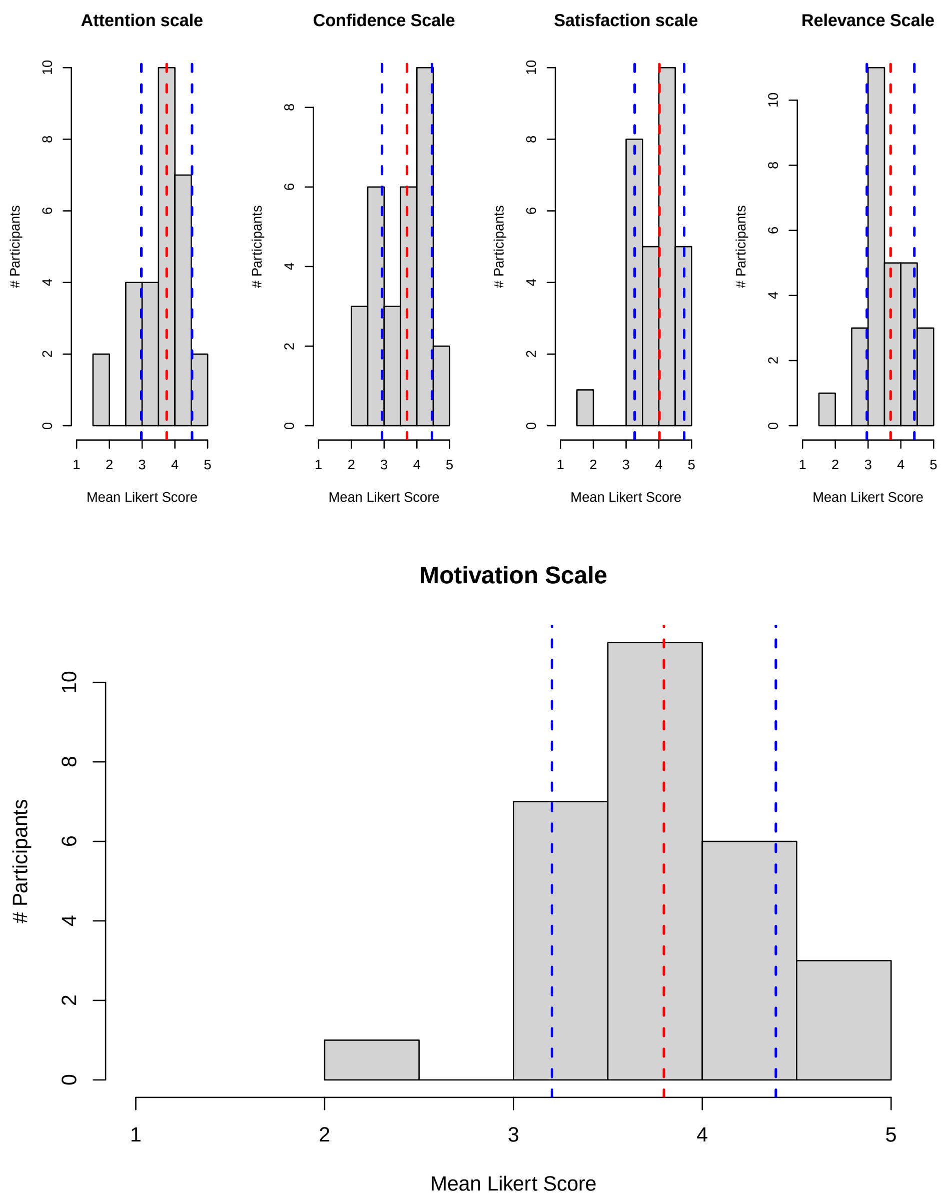


FIGURE 6.12: Histograms of the four dimensions of the instructional motivation survey and the complete scale. The red vertical line is the mean of the scale, while the blue lines depict one standard deviation around the mean. As seen in Table 6.9, the relevance scale has a low internal consistency.

observed. These correlations are shown in Figure 6.14. It is visible that the GL is positively, but very weakly correlated with all four motivation scales, but none of the correlations is significant.

6.6.6 Prior knowledge questionnaire

Before starting the learning material, students filled out the prior knowledge questionnaire described in Section 5.3.1 and presented in the appendix table B.2, which was based on the questionnaire used in the second student study but used single-

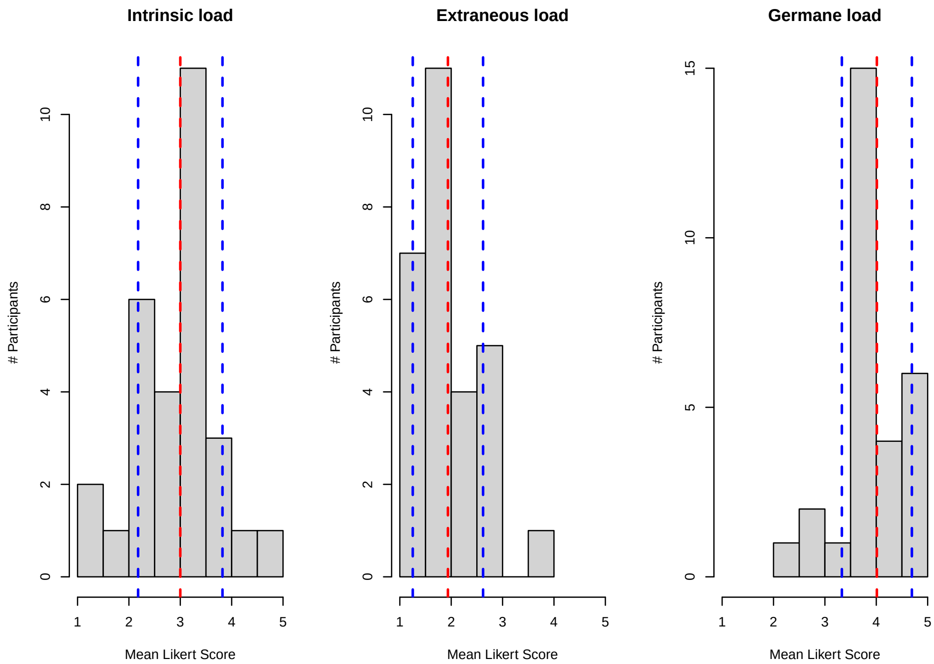


FIGURE 6.13: Histograms of the three dimensions of the Cognitive load questionnaire. The red vertical line is the mean of the scale, while the blue lines depict one standard deviation around the mean. As seen in Table 6.10, the internal consistency of the extraneous load scale is low.

and multiple-choice instead of open questions to make the scoring more objective. The items covered similar topics, but the phrasing of the items was different. The summary of the item analysis of this questionnaire is shown in Table 6.11. This analysis reveals that two of the items were fairly difficult ("Which of the following particles is an interaction particle?" and " Which of the following particles has the smallest mass?") as they had a score p of below 0.3 (Ding & Beichner, 2009). Also, the discriminatory power of the second item ("Which of the statements about energy is true?") was too low, possibly because it was too easy. The point biserial coefficient of all the items was in an acceptable range. However, the test's internal consistency was very low with $\alpha = 0.60$ and KR-20 = 0.36, while the overall discriminatory power was just outside the acceptable range with $\delta = 0.86$.

Like in the previous study, the overall test score is slightly negatively correlated with the intrinsic load measured by the cognitive load instrument. This correlation is shown in Figure 6.15. It is visible that the trend is in the right direction, even though it is far from being significant.

6.6.7 Students' information processing of Feynman diagrams

To analyse the eye-tracking data, so-called "areas of interest" (AOI) were defined on the task stimuli which were then taken as units to count the numbers of fixations,

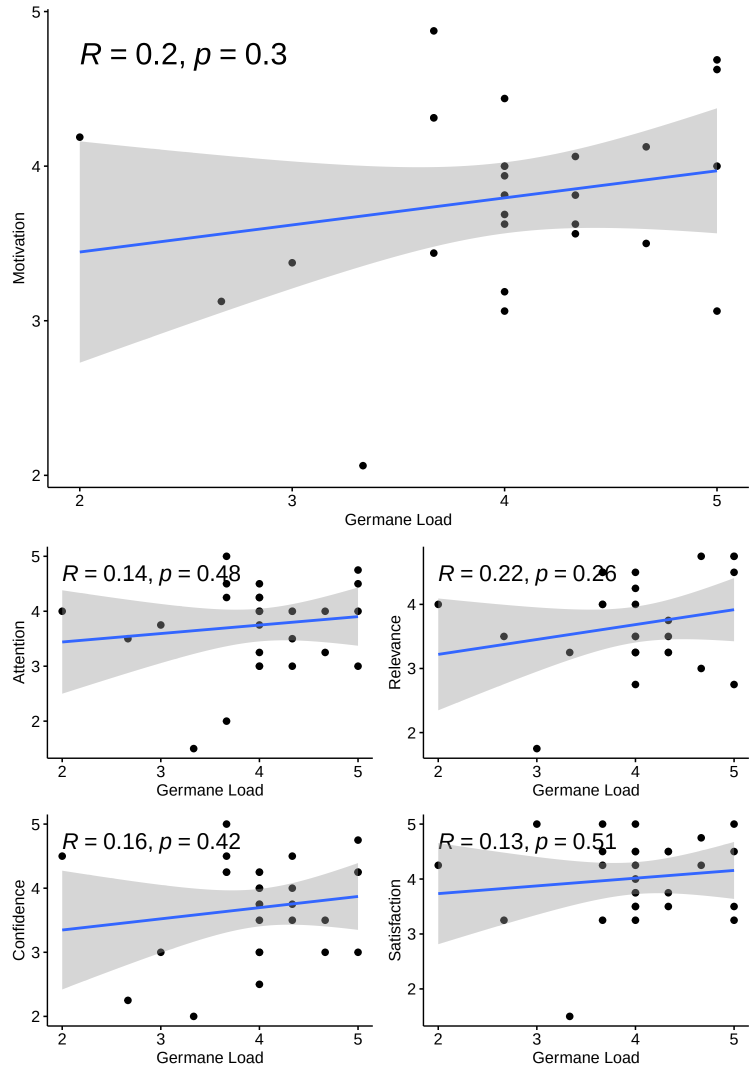


FIGURE 6.14: Scatter plots of the germane load measured by the cognitive load questionnaire against the motivational scales measured by the instructional motivational survey. At the top, the correlation is shown with the overall motivation, measured as the average between all scales, while on the right, the correlations with the four single scales are shown. The p-values reveal that none of the correlations is significant.

dwelling time, and transitions between them. There are two different levels of analysis. On the global level, each large stimulus element, i.e., answer options and text elements, defines an AOI, as exemplified on the left side of Figure 6.16 for the second diagram task. Each charge annotation defines an AOI on the local level, as shown on the right side of the same Figure for the same stimulus. The AOIs for the other tasks were defined accordingly.

As described in Section 4.4.3, different cognitive processes are connected to different eye-tracking measures. The following hypotheses describe how the measures are connected to learners' cognitive processing. The first three of the following hypotheses refer to fixation metrics, of which the calculation is described in Section

Item	Correct	Partially correct	Incorrect	p	D	r_{pbi}
1	7	28	0	0.6	0.37	0.72
2	30	0	5	0.86	0.22*	0.44
3	6	0	29	0.17*	0.56	0.66
4	7	0	28	0.2*	0.44	0.54
5	23	0	23	0.66	1	0.73
Total	73	28	85	α	KR-20	δ
				0.6*	0.36*	0.86*

TABLE 6.11: Psychometric analysis of the items of the prior knowledge questionnaire in the order presented in Table B.2. Students could give partially correct answers only in the first item. p denotes the average score obtained on an item, D denotes the discriminatory index and r_{pbi} denotes the point biserial coefficient. α denotes Cronbach's α, KR-20 denotes the Kuder-Richardson index, and δ denotes Ferguson's δ. Values marked with an asterisk are outside the range proposed by Ding and Beichner (2009) as acceptable.

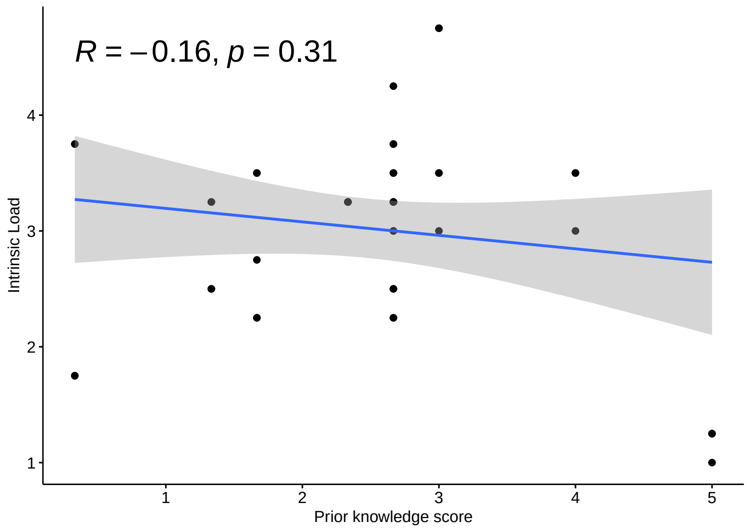


FIGURE 6.15: Scatter plot of prior knowledge against the intrinsic load. The correlation is not significant as denoted by p.

5.2.6, the last four refer to transition metrics, which are described in Section 5.2.7, while the last of them describes representation-specific behaviour:

- Perceptual processing of the different options is connected to the respective dwelling times, therefore the first hypothesis refers to the total dwelling time on correct and incorrect options. More perceptual processing of the correct answer is hypothesized to be related to higher answer probability, as found in most studies examined by Hahn and Klein (2022).

  H1 Students' time on correct options is correlated with their answer correctness.

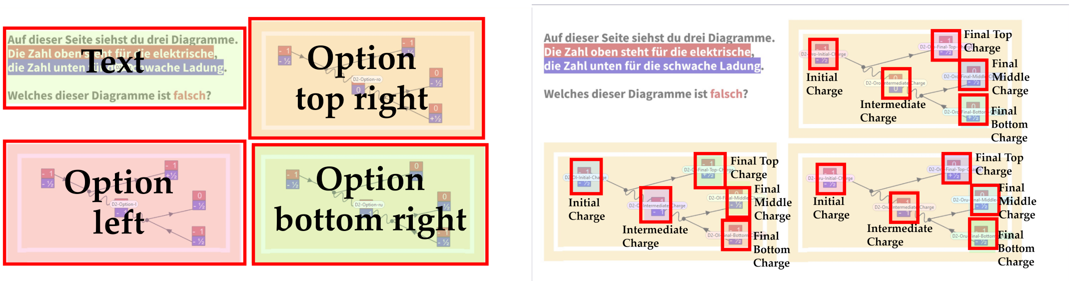


FIGURE 6.16: Definition of AOIs on the global (left) and local (right) level on the example of task D2 (cf. Figure 6.10, right).

- The second hypothesis contrasts the perceptual processing of elements relevant to solving the problem in question with those less relevant. The pilot study (Section 6.3) already showed that novices spent more time on less relevant elements of the diagrams than experts. The examples should have helped to show that the charge annotations are relevant to solving the problems. Therefore, the students are hypothesized to spend more time on the charge annotations of the diagrams in the posttest than in the pretest. Due to the testing effect, this difference might also be prevalent between the first and second views of the task regardless of seeing the examples.

  H2 The ratio between the time spent on the charge annotations and the time spent on the whole stimulus is higher in the posttest and the second view of the task than in the pretest and the first view.

- The next hypothesis refers to the perceptual processing of the significant charge annotations that needed to be spotted in the "vertex" and "diagram" tasks (for example, in D2, the three charges on the left of the left option do not add up). It is hypothesized that perceptual processing of these stimulus regions is related to a higher probability of correctly answering these tasks.

  H3 The time spent on the significant charges determining the correct option is related to the answer correctness on a task.

- The next hypotheses refer to the integrating processes of students, which are connected to transition measures (cf. Alemdag and Cagiltay (2018)). These processes are related to transitions between different elements. As introduced in Section 4.4.3, the transitions are quantified in two ways. One way is to compute the fixation/transition ratio (FTR), which distinguishes focused (higher FTR) from comparative (lower FTR) behaviour. The other way is the transition entropy, which distinguishes exploratory (higher entropy) from focused (lower entropy) behaviour. Both can be calculated on the local and global scales. Since both measures are relatively new and haven't been used widely in physics education research (cf. Hahn and Klein (2022)), there is no clear hypothesis as to what is to be expected. For the FTR, Rodemer et al. (2020) found a high FTR, i.e., focused behaviour is more prevalent in beginners when confronted with a task that requires heavy comparison between AOIs. In contrast, Holmqvist et al. (2011) found a focused behaviour to be more prevalent in high-ability than low-ability students, who rather show an "overview behaviour" when confronted with a mathematical task. Therefore, a preliminary hypothesis would be that focused behaviour leads to more successful task-solving on a global scale. In contrast, a comparative behaviour might be beneficial on the local scale. For the transition entropy, Mozaffari Chanijani et al. (2016) found a higher entropy to be more prevalent in novices, which is why higher-ability students are hypothesized to have a lower transition entropy. Also, the transition entropy is hypothesized to become lower over time, as found by Jordan and Slater (2009) in a virtual environment. However, the hypotheses about the transition entropy are only valid for the global scale, while the local scale seems more exploratory.

H4a The FTR should be higher on the global scale among successful students since students need to focus on single options to find the solution.

H4b The FTR should be lower on the local scale among successful students since students need to compare different charge annotations to find the solution.

H5 The transition entropy on the global scale should be lower in the posttest and the second view of the task because the students have learned how to examine the diagram.

H6 The transition entropy on the global scale should be lower for successful students as those examine systematic viewing behaviours.

- The final hypothesis refers to representation-specific gaze patterns. Since for a successful completion of the tasks, different charges have to be compared with each other systematically, it is hypothesized that productive viewing behaviour is reflected in a high triple transition ratio (see page 52).

H7 The triple transition ratio should be higher for successful students than for unsuccessful students.

Item analysis of the tasks

The task answer behaviour in the first part of the learning environment is analysed using classical test theory using the indicators introduced in Section 5.4.1. The results of this analysis are shown in Table 6.13. In the table, an asterisk marks all the values outside the acceptable range according to Ding and Beichner (2009). However, according to Adams and Wieman (2011), it could be desirable for some items to lie outside of an acceptable range in some cases. All but the first two completion tasks showed suitable variability and discriminatory power when the answers were taken together. Only Cronbach's α was a bit off. However, as discussed by

Adams and Wieman (2011), a high α would denote many redundant items. As the test was not designed to measure a single construct, a high internal consistency is not expected.

For the difference between the three task categories, a paired t-test revealed a significant difference between the answers to the completion tasks ($M = 0.87, SD = 0.23$) and the vertex tasks ($M = 0.68, SD = 0.33$), $t(48) = 3.3 p = .002$ with a small effect size of $d = 0.47$ as well as between the completion and the diagram tasks ($M = 0.60, SD = 0.33$), $t(48) = 4.2, p = .0001$ with a moderate effect size of $d = 0.6$. However, there were no significant differences between the answers in the vertex and the diagram task ($t(48) = 1.5, p = .15$).

For the individual tasks in the vertex and diagram category, the test revealed no significant difference in the vertex category ($M = 0.75, SD = 0.43$ and $M = 0.61, SD = 0.49$ for V1 and V2, respectively, $t(48) = 1.5, p = .13$). In contrast, there was a significant difference between the diagram tasks ($M = 0.77, SD = 0.42$ and $M = 0.44, SD = 0.50$ for D1 and D2, respectively, $t(47) = 3.5, p = .001$ with a small effect size of $d = 0.48$. The difference between the most difficult task, D2, and the second-most difficult task, V2, was also insignificant ($t(47) = 1.8, p = .07$). In summary, the completion tasks were significantly easier than the other two task categories, while the second diagram task was significantly more difficult than the first "diagram" task.

Group	Test	N	Mean	SD
Pre-Post	Pretest	17	4.4	1.6
Pre-Post	Posttest	17	5.3	1.3
Post-only	Posttest	16	5.5	1.3
Total	Posttest	33	5.4	1.3
Total		50	5.1	1.5

TABLE 6.12: Summary of the test scores of the different groups

In the overall performance, which is summarised in Table 6.12, a t-test revealed a significant difference between the scores of the pre-post group in the pretest and the post-only group ($t(29.2) = 2.2, p = .036$ with a moderate effect size of $d = 0.77$). However, while there was no significant difference between the two groups in the posttest, a paired-sample t-test revealed a slightly significant difference between the scores of the pre-post group in the pretest and the posttest ($t(16) = 2.2, p = .043$ with a moderate effect size of $d = 0.53$). The means and standard deviations of the group/test combinations are shown in Table 6.12.

To check the consistency between the prior knowledge questionnaire and the questions in the learning environment, Figure 6.17 shows the correlation between the scores the students obtained in these two instruments. As the diagram shows, the correlation is significant with a high prior knowledge score fairly well correlated with a high test score. In all scatter plots where test scores are plotted against student scores, the test scores in the pretest are considered for the students in the pre-post group.

Task	Prepost-Pretest			Prepost-Posttest			Post-Posttest			Total		
	p	D	r_{pbi}	p	D	r_{pbi}	p	D	r_{pbi}	p	D	r_{pbi}
C1	0.93*	0.25*	0.60	1*	0*	0*	1*	0*	0*	0.98*	0.07*	0.43
C2	0.88	0.5	0.62	0.94*	0.2*	0.65	1*	0*	0*	0.94*	0.2*	0.54
C3	0.59	0.83	0.69	0.76	0.5	0.56	0.71	0.5	0.36	0.69	0.54	0.56
V1	0.59	1	0.85	0.76	0.3	0.46	0.93*	0.33	0.33	0.76	0.58	0.64
V2	0.53	0.07*	-0.05*	0.65	1	0.85	0.67	1	0.75	0.61	0.60	0.53
D1	0.59	0.27*	0.38	0.82	0.6	0.71	0.93*	0*	0.16*	0.77	0.45	0.65
D2	0.47	0.43	0.44	0.38	0.04*	0.05*	0.47	0.67	0.67	0.44	0.37	0.41
Total	α	KR-20	δ	α	KR-20	δ	α	KR-20	δ	α	KR-20	δ
	0.43*	0.72	0.89*	0.43*	0.70	0.79*	0.53*	0.81	0.84*	0.52*	0.77	0.90

TABLE 6.13: Analysis of the questions within the learning environment according to classical test theory. C1, C2, and C3 denote the Completion tasks as exemplified in Figure 6.8, V1 and V2 denote the Vertex tasks as exemplified in Figure 6.9, and D1 and D2 denote the Diagram tasks as presented in Figure 6.10. p denotes the average score obtained on an item, D is the discrimination index, r_{pbi} the point biserial coefficient, α denotes Cronbach's α, KR-20 is the Kuder-Richardson index, and δ denotes Ferguson's δ. The table shows the values for the three group/test combinations (pre-post in the pretest and posttest and post-only in the posttest). The Total column shows the values for all the answers taken together. N was 17 in the case of the pre-post group and 15 in the case of the post-only group. Values outside the acceptable range according to Ding and Beichner (2009) are marked with an asterisk.

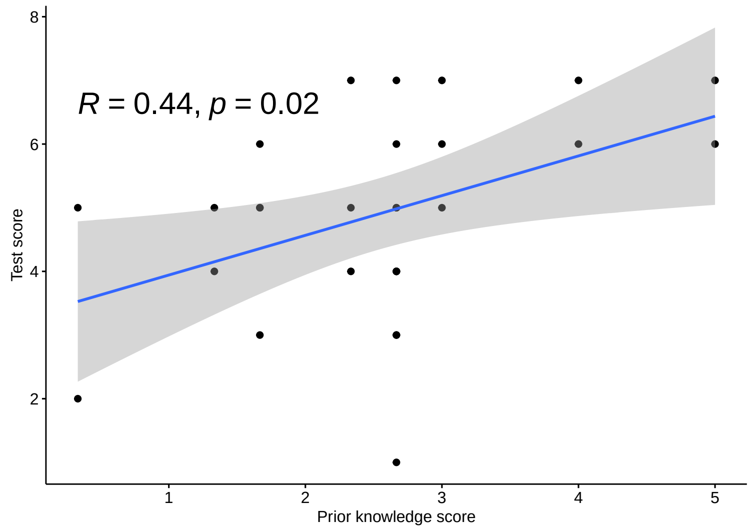


FIGURE 6.17: Scatter plot of prior knowledge against test scores of the students in the learning environment questions.

Attention distribution.

The first hypothesis concerned the ratio between the time spent on the correct option and the whole time spent on the stimulus. To obtain an overview, Figure 6.18 shows the correlation between the average ratio and the test score and the intrinsic load, where the former serves as a performance indicator and the latter as an indicator of perceived difficulty. The diagrams show that the test score significantly correlates with the average relative time spent on the correct options, as denoted by Pearson's correlation coefficient. However, the correlation is not significant with the intrinsic load. A per-stimulus analysis of this quantity is shown in Figure 6.19 (without the first completion task, as there was only one wrong answer among those of whom eye-tracking data could be collected). T-tests yielded significant differences among those who answered correctly and incorrectly in the third completion task, both vertex tasks, and the second diagram task. The results of the t-tests are shown in Table 6.14. In summary, students who answer correctly generally spend more time on the correct option than students who answer incorrectly. Therefore, H1 can be accepted. It is interesting to note that the hypothesis does not hold for the first diagram task. The item analysis showed that this task was significantly easier than the second diagram task, but not significantly easier than the first vertex and the third completion task (cf. Table 6.13). Therefore, the difference in the viewing behaviour can only in part be explained by the difference in difficulty.

Stimulus	t	df	p	d	magnitude
C2	1.1	1.4	0.4		
C3	5.4	16.6	$5 \times 10-^{-5}$	1.96	large
V1	5.7	21.5	1×10^{-5}	1.85	large
V2	2.9	38.2	0.006	0.89	large
D1	0.4	9.4	0.7		
D2	3.8	22.2	0.0009	1.29	large

TABLE 6.14: Results of a Welch-corrected t-test for the average relative time on the correct option in the respective stimulus between those who answered correctly on that stimulus and those who did not across both groups and both tests. t denotes the test statistic, df the number of degrees of freedom, p the p-value, and d Cohen's effect size, which is only calculated for significant differences and used to determine the magnitude of the effect (last column).

The second hypothesis concerned the students' relative time on the charge annotations. Figure 6.20 shows the difference in the average ratio across all stimuli between the three group/test combinations. There is only a slightly significant difference in the paired sample t-test between the pre- and the posttest of the pre-post group ($t(13) = 2.2, p = .049$ with a moderate effect size of $d = 0.58$). However, Figure 6.21, where heatmaps of the first task are shown, indicates that students spread out their gaze more in the pretest (the heatmap at the top of the figure) than they do in the posttest (the heatmaps at the bottom of the figure). A per-item analysis reveals that

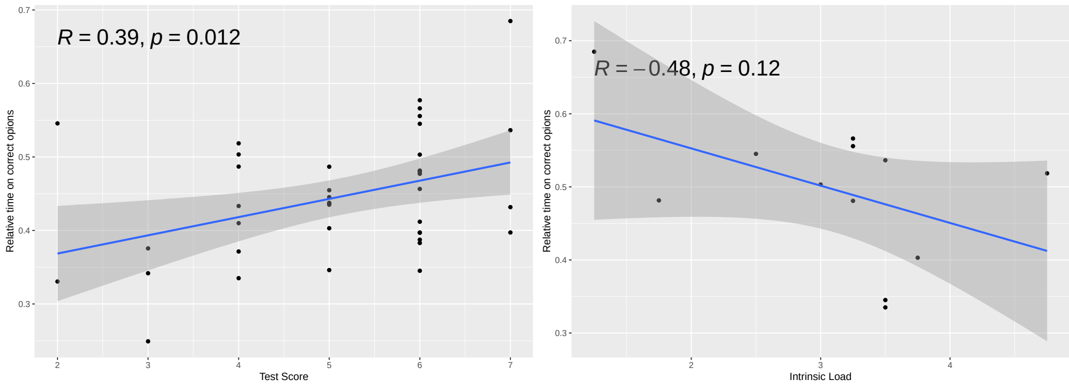


FIGURE 6.18: Scatter plots regarding H1 between the average relative time spent on correct options (y-axis) and the overall score (top) and the intrinsic load (bottom).

there is indeed a significant difference for the first two stimuli. A paired-sample t-test yields $t(13) = 2.8, p = .016$ with a moderate effect size of $d = 0.74$ for stimulus C1 and $t(13) = 5.3, p = .0002$ with a large effect size of $d = 1.41$ for stimulus C2. The effect was not significant when comparing the pre-post group in the pretest with the post-only group ($t(22.6) = 1.8, p = .09$ for stimulus C1 and $t(23.9) = 1.9, p = .08$ for stimulus C2). When comparing the two groups in the posttest, the difference was not significant for stimulus C1 ($t(24.5) = 0.4, p = .7$ but slightly significant for C2 ($t(22.6) = 2.1, p = .046$ with a moderate effect size of $d = 0.79$). Thus, H2 can be tentatively accepted.

The third hypothesis zooms even further into the stimuli. It is concerned with the time students spend on the significant charge annotations which need to be found to determine the correctness of the respective option. Figure 6.23 shows the average relative time spent on these "significant charge annotations" relative to the time spent on the respective option averaged throughout all stimuli in correlation with the overall test score (top) and the intrinsic load (bottom). In this case, there is no significant correlation. In the case of the intrinsic load, there is a trend visible in the direction that those with a higher load, and hence a greater perceived difficulty, spent less time on the significant charge annotations, but this trend is not significant. H3 has to be rejected.

Transition measures.

The fourth hypothesis was split into two parts. The first part concerned the fixation/transition ratio on the global scale, whereas the second part concerned the FTR on the local scale. Figure 6.24 shows the scatter plot of the average FTR throughout all stimuli against the overall test score and the intrinsic load. The figure shows a positive correlation with the test score, which means that students who perform better show a more focused viewing behaviour. This trend is also visible when correlated with the intrinsic load, where a more focused behaviour is correlated with

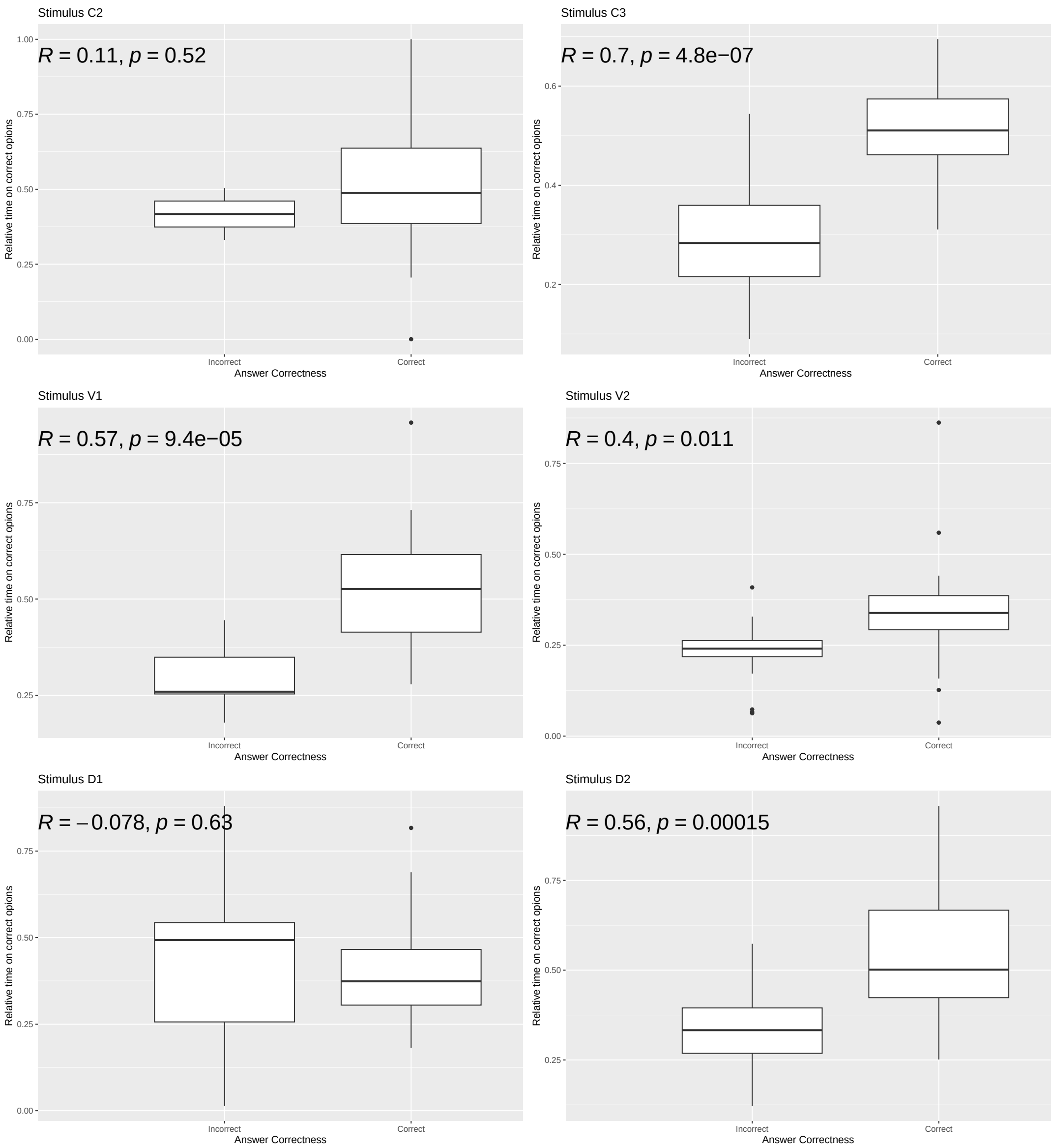


FIGURE 6.19: Boxplot diagrams regarding H1. The y-axis shows the dwelling time on the correct options as a fraction of the total time spent on the stimulus. The incorrect options are on the left, and the correct options are on the right. Task C1 is omitted in this figure since there were too few incorrect answers.

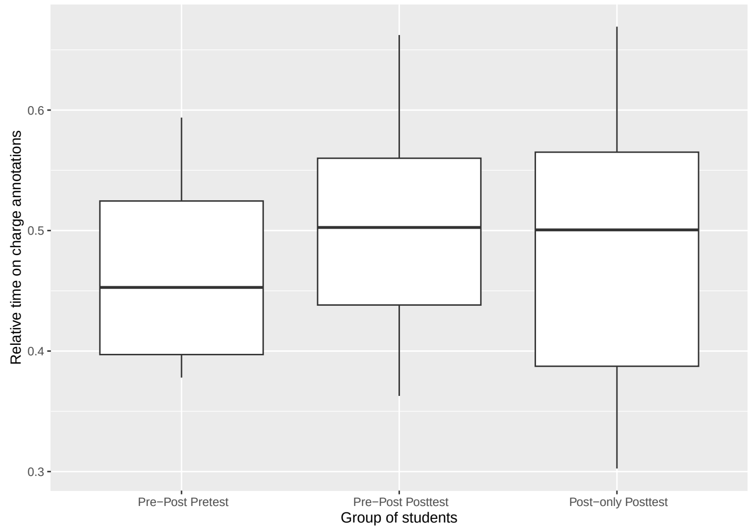


FIGURE 6.20: Boxplot diagrams regarding H2. The y-axis shows the average dwelling time on the charge annotations as a fraction of the total time spent on the corresponding stimulus. The left box summarizes the quantity for the pretest of the pre-post group, the centre box for the posttest of the pre-post group, and the right for the posttest of the post-only group. T-tests reveales a slightly significant difference with a moderate effect size between the pretest and the posttest of the pre-post group.

FIGURE 6.21: Heatmaps of the first completion task. At the top of students in the pretest of the prepost-group, on the bottom left of students in the post-only group in the posttest, and at the bottom right of students in the posttest of the prepost-group. Note that the attention is more spread out in the top heat map than in the bottom two. A quantification of this difference is shown in Figure 6.22.

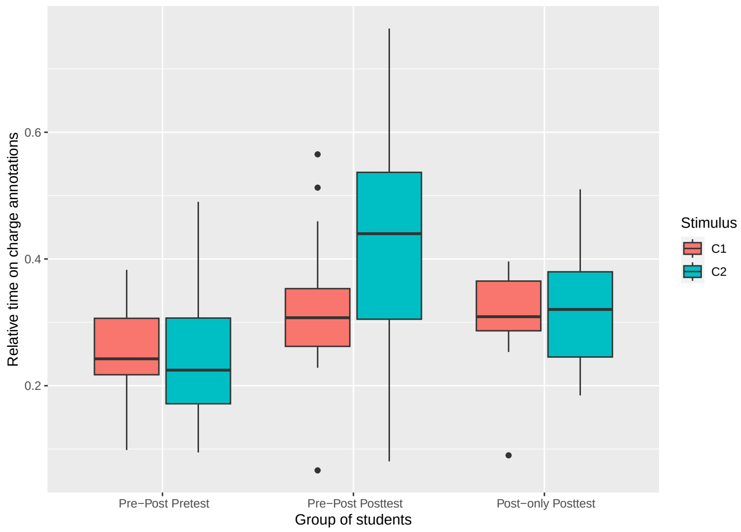


FIGURE 6.22: Relative students' duration on charge annotations in stimulus C1 and C2 regarding H2. The left boxes summarize the quantity for the pretest of the pre-post group, the centre boxes for the posttest of the pre-post group, and the right for the posttest of the post-only group. The red boxes signify the ratio for stimulus C1, the blue for stimulus C2. T-tests revealed a significant difference between the pretest of the pre-post group and the posttest of the pre-post group on both stimuli and between the posttest of the pre-post group and the posttest of the post-only group on stimulus C2.

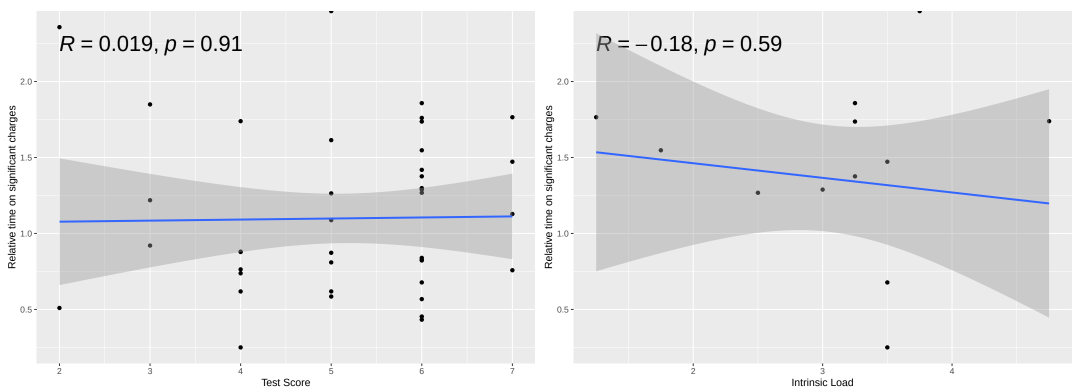


FIGURE 6.23: Scatter plots regarding H3. The y-axis shows the average dwelling time on the "significant charge annotations", which needed to be examined as a fraction of the total time spent on all the charge annotations in that option. The x-axis shows the overall score (top) and the intrinsic load (bottom).

less intrinsic load and, thus, a lower perceived difficulty of the tasks. The latter trend, however, is not significant. As for the local FTR, it was hypothesized that a higher test score is connected to a more comparative behaviour and, hence, a lower FTR. Even though there is a visible trend in that direction, this is not significant. When compared with the intrinsic load, there is no correlation whatsoever. A per-item analysis via t-test showed that there is a significant difference in the global FTR between those who correctly and incorrectly answered the task only for the second vertex stimulus ($t(31.5) = 3.7, p = .0009$ with a large effect size of $d = 1.09$). In contrast, in all other stimuli, the difference is not significant. In the case of the local FTR, there is no significant difference on the individual stimulus level. Therefore, the data only tentatively support hypothesis 4a, whereas hypothesis 4b has to be rejected.

The fifth and sixth hypotheses concerned the transition entropy. Figure 6.25 shows the distribution of the average transition entropy on the global scale in the three group/test combinations. A paired-sample t-test revealed that students of the pre-post group exhibited a significantly lower transition entropy in the posttest than in the pretest ($t(14) = 2.6, p = .02$ with a moderate effect size of $d = 0.67$). The differences between the other combinations were not significant. There was a barely significant difference on the individual item level only on the second vertex stimulus ($t(14) = 2.2, p = .049$ with a moderate effect of $d = 0.56$). H5, therefore, can very tentatively be accepted.

Concerning H6, Figure 6.26 shows the average transition entropy on the global scale plotted as a function of the overall test score and the intrinsic load. The plots reveal slightly significant correlations in both cases, which suggests that more successful students and students who perceive the tasks as easier show a lower transition entropy on the global scale. A per-item analysis, again, yields a significant difference among those who answered correctly and those who answered incorrectly only for the second vertex stimulus ($t(37.8) = 3.6, p = 0.001$ with a large effect size of $d = 1.11$). In contrast, there is no significant difference in global transition entropy between those who answer correctly and incorrectly in all other stimuli. A similar but less pronounced picture is prevalent on the local scale. Figure 6.27 shows the same correlations as above but for the transition entropy on the local scale. In this case, a per-item analysis also yields one slightly significant difference between those who answered incorrectly and those who answered correctly, but on the second completion stimulus ($t(31.3) = 3.2, p = .003$ with a moderate effect size of $d = 0.75$). Therefore, H6 can tentatively be accepted on both the global and local scales.

Triple transition ratio

The seventh hypothesis concerned a representation-specific measure, the "triple transition ratio". Figure 6.28 shows this measure as a function of performance indicator,

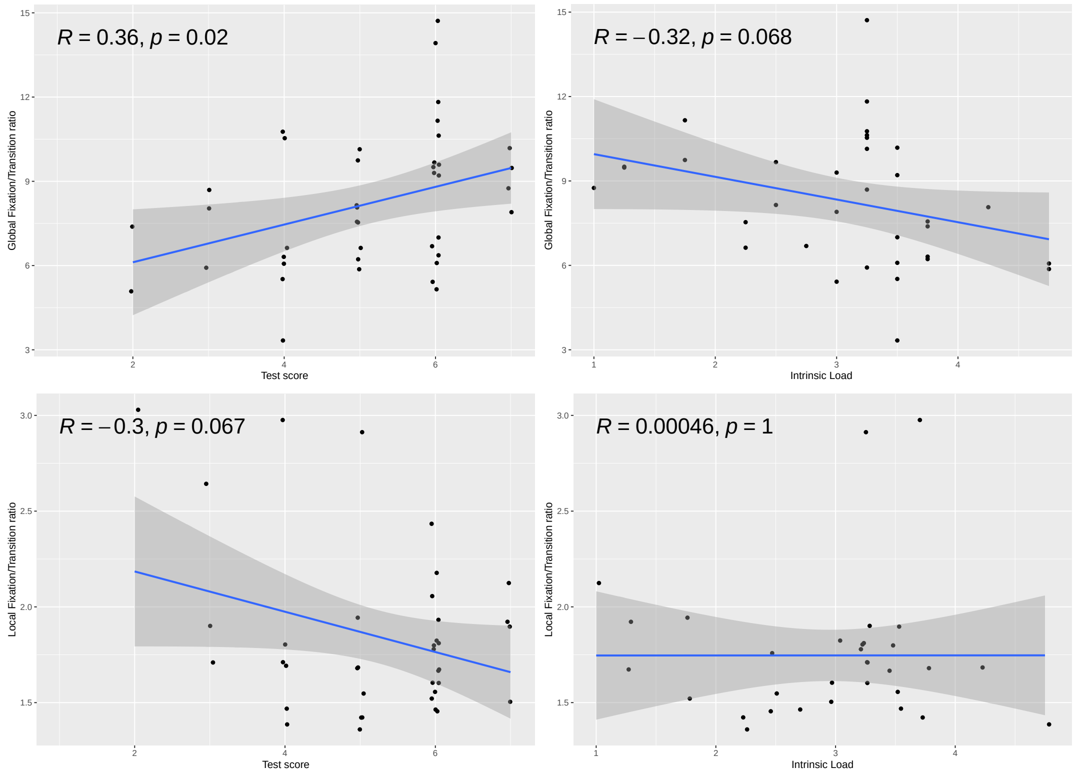


FIGURE 6.24: Scatter plots regarding H4a (top) and H4b (bottom) The plots show the correlation between the average fixation/transition ratio (y-axis) and the overall score and the intrinsic load. The upper two plots show the FTR on the global scale, whereas the lower plots show the FTR on the local scale.

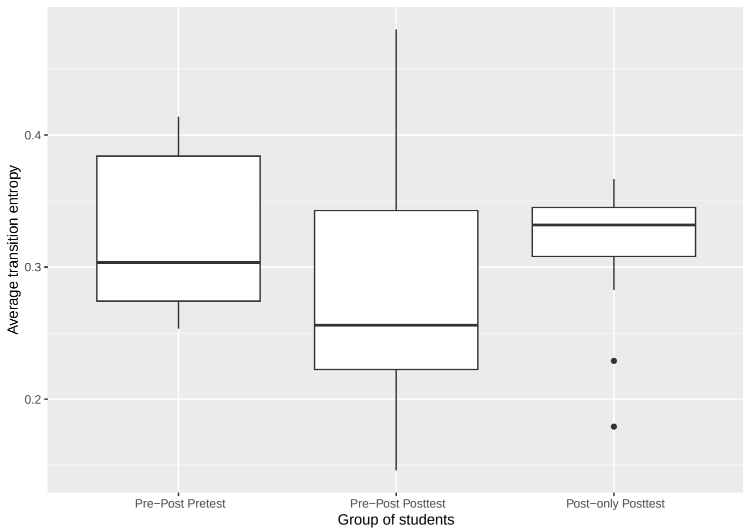


FIGURE 6.25: Boxplot diagrams regarding H5. The y-axis shows the average transition entropy across all stimuli. The left box summarizes the quantity for the pretest of the pre-post group, the centre box for the posttest of the pre-post group, and the right for the posttest of the post-only group. T-tests revealed a slightly significant difference between the pretest and posttest of the pre-post group.

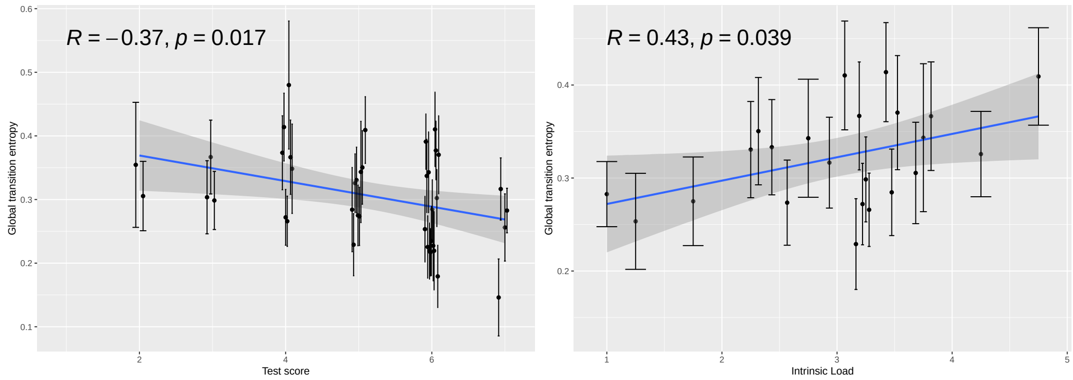


FIGURE 6.26: Scatter plot regarding H6. The y-axis shows the average transition entropy on the global scale, and the x-axis shows the overall test score (left) and the intrinsic load (right).

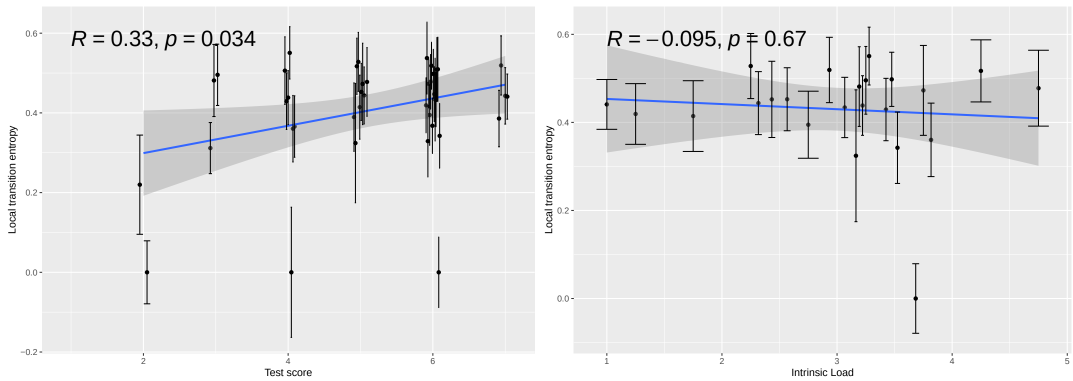


FIGURE 6.27: Scatter plots of the average transition entropy on the local scale against the overall test score (left) and the intrinsic load (right).

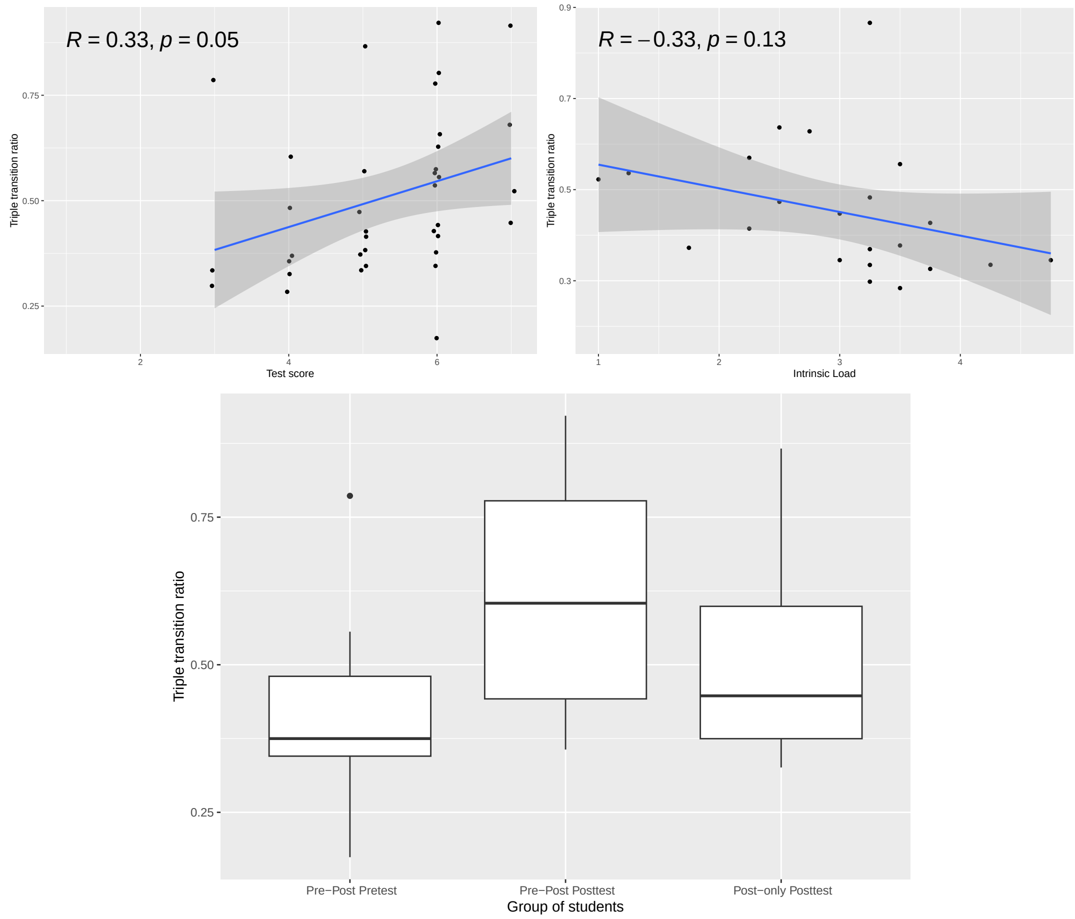


FIGURE 6.28: Scatter plot of the average triple transition ratio and the test score (top left) and the intrinsic load (top right). The bottom plot shows the triple transition ratio as a function of the group/test combination. The left box summarizes the quantity for the pretest of the pre-post group, the centre box for the posttest of the pre-post group, and the right for the posttest of the post-only group. T-tests revealed a significant difference between the pretest and the posttest of the pre-post group.

perceived difficulty, and group/test combination. The data suggest a barely significant positive correlation of the triple transition ratio with test scores and a non-significant negative correlation with the intrinsic load. This aligns with the hypothesis that higher-ability students show a higher triple transition ratio. A per-item analysis yielded no significant differences on an individual-stimulus level. Therefore, the hypothesis H7 cannot be conclusively accepted. As for the temporal evolution of the triple transition ratio, the results look more conclusive. The bottom plot shows the triple transition ratio in the three group/test combinations. A paired-sample t-test yields a significant difference between the pre- and posttest of the pre-post group ($t(11) = 6.6, p = 4 \times 10^{-5}$ with a large effect size of $d = 1.9$), but no significant differences between the other combinations. Therefore, it can be tentatively concluded that students show a higher triple transition ratio after seeing the examples and solving a previous task.

Discussion

First, the results from Table 6.13 show that the different tasks had varying degrees of difficulty. The difficulty of the decay-type diagram already introduced in the first version of the learning material (cf. Section 6.4.2) is higher than all the other tasks used within the current version of the learning material. This confirms the findings from earlier studies that the decay-type diagram seems to require more representational competencies than other diagrams, such as the scattering-type diagram.

The eye tracking data of this study allow a closer look into the processing demands of Feynman diagrams and give first insights into possible representation-specific gaze patterns.

The acceptance of H1 (cf. Figure 6.19) means that, generally, it requires some perceptual processing of an option to infer its correctness. Because of this, students who solve a task correctly generally spend more time on the correct option. The interesting part of Figure 6.19, however, is the difference between the decay-type (D2) and scattering-type (D1) diagram, as in the latter diagram, such a difference is not observed. This is in line with the fact that task D1 was one of the easiest tasks in the material (cf. Table 6.13). However, the tasks C3 and V1 were in a similar range of difficulty as D1 and they show differences between those who were answering correctly and incorrectly. This could be explained by the fact that C3 and V1 are answered before D1, so the students already had the opportunity to train on less complex tasks before they arrived at D1, where they did not need as much perceptual processing as before. This result suggests that even without showing solutions students can build up some visual fluency through solving tasks which is in line with the conceptualization of representational competencies by Rau (2017) (cf. Section 4.3).

The tentative acceptance of H2 confirms a tentative finding from the pilot study (Section 6.3) that novices tend to distribute their attention more widely across the stimuli. In the pilot study, experts focused more on the vertices as these were the relevant elements, whereas novices seemed more distracted by other elements. This result is now partly reproduced (cf. Figure 6.22) as in the first two tasks, the students spent significantly less time on the charge annotations, which are the relevant elements now, if they have not seen examples before than they do afterwards. There are no significant differences in the later tasks. This result suggests that students learn over time where to focus their attention.

The rejection of H3 means that students' solving probability does not depend on them spending significantly more time on the relevant charge annotations. The tasks were made so that, theoretically, one could have concluded that a particular diagram is the correct answer upon checking the significant charge annotations. The fact that there was no difference in the relative time spent on the significant charge annotations shows that students generally did not jump to these conclusions. Instead, they checked the whole diagram before concluding it was the correct option. In future research, it would be interesting to see if there is a difference when experts are confronted with a similar task.

The fourth to sixth hypotheses regard transition measures and are concerned with integration processes and viewing behaviour. The fixation/transition ratio discriminates comparative from focused behaviour. It was hypothesized that the behaviour differs on the global and the local scale. While for the global scale, some evidence for the connection between more focused behaviour and answer correctness could be found, this was not the case on the local scale (cf. Figure 6.24). A probable explanation for this could be a bottom effect. While the FTR can become theoretically infinitely large, it has a lower boundary at one since there cannot be more transitions than fixations (cf. the calculation of the FTR in equation 5.1). When looking closely at the two plots at the bottom of Figure 6.25, it is visible that the local FTR in most cases is already lower than two, i.e., there is more than one transition per two fixations on average. Since the AOIs in question are relatively small (cf. Figure 6.16), there are rarely more than two consecutive fixations recorded within one AOI; therefore, the FTR is already at the lower edge for all the students and differences do not become apparent in this measure even if there were a difference in the viewing behaviour.

For the transition entropy, the picture was also mixed. H5 stated that there would be a temporal development in entropy in that later views exhibit a more systematic viewing behaviour and, hence, a lower entropy. A significant difference between the group/test combinations in Figure 6.25 could only be found between the pre- and posttest of the Pre-Post group, which means that students exhibit a more systematic viewing behaviour after seeing examples and solving the tasks. The effect, however, is barely significant and relatively weak. A possible explanation could be that the intervention time used in this study was too short to affect viewing behaviour significantly. However, the tentative acceptance of H6 shows that a lower transition entropy is indeed connected to a more successful solving of the tasks on a local and global level of the stimulus. Therefore, it can be inferred that systematic viewing behaviour is connected to more successfully solving tasks with Feynman diagrams. An inference about the direction of the causation can not be made. It could either be that high-ability students who would have solved the tasks better exhibited a more systematic viewing behaviour or that a systematic viewing behaviour was helpful in successfully solving the task.

When interpreting the results of this study, it has to be kept in mind that parts of the results are representation-specific while parts are specific to the type of task. Prior research suggests that representation-specific gaze patterns are a better performance indicator (cf. Hahn and Klein (2023)). The results of this study point towards a direction of what could be a valid representation-specific gaze pattern. As shown in

Figure 6.28, the triple transition ratio is loosely connected to a more successful solving of the tasks, even though the correlation is barely significant. It is visible, however, that students exhibit a higher triple transition ratio after seeing examples and solving the tasks. Therefore, this viewing behaviour is fostered by the intervention and the test itself. Future research must confirm whether this viewing behaviour benefits students when solving tasks with Feynman diagrams.

The full version of the learning material, the answers to the questionnaires, eye tracking data as well as analysis scripts are publicly available. The links can be found in appendix E.

6.6.8 Conceptional challenges with particle interactions

The students' answers as free-text answers to the comprehension questions (cf. Table 6.7) can be categorised into adequate and inadequate conceptions. The following description first explains the conceptions underlying the inadequate conceptions and then describes the challenges exhibited by the students. As an overview, Table 6.15 summarises these conceptual challenges. Since there are only one or two mentions of each inadequate conception, this is a qualitative description of the conceptions being present and cannot be a statistical description of the prevalence of these conceptions. Also, it cannot be an exhaustive catalogue of possible inadequate students' conceptions but rather a collection of what could be present. The following description emphasizes quotes from the free-text answers to the comprehension questions that the author translated.

Diagram elements.

Before going into the details of inadequate conceptions related to particle physics concepts, there are some inadequate conceptions about the terminology of the diagram. One student stated that *"a connection [between two knot-points] is often an interaction particle, but sometimes a vertex"* which shows confusion about the term "vertex". The student probably refers to a straight line as a vertex, while what they refer to as a "knot-point" is referred to as a "vertex" in the learning material. Besides mere confusion of terms, there are also inadequate conceptions concerning the function of the different elements. One student stated that *"particles are evenly distributed in a diagram"* which refers to the notion that the precise form of the diagram matters, which is not the case. Another easily confusing element is the arrows of the lines in the diagram. While they depict the distinction between (matter) particles and antiparticles, one student stated that *"an arrow to the left stands for 'in' and an arrow to the right stands for 'out'"*. This confusion is likely rooted in the fact that incoming and outgoing particles were also a topic in the learning material, but rather on the subject of sides of the diagram. The sides of the diagram were the source of a conceptual challenge as well, as one student referred to them as *"[depicting the electrical charges]*

Topic	Conceptual Challenge	#
Diagram elements	confusion between line and vertex	1
	particles are ordered in a diagram	1
	arrow directions	1
	sides of the diagram	2
Particles	Confusion of interaction particle and anti-particle	1
	anti-particles are the opposite of "real particles"	1
	correspondence between particle and interaction particle	1
	interaction particles are difficult to observe	2
Interactions	Names of interactions	1
	Interaction takes place where the interaction particle is shown	1
	Vertices are where forces act	1
	energy interacts	1
	Particles change their shape	1
	Electrons are neutralised	1
	Electrons transform into photons	1
	Particles transform into anti-particles in the weak interaction	1
	Confusion as to what is exchanged	1
	Interaction particles are responsible for energy conservation	1
Charges	Confusion of charge and energy	1
	Charges move in a diagram	1
	Charge is transformed	1
	Charges are measured between vertices	1

TABLE 6.15: Inadequate conceptions mentioned in the free-text answers to the comprehension questions after the learning material (cf. Table 6.7). The right column shows the number of occurrences of that conception among the answers.

on the one side [and] on the other side the weak charge" and being *"connected [...] by interactions (the wavy lines)"*. These statements acknowledge that the diagram's sides have a meaning (depicting incoming and outgoing particles) but misattribute them to depicting the different charges or the particles interacting.

Particles.

As Feynman diagrams are a representation of the interactions between different particles using different types of lines, it can be argued that the lines are a schematic representation of particles, even though it is important to note that they are not "picture-like" representations (for a more detailed analysis of these interpretation problems see Section 2.1.3). In any case, students are taught in the material that different lines stand for various kinds of particles and in concrete examples, the lines are directly attributed to concrete particles. Thus, there are different conceptions about the nature of particles present in the students' answers.

Many of these inadequate conceptions are language confusions between different terms, such as referring to interaction particles as anti-particles (*"particles which are in interaction with each other emit an anti-particle"*), even though both are entirely different classes of particles, which are depicted differently in a diagram. Also, anti-particles are opposed to "real particles" (*"arrows to the right describe real particles, arrows to the left depict anti-particles"*), even though anti-particles are also indisputably real. The notion of a "real particle" could be grounded in the term virtual particle in the learning material.

There are more inadequate conceptions concerning the nature of interaction particles. One is that there is a correspondence between particles and interaction particles (*"each particle mostly has a specific interaction particle"*). While there is a correspondence between interactions (or charges) and interaction particles, there is no such thing between (matter or anti-) particles and interaction particles since different interaction particles can mediate interactions between particles. Another inadequate conception of interaction particles addresses their observability. Two answers described interaction particles as *"a state which is difficult to observe"* and *"faster particles which are also shorter-lived"*. This is an inadequate conception, as interaction particles are not always short-lived. Photons have an infinite lifetime and are not difficult to observe. However, the notion might also result in interaction particles being predominantly depicted as virtual particles in the learning material, which are inherently impossible to observe directly.

Interactions.

Particle interactions are the most central part of the learning material's second part and the explanations' core. In general, particle interactions are explained as mediated by interaction particles, i.e., an exchange of energy and momentum. As this is a new and relatively abstract concept, it is not surprising that most of the inadequate conceptions stated by the students are around this topic. As before, there is mere language confusion, in this case, falsely remembering the names of interactions (*"proton and neutron interaction and electron interaction"*instead of weak and electromagnetic interaction). Some of the inadequate conceptions refer to the depiction of the interaction by the diagram, such as the conception that the depiction of the interaction particle in the diagram shows the place of the interaction (*"[The interaction particle] shows at which place the interaction occurs"*) or that the vertices show where forces act (*"crossings in which the respective force acts"*). In contrast, the diagram is just a schematic depiction of the particle interaction. However, the learning material suggests interpreting the lines as the respective particles, which is reflected in these answers.

Most inadequate conceptions about this topic are about the mechanism of how the interaction takes place since this is the most abstract part of the explanations of the learning material. [8] One student stated that *"the diagram shows how the energy interacts"* which can be interpreted as seeing energy as part of a particle which interacts. In contrast, it is the whole particles that interact with each other. It was also stated that *"the interaction particle describes the shape the particle takes during an interaction"* and that *"electrons need to be neutralised to be transformed"* which both share the notion that particles change their properties during a particle process. More adequately, a specific particle has some unchangeable properties, such as its different charges. In most particle processes, the particles transform into other particles with different properties. This conception is reflected in some of the statements but is inadequate as to how these transformations take place. E.g., when students state that *"in the electromagnetic [interaction] electronic particles are transformed into photons"* and that a *"transformation from a particle into an anti-particle"* takes place in the weak interaction, they are not entirely wrong, but provide an incomplete description of how particles transform. Another statement referred to the role of interaction particles in particle processes. In stating that their *"role is to make sure no energy is lost"*, the conception attributes a fundamental property of every interaction to just the interaction particles.

Charges.

As charges are a fundamental concept when learning about particle physics (cf. Section 2.2.2), they played a significant role also in this learning material. They were the source of inadequate conceptions in the students' statements.

First, there was confusion between the terms charge and energy when a student stated that *"charges need to cancel out each other for energy conservation to be valid"*. This confusion is probably rooted in the fact that students learn about conservation

[8] There is a more detailed account about this part in Section 2.1.3, but Feynman diagrams themselves are the compromise between a "cartoon-like" and a technically correct but purely abstract description.

laws first and often only in the context of energy conservation. Therefore, energy is associated with the term conservation.

Another issue which causes inadequate conceptions is the question of what happens to charges in a particle process. One student states that *"the movement of charges"* is depicted in the diagram, while another one states that the diagram is *"somehow about the transformation of charges"*. The underlying conception might be that charges themselves are manipulated in the diagram rather than being properties of the particles and the particles being transformed.

Another statement reads that *"the charge is measured at specific points, namely at connections between vertices"*. The underlying inadequate conception is that charges could be measured within a diagram, which is not the case. The only points where a measurement would be possible are the initial and the final states, not in between the vertices.

Discussion

Most of the conceptual difficulties unravelled in this study are concerned with the nature of interactions, which is in line with this being the main topic of the interactive part of the learning material where the questions were focused. However, the same number of mentioned challenges were concerned with three other topics, namely diagram elements, the nature of particles, and charges. The latter topic was the central part of the first part of the learning material. In contrast, the other two topics were not the main topics of the learning material but were mentioned mostly throughout the interactive part.

While there is a vast body of research about students' conceptions of the particulate nature of matter (cf., Harrison and Treagust (2003) for a comprehensive review), this is mainly from a chemistry perspective. In contrast, research on students' conceptions of subatomic particles is sparse. Wiener et al. (2015, 2017b) tried to introduce lower-secondary students to the quark model of subatomic particles using purely typographic illustrations and found that the students generally accept key ideas about the subatomic nature of matter. However, they found linguistic accuracy problems, like using subject-specific terms. Similarly, Gourlay (2016) found in her study on students' understanding of particle physics that students often lack clarity in terminology. Specifically, she found problems in students' understanding of annihilation and pair production. Tuzón and Solbes (2016) found that students' understanding of particle physics is primarily based on a classical picture, with some ideas from a modern science view apparent but not stable. On a diagram side, Dorris and Rau (2022) analysed students' view of the atomic orbital energy diagram and found that the arrows depicting the spin states of the electrons are a critical source of misconceptions. The current study's results are essentially in line with these prior findings.

Students generally seem to accept the idea of interaction particles but exhibit conceptual challenges, partly based on terminology, partly on classical ideas of how particles work, and partly on a literal view of the diagrams.

The conceptual difficulties exhibited by the students can be categorised into four topics, which are well-reflected in the learning material. However, none of the inadequate conceptions are mentioned more than twice, and most are mentioned just once. On the one hand, these results of the current study can only be exploratory and point to the direction of a catalogue of conceptual difficulties with Feynman diagrams; therefore, a quantitative analysis of the conceptions would be out of scope, on the other hand, it suggests that none of the observed conceptions are inherently stable or widely prevalent among students. The method also does not allow us to reliably distinguish between linguistic problems and evident conceptual difficulties. Future research is needed to investigate these difficulties in order to inform educational practices in teaching particle physics or to create a test instrument.

Even though the results are preliminary, some conclusions regarding using Feynman diagrams in teaching can be drawn, particularly on challenging terms, such as different kinds of charges and particles or elements that complicate a diagram. These issues will be discussed in Section 7.1.2.

Chapter 7

Discussion

This research project about designing learning materials for using Feynman diagrams in physics education is the first to systematically attempt to do so, making it an essential contribution to the field of physics education. Additionally, using the design-based research approach allows for potentially producing a valuable product for practice. Furthermore, eye-tracking data collection allows for valuable insights into students' thought processes.

In this chapter, I am discussing the outcomes of the studies presented in the previous chapter with respect to the learning goals stated in Chapter 3. The concrete outcomes of the studies were already discussed in the discussion parts of the single studies in the previous chapter. See Section 6.1 and especially Section 6.2 for the discussion of the results from the expert study, Section 6.3 and the discussion part therein for the pilot study, Section 6.4.4 for the first student study, Section 6.5.3 for the second student study, and Sections 6.6.7 and 6.6.8 for the discussions of the two respective research questions of the third student study.

7.1 Research goal 1: Opportunities and Challenges of Teaching with Feynman Diagrams

The first research goal as stated in Section 3.1 was to answer the question *whether* Feynman diagrams can be used to teach particle physics concepts. To this end, possible learning goals and challenges of using Feynman diagrams were investigated.

7.1.1 Learning Goals

As stated in Section 6.2, four learning goals on what can be achieved with Feynman diagrams from the expert interviews have been derived from the expert interview study and literature about teaching particle physics. In this section, I will revisit them and discuss their suitability for high-school students.

Learning goal 1: Charges

The first learning goal was for students to be able to use Feynman diagrams to check if the charges are conserved in a given particle process. The idea of this learning goal was to convey the importance of the charge concept as well as the concept of conservation laws. Charges play a significant role in the teaching of particle physics, as, e.g., Lindenau and Kobel (2019) use charges as one of the three basic concepts in their educational reconstruction of the SMPP, which is interconnected with the concepts of fundamental interactions and elementary particles. Other than that, charges are rarely explicitly mentioned but rather than as a property of particles without assigning much meaning to it (e.g., in McGinness et al. (2019) and Pavlidou and Lazzeroni (2016) the charges are part of the puzzle pieces/playing cards). Furthermore, the curriculum study of Kranjc Horvat et al. (2022) showed that while the electric charge is discussed (or, at least mentioned) in all of the investigated curricula, the conservation of charge(s) is only mentioned in just over a third of them. In contrast, almost all of them mention the conservation of energy and (linear) momentum. Other charges than the electric one are not discussed in any of the curricula, even those with a dedicated topic for particle physics.

This points towards the fact that the notion of the conservation of charges is a topic worthwhile to discuss with high school students: They can apply prior knowledge (knowledge of the electric charge and conservation laws, sometimes even of the conservation of electric charge) to something new (the application of conservation laws to something else than kinematic variables and at least the existence of other than electric charges). As already stated in Section 6.2.1, within this learning material, only two of the charges and interactions are discussed, as the message is that there are different ones, and a third one would impose an unnecessary layer of complexity.

Regarding the question of to what extent this learning goal was achieved with the learning environment, the answer is a tentative yes. It was primarily tested summatively by using tasks with Feynman diagrams in all three student studies (cf. Sections 6.4, 6.5.2, and 6.6.7). While there were apparent ceiling effects in the first two studies, which a "teaching-to-the-test" effect could explain, the results from the third study showed a clear picture that students, in principle, had no problem applying the concept of charge conservation to the given tasks. Still, they had diagram-specific difficulties, which will be discussed further below.

Learning goal 2: Interaction particles

The second learning goal was for students to understand the concept of interaction particles and their role in particle physics. As with the previous one, this learning goal was also partly based on one of the three basic concepts used by Lindenau and Kobel (2019). However, from the literature, it is not as straightforward as the previous learning goal that this is a worthy learning goal. Some educational approaches favour the field-theoretic description over the description with the image of particles (e.g., Allday (1997), Bertozzi (2013), Daniel (2006), Hobson (2010), and Organtini (2011)). These approaches, however, do not disregard the interaction particle altogether, they instead introduce the particle as some "artefact" of the field; hence they still claim that the interaction particle which is exchanged as a virtual particle is real. Some philosophers of quantum field theory dispute this notion of reality (see the discussion in Section 2.1.3). Their arguments are based on the fact that Feynman diagrams, in their initially intended way, are mere representations of terms in a perturbative series, and therefore, the only object with a claim to reality is the squared sum of all diagrams contributing to a specific process. This argument is addressed with the third learning goal. Within the learning material presented in this thesis, a line more closely related to that of Stöltzner (2018) is followed. Feynman diagrams can have different functions while especially physicists pragmatically see them, in line with the second function Stöltzner presents, as a representation of a particular process which in fact would be described by an (infinite) series of diagrams, but which is represented usually only by the lowest-order diagram.

These considerations aside, the interaction particle is a useful model that can be utilized to motivate the similarities and differences between the fundamental interactions and the peculiarities of particle physics, namely the locality of an interaction. In contrast to Newton's description of gravity, where two bodies act on each other from a distance, an interaction in particle physics is always described by the exchange of (virtual) particles. In some texts, this is explained by an analogy of two boats on a lake or ice skaters which throw balls at each other to model a repelling force or boomerangs in opposite directions to model an attractive force (cf., e.g., Allday (1997) for a critical review of this analogy). However, this analogy reinforces a naive conception of particles as small balls, which is incompatible with more modern particle conceptions where particles are artefacts of quantum fields (see above). Therefore, a particle description mentioning the fundamental interaction is preferred, which could be expanded to mention the underlying field.

Throughout the learning material, students learn the lowest-order diagrams for different processes, which are mediated by different fundamental interactions. As stated by the theory of Social Semiotics (cf. Section 4.3), practice is essential in acquiring fluency in discerning semiotic resources such as Feynman diagrams. Therefore, in this part, students must learn many examples of processes and their corresponding lowest-order diagrams to understand better how they work.

The Feynman diagrams are explained in a "story-like" manner, meaning the process at different vertices is mentioned (emission/absorption of an interaction particle, annihilation, pair production). This story-like approach breaks down when internal lines are depicted vertically, such as in the typical lowest-order diagram of the process $e^+e^- \rightarrow \gamma\gamma$ (Figure 7.1). For internal lines, the Feynman rules (cf. Section 2.1.2) do not distinguish between particle and anti-particle, hence it does not matter whether a virtual particle is depicted as electron or positron.

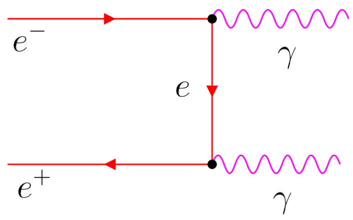


FIGURE 7.1: Lowest-order diagram of the process $e^+e^- \to \gamma\gamma$. The arrow of the virtual particle could be omitted.

Therefore, a depiction with a vertical line poses no problem for physicists who use that diagram, but for telling the "story" of that diagram, it does. The diagram could be drawn differently with the internal line tilted so that the processes could be described as an electron emitting a photon and then annihilating with the positron or vice versa. To mark the fact that it does not matter whether it is a particle or an anti-particle, the annotation of the internal line in Figure 7.1 only reads e, in contrast to e^- and e^+ for the external lines.

It cannot be conclusively said whether the learning goal was achieved. It was only systematically tested in the third student study (cf. Section 6.6.8). The answers students gave mainly aligned with the messages in the learning material. However, the students' answers still had inadequate conceptions regarding the nature of interaction particles and interactions. However, none of these inadequate conceptions seemed widespread or inherently stable (cf. Table 6.15).

Learning goal 3: Mathematical nature of modern physics

The third learning goal was split into two and concerned the superposition of several diagrams. With this learning goal, the considerations discussed in Section 2.1.3 and repeated in the previous discussion are addressed. It also addresses the need to discuss the Nature of Science (cf. Section 2.2.3) in highlighting the model character of Feynman diagrams and the profoundly mathematical nature of modern physics. Considering this, it could be argued that the mechanism addressed with this learning goal should be introduced at the very beginning to convey an image of Feynman diagrams closer to their actual use. However, it was decided to introduce the two learning goals mentioned above before mentioning the superposition of diagrams to give students a working knowledge of particle interactions, which they can then apply. This is in line with the pre-training principle from Mayer (2002a) (cf. Section 4.3), which states that the essential terms and concepts need to be introduced before they are required.

Even though this learning goal aims at conveying a message from modern physics where advanced mathematics plays a critical role[1], the mathematics which is needed to achieve this learning goal is not too complicated for upper secondary school. As stated in Section 6.2.3, the mathematics in the first part of the learning goal are reduced to the binomial formula, which should be a topic in lower secondary school.

[1]The mathematical account presented in Section 2.1.2 only gives a glimpse into the mathematics which are needed to calculate cross-sections in real physics problems efficiently, but already there delta distributions and multi-dimensional integrals are introduced which go beyond school mathematics.

The mathematics in the second part of this learning goal is slightly advanced as it includes infinite series. However, this topic should be part of the mathematics curriculum in upper secondary school.

The learning goal also connects the topic of particle physics with the topic of quantum mechanics. Most students will see the double-slit experiment if they discuss quantum mechanics in their physics course (Krijtenburg-Lewerissa et al., 2017). The superposition notion of Feynman diagrams offers an interesting view of this experiment. While the double slit experiment is usually taken as a pivotal experiment to show the wave nature of light, it can also be viewed in the particle picture if it is assumed that the photon neither takes one slit or the other, but both and it "interferes with itself". In the same way, neither the one diagram nor the other is realised to turn the initial state into the final state, but a superposition of both – or rather all infinitely many.

However, the second part of this learning goal offers another opportunity to discuss NoS as it highlights the approximate nature of scientific theories. It can be seen that the Feynman diagram method will never give exact results. However, the calculations obtained with this method provide the most accurate results of any physical theory, which are still confirmed (cf. Fan et al. (2023) for an example of a recent update of the measurement of the magnetic moment of the electron, which already was a benchmark for the precision of the theory during the early discussions of QED). At the same time, discrepancies give hints for modifications of the theory (cf. Wolchover (2021) for an article about the discrepancy between the calculation and experiment of the muon's magnetic moment).

Learning goal 4: Particle discoveries

The fourth learning goal was for students to be able to apply the knowledge they have learned thus far to the work of scientists. The idea is that they can trace the different stages of the discovery of a particle with the example of the Z boson and, by that, make the connection between the theory, illustrated by Feynman diagrams – where one Feynman diagram is the representative of a series of diagrams as stated in the above discussion of the second learning goal – and the experiment, illustrated by bubble chamber pictures, event displays and cross-section diagrams. This lets students connect multiple different representations where the representations have functions in line with the DeFT framework by Ainsworth (2006) (cf. Section 4.3), such as constraining the meaning of a diagram or connecting with it to construct a deeper understanding.

This learning goal conveys an essential message about the Nature of Science as it shows the close connection between theory and experiment. Section 6.2.4 shows the different NoS aspects this learning goal conveys. In principle, it could also be shown that science is no linearly advancing field where one discovery always builds on the previous one. For this message, however, a more historically accurate account would need to be given, showcasing the different failures and sidetracks of the discovery. This account, however, would require another PhD thesis in the history of science and would be far beyond the scope of this work. The story told within the learning material presented here is rather a "quasi-history" as it is often told in physics textbooks (cf. Passon (2022) or Griffiths (2008, p. 13))[2]. This could be a problem if the reader is unaware that the historical account is shortened because inadequate ideas about scientific progress could be enforced. In the worst case, it could even transform into an anti-scientific attitude: Current scientific achievements are not valued as much as historical achievements because, for historical achievements, one often already knows the larger context, while for current achievements, the "bigger picture" seems a bit "blurry" which could result in the perception that current achievements are less relevant than historical ones. This perception could be avoided by making learners aware that people did not have the bigger picture at the time of historical accomplishments and that today's science is constantly progressing, and a bigger picture of today's achievements will only form in the future.

7.1.2 Challenges

The learning goals discussed in the previous section pose opportunities for teaching particle physics with Feynman diagrams. This section discusses the challenges of using Feynman diagrams with upper-secondary students.

Charge conservation

As discussed above, the first learning goal was generally achieved because students understood that charges must be conserved. However, two challenges became apparent in incorporating this learning goal into the material.

First, it was observed that some students refer to the charge as energy, which is likely because they know about conservation laws from learning about energy conservation. This could lead to confusion between the two terms energy and charge. The difference between these terms is essential, as the energy of a particle can change, e.g., by accelerating it, and isn't even invariant, i.e. it is different in a different frame of reference. In contrast, the charge is an unchangeable property of a particle, which means that if the charge changes, the particle changes. This leads to two completely different consequences of the conservation laws: Energy conservation forbids specific energy ranges for particles in the initial state, while charge conservation forbids certain particle combinations to occur (which energy conservation only does indirectly by setting an upper boundary on the rest mass of particles in the final state depending on the energy of the particles in the initial state).

[2] Of course, also the historical account told in Section 2.1.1 is such a quasi-history.

Second, when using the diagrams as they were used in the current learning material, there is the possibility that students ignore the fact that the numbers notated at the particles are charges and, therefore, think of the tasks as simple context-free mathematics tasks. This could be avoided in a future version of the learning material by providing more context to the numbers, such as giving units[3] and featuring it more prominently in later stages to motivate better why charge conservation is a critical feature of Feynman diagrams. In the current version, charge conservation is mentioned in the explanation of one process in the second part (cf. Section 6.6.3), where the fact that the muon has to transform into a muon neutrino after emitting a W boson is explained by the fact that the latter has a non-zero charge, so after the emission, another particle must be present. Hence, the total charge is still conserved.

Interaction particles

As for the second learning goal, it can’t be said conclusively whether it was achieved. Several points turned out challenging regarding this topic. Launer (2020) analysed different concepts of "interactions" in physics and tested them for suitability for teaching to upper secondary students. He concluded that classical interaction concepts of forces and fields are comparatively easy to teach. In contrast, modern concepts such as the geometrical concept, which describes gravity in the framework of general relativity and the idea of interaction particles, are methodically more challenging.

While this learning goal introduces the photon as the interaction particle of the electromagnetic interaction, students might have heard about the photon as "the particle light is made of" – this, or a version of this, was the answer stated most often to the question "What is a photon"?" in the second student study. As Passon (2022) points out, there is a difference between the original conceptualisation of the photon as the "light quantum" and the modern conceptualisation as a vector boson responsible for mediating the electromagnetic interaction. Furthermore, the fact that interaction particles are drawn as "waves" in a Feynman diagram could also contribute to a confusion around the topic of "wave-particle duality" which already is a major source of misunderstanding within the teaching of quantum physics (for a comprehensive review, see Krijtenburg-Lewerissa et al. (2017)).

Another conceptual challenge concerning this topic is the distinction between virtual and interaction particles. In the two diagrams primarily shown in this learning material, the scattering-type and the decay-type diagram, every interaction particle is also a virtual particle, i.e., depicted as an inner line, and vice versa. In these cases, the explanation that the exchange of an interaction particle mediates an interaction is valid. It is, however, not a general rule. See, for example, the lowest order diagram for the electron-positron annihilation (Figure 7.1). In this diagram, the inner line, i.e., virtual particle, is an electron or positron, and therefore, no interaction particle.

[3]Some particle physicists prefer the charges being unitless

In contrast, the particles in the final state are two photons introduced as interaction particles.

An essential part of the learning goal was also conveying the differences between two fundamental interactions. One reason why the electromagnetic and weak interactions are so different is grounded in the distinct properties of their interaction particles. While the photon is massless and therefore has an infinite reach, the interaction particles of the weak interaction are heavy, which is the reason for the low effective strength of the weak interaction and its limited reach. Also, the W bosons have electric and weak charge contrary to the photon, which is why the weak interaction mediates particle transformations via the $W^{+/-}$ boson (such as, most prominently within the learning material, the muon transformation or the $\beta^{+/-}$ transformation). This notion requires a more profound knowledge of how the interaction works and how the properties of the interaction particles influence the properties of the interaction. In the current learning material, this issue is not addressed, which could lead to very superficial learning of the topic, as is indicated by some of the students' answers to the comprehension questions on page 137, such as ideas that interaction particle is responsible for energy conservation, or particles transform into anti-particles.

Conceptually challenging parts of Feynman diagrams

Feynman diagrams might pose some particular challenges besides the challenges directly connected to the above-mentioned learning goals, as discussed below. An issue that became apparent in the first and second student studies but was systematically analysed in the third was the apparent increased intrinsic cognitive load the decay-type diagram posed compared to the scattering-type diagram. This becomes apparent when comparing the stimuli D1 and D2, i.e., a typical scattering-type diagram and a typical decay-type diagram in Table 6.13. It is demonstrated that students have more difficulty examining the decay-type diagram correctly than the scattering-type one. In the eye-tracking data, this difference becomes evident when comparing these two stimuli in Figure 6.19. Here, it is shown that students who solve a task with the decay-type diagram correctly spent more time on the correct option than the wrong one. This trend is not observed with the scattering-type diagram. This fact suggests that more mental capacity is needed for the decay-type diagram to examine the diagram thoroughly. As the extraneous load is hypothesized to be constant throughout the tasks and the germane load does not change drastically, the inference is that the intrinsic load posed by the decay-type diagram is higher than that of the scattering-type diagram. A possible reason for this could be the different sequential order of the diagrams: While a scattering-type diagram is strictly sequential, meaning that two transformations happen strictly after each other, the decay-type diagram does not have such a strict order. Even though the two vertices are "ordered", meaning first, the W particle is created. Then the W particle transforms into the particle-anti-particle pair, and a particle of the final state is already created at the first vertex, which is why the result of the first vertex seems to "interfere" with the result from the second one.

Students should obtain meaningful "chunks" of a representation for perceptual fluency. These parts of the representation are seen as one piece at a time. In the case of Feynman diagrams, these chunks could be vertices, i.e., three particles which meet at one point. It was tried in the first draft of the learning material to highlight these chunks (cf. Figure D.1). However, throughout the first student study, students ignored these highlights mainly because they had already solved the task and weren't interested in the solution anymore. In the second version of the learning material used in the third student study, highlighting was abandoned to show correct and incorrect vertices (cf. Figure E.1). The results demonstrated in Figures 6.22, 6.25, and especially 6.28 (bottom) show that there are indeed slight effects of seeing the examples which are not merely a testing effect. Especially the representation-specific triple transition ratio significantly changed after seeing the examples and the tasks, compared to seeing the tasks or the examples alone.

To summarise, a factor complicating diagrams is "different things happening simultaneously", i.e. above each other. In practice, this does not matter, of course, since the time order between the initial and the final state does not matter (cf. Section 2.1.3), but it does when explaining it "by telling its story". This, again, becomes particularly challenging when dealing with diagrams like those in Figure 7.1, where two vertices are on top of each other. One student who examined this diagram was even confused about the time direction in the diagram and suddenly wanted to read it from the top to the bottom because they tried to make sense of it but were utterly confused even though they had answered everything perfectly correctly and made the correct deductions [4].

Another issue with Feynman diagrams which has been discussed, among others in the expert interviews (cf. Section 6.1) is the issue of arrows. While they are an essential element in distinguishing particles from anti-particles, they could also lead to confusion among students. The interview results from Section 6.6.8 show that some students confused the arrow directions for incoming and outgoing particles. As I was concerned about this, I omitted different arrow directions in the first version of the learning material. I only added arrows in one direction, as done in the material by Berg and Hoekzema (2006). In a refined version of that material, I omitted the arrow directions entirely but introduced them from the beginning in the second version (cf. Figure E.1). This was done to keep the diagrams introduced for the first learning goal connected to later ones where the arrow directions are needed. The meaning of the directions is then only introduced at the start of the second part, aiming for the second learning goal. The function of arrows in diagrams that do not

[4]The data analysis of the interview where this happened is not part of this thesis.

depict a direction of movement is also investigated by Dorris and Rau (2022), who investigates the atomic energy orbital diagram. They concluded that arrows are a critical source of conceptual challenges students exhibit when interpreting these diagrams. While only having limited evidence, I can tentatively support the issue in the case of the Feynman diagram.

Conceptually challenging terms connected to Feynman diagrams

Several conceptual challenges occurred within the learning material because of terms that were either unknown or similar to known terms, which caused a certain level of confusion. These terms are discussed here. One of the confusions was already mentioned above: the charge-energy confusion. This is caused not by semantics but by a functional similarity between the two terms.

Closely connected to this is a confusion between weak charge and spin. This is also rooted in a functional similarity, as some students have heard about the concept of spin and might know that this is often a half-integer, noted as "spin up" or "spin down". What is introduced as the weak charge in this learning material is the third component of the "weak isospin", whose description is mathematically similar to that of spin but is a dimensionless number and physically not connected to any type of angular momentum. Therefore, its description as "charge" makes more sense than that as "spin". Also, as it is conserved in interactions with the weak, electromagnetic, and strong fields,[5] it is a physically meaningful example of charge conservation. Also, students may have heard about the weak interaction, or primarily "weak force", and connect the term to the concept, but usually without making a deeper connection.

Another conceptually challenging topic is the different classes of particles. As Feynman diagrams are a central topic in particle physics, the concept of a particle should ideally be introduced first. However, the question of what a particle really *is* can spark heated discussions among particle physicists (Wolchover, 2020). It is consensus, however, how particles can be described mathematically and which types of particles exist. Therefore, it can be worthwhile to postpone the question of the "true nature of a particle" to a point where learners already have some familiarity with the particle concept. This familiarity could be acquired by using Feynman diagrams. However, introducing different particles through simple naming of the elements of a Feynman diagram could lead to rote learning of the names. The results in Section 6.6.8 show that different categories of particles are easily confused, such as anti-particles and interaction particles, or anti-particles and virtual particles (when anti-particles are opposed to "real particles").

[5] not necessarily in interactions with the Brout-Englert-Higgs field, which does not play a role within this learning material

7.2 Research Goal 2: Suggestions for Practices when Teaching with Feynman Diagrams

The second research goal was to investigate how Feynman diagrams can be effectively used for teaching. Some suggestions can be distilled from the opportunities and challenges discussed in the previous section. These serve as tentative answers to the research question stated in Section 3.2 but must be tested in future studies.

Introduce diagrams in order of progressing difficulty

From the data presented in the third student study, it can be inferred that diagrams with a vertical intermediate particle are more challenging for students. Based on this observation, I would suggest the difficulty progression proposed in Figure 7.2, which starts with single vertices to explain the notion of interactions depicted by lines meeting each other and the principle of charge conservation, then build from these the scattering-type diagram with a horizontal intermediate particle. A more complex diagram than that is the decay-type one since one of the final state's particles is already produced at the first vertex. An alternative depiction of the scattering-type diagram can be introduced where the interaction particle is vertical. With this diagram, the notion that a Feynman diagram tells a continuous story can be challenged, as it is not clear which of the particles emits and which absorbs the interaction particle, so this diagram can be used to explain that between the initial and the final state, the time order does not matter.[6] This can then be expanded using the annihilation diagram in which it is unclear whether the intermediate particle is an electron or a positron.

These are only the leading-order diagrams ordered for progressing difficulty. Higher-order diagrams have to be considered when introducing the third learning goal, like those in Figure 7.3. These are diagrams with more than two vertices, i.e., more than one virtual particle. These virtual particles can be additional photons (like in the top two diagrams) or loops (in the bottom).

Vary the strategies on how to explain the diagrams

A second suggestion for practice based on the results of the second student study is the visual strategies, which should be promoted when using Feynman diagrams. When diagrams are used in practice, perceptual fluency is required. To acquire these, learners need many simple tasks, such as those described in the eye-tracking studies within this thesis. One outcome of these tasks might be an understanding of charge conservation, but at the same time, learners get familiar with the diagram

[6] In physics calculations, when the two diagrams depict the same process, they are usually called the s-channel and the t-channel after their so-called *Mandelstam variables*. s denotes the square of the centre-of-mass energy, and t is the square of the 4-momentum transferred between the two particles.

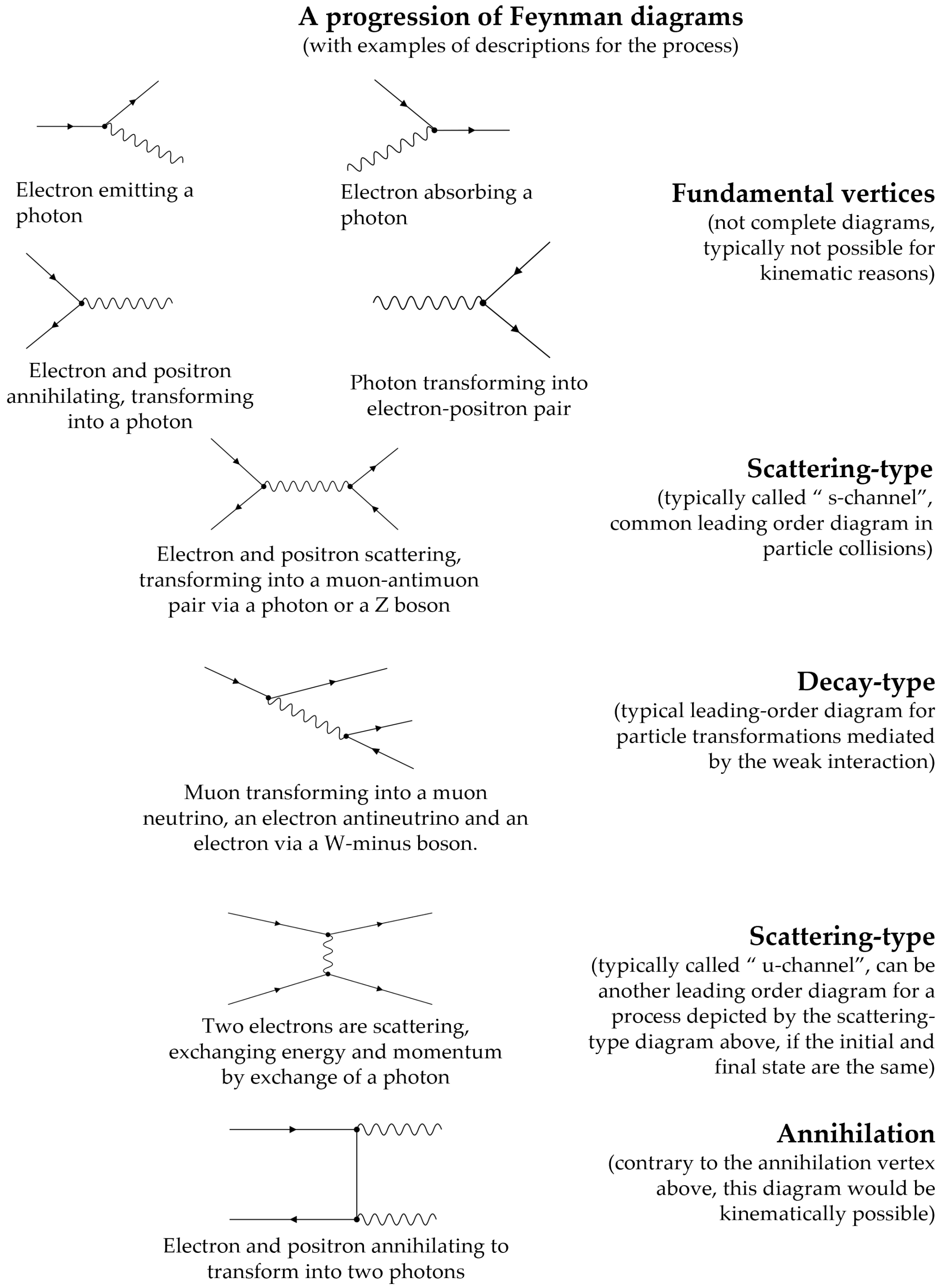


FIGURE 7.2: A proposed difficulty progression of Feynman diagrams. The texts below the diagrams describe a possible example for which the diagram could be used. The diagrams are ordered from top to bottom, increasing the difficulty level.

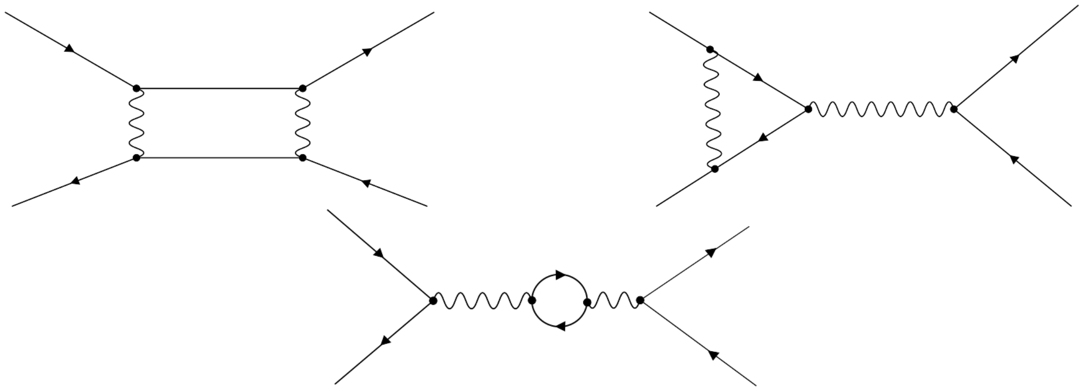


FIGURE 7.3: Three examples for second-order diagrams. The diagrams on the top have an additional interaction particle (e.g., a photon); the diagram at the bottom is a so-called loop diagram. The loop in the middle denotes a virtual electron-positron pair which annihilates directly after being created.

and develop an ease of reading it. As analysed in Section 6.5.2, students use different strategies when examining the diagrams. Through this change of perspective, they could gain a deeper understanding of the diagrams. This could be fostered, for example, by highlighting different "chunks" in different diagrams as visual cues for learners. All in all, through these strategies, learners get a good idea of which diagrams are allowed and which are not.

The second learning goal requires learners to develop a deeper understanding of the processes represented by the diagram. Since it is difficult to give a formal definition for what constitutes a particular fundamental interaction without introducing the mathematical apparatus[7], learners need to understand what kind of processes are mainly mediated by the different interactions to understand what the interactions are. Therefore, for this learning goal, learners should also get in contact with many different diagrams, as suggested by the learning material (cf. Section 6.6.3).

Pre-train the critical terms and definitions

As suggested by the pre-training principle (cf. Sections 4.3,5.1.2), learners should learn about the critical definitions and terms before they need them. That means that they should not only read about the terms but also get familiar with them. For the first learning goal, this means that they should get familiar with the two charge concepts, i.e., they should have some preparatory tasks where they are supposed to identify the electric and the weak charges of a particle; if they are not annotated before they use them to check a diagram for charge conservation.

For the second learning goal, they would need to get tasks to identify different parts of some diagrams to get familiar with the terms "matter particle", "anti-particle", and "interaction particle" and with some examples of the different categories (like

[7] A humoristic account of the dilemma of "explaining the four fundamental forces" was given by Munroe (2015a).

electron, muon, proton for matter particle, positron, anti-muon for anti-particle, or photon for interaction particle).

Possible confusions, like interaction particle vs. virtual particle or charge vs. energy, should also be directly addressed when introducing the terms so learners are aware of them.

7.3 Future directions of study

The studies in the current project were not designed to allow systematic comparisons between the different versions of the learning material. Instead, each study examined challenges and potential practices for teaching particle physics with Feynman diagrams.

Furthermore, they point out possible directions for future studies. In this section, I want to sketch out some potential future work regarding the role of Feynman diagrams in physics education.

7.3.1 Drawing diagrams

First, the studies presented in this thesis only investigated the perception of Feynman diagrams. It is, however, well established in education research that students learn better by creating a representation of themselves. In the taxonomy of learning (Anderson & Krathwohl, 2001; Mayer, 2002b), the cognitive processes subsumed under the category "create" comprise the category which is most likely to promote "meaningful learning". For an overview of the functions that drawing has in learning science, see Ainsworth et al. (2011). Therefore, performing studies where learners draw Feynman diagrams themselves would be desirable. This is expected to increase the essential processing while engaging in the activity. Many educational approaches presented in Section 2.2.2 already introduced Feynman diagrams in a "hands-on" approach, i.e., by a pen-and-paper activity, via card games or mechanical models. Some approaches could be evaluated regarding their learning effectiveness or motivational and affective factors.

7.3.2 Development of an assessment instrument

Furthermore, the studies presented in this work focused on teaching charge conservation with Feynman diagrams. Even though one part of the third student study was dedicated to the second learning goal, it was the main focus of the eye-tracking research. Future studies should investigate the second, third, and fourth learning goals.

An essential step towards this goal would be developing an assessment instrument for particle physics. Within the framework of this project, a full-scale development of an assessment instrument was out of scope. Still, on a smaller scale, two instruments were tested as prior knowledge instruments in the second and third student studies. The first was based on an instrument proposed by Tuzón and Solbes (2016) with some questions slightly rephrased, and I added one question. The results of the first use of the adapted instrument look promising in general (cf. instrument indicators summarised in Table 6.3), even though the newly added item did not perform well in several indicators (difficulty and discriminatory power). I tried to develop the instrument further by omitting some items and adding new ones. Still, the results of the use in the third student study reveal that this instrument did not work as well as the previous instrument (cf. Table 6.11). In future research, development should focus on the initial instrument and try to develop it further in line with literature suggestions of test instrument development (cf. Adams and Wieman (2011)), e.g., by conducting interviews with experts and students. However, as it was not the main scope of this project, the development was more focused on the learning material.

7.3.3 Limitations of using Eye Tracking in PER

As noted in Section 4.4, eye tracking has a successful history in PER. However, there are also limits to its usefulness. As stated in the introduction to Chapter 4, I tried to use eye tracking as the predominant research method of the DBR project. It was possible to make statements about students' cognitive processes when solving tasks with Feynman diagrams or their strategies, as stated in Sections 6.5.2 and 6.6.7. These results could draw some inferences regarding potential difficulties with the diagrams (cf. Section 7.1.2).

A possible research direction with eye tracking that has not been carried out during this project would be to investigate the expert's gaze on Feynman diagrams profoundly. The work presented in this thesis provides a starting point for which processes could be explored. For example, the issue pointed out in the pilot study and further investigated in H2 of the third student study is that experts, or students with minimal training, focus more on the relevant parts that could be examined. Also, the finding from H7 from the third student study that the triple transition ratio increases with training could be worth further examination to find representation-specific gaze patterns beneficial for learning with Feynman diagrams. These gaze patterns could be exploited by using representation-specific cueing, which guides students when examining Feynman diagrams.

Chapter 8

Further Directions

In this chapter, I present some ideas that could guide the development of future versions of the learning material, primarily focusing on exploiting the third and fourth learning goals, which have not been implemented yet.

8.1 Use of Feynman diagrams as calculational tools

The third learning goal highlights the mathematical nature of modern physics and the fact that Feynman diagrams are primarily used in particle physics as a calculational tool. This can be addressed in several ways.

8.1.1 Connection between Feynman diagrams and approximations

It is described in Section 6.2.3 how Feynman diagrams are used to approximate processes. A ubiquitous example used in physics is using a Taylor series to approximate the behaviour of a function near a certain point, for example, to approximate the function $\sin(x)$ near 0 as x by calculating its Taylor series until the second order. These and more complex examples could be used to motivate the use of approximations in physics, which are then compared to Feynman diagrams. Feynman diagrams offer an instance of approximations typical for modern science in that they are used in theoretical calculations in current research. Contrary to the often-painted picture that in contemporary science, there are no uncertainties, Feynman diagrams offer an example that also theoretical calculations are often, if not always, subject to uncertainty.

8.1.2 Connection between Feynman diagrams and programming

Feynman diagrams also offer the opportunity to show the relation between pen-and-paper estimations and calculations with a computer. Even though many physicists draw Feynman diagrams on blackboards or scribble them into notebooks by hand, calculations are usually carried out with a computer, where the rules of how to carry out the calculations are encoded into the diagram. For example, a program like *FeynGame* directly prints the probability amplitude for a given Feynman diagram and, through that, can help translate the diagrammatic into the mathematical representation. Harlander (2021) describes in his overview of Feynman diagrams how they are used in today's calculations: For higher-order calculations, it would not be practical anymore to draw all the diagrams (as an example, he states that the comparatively simple process $e^+e^- \rightarrow e^+e^-$ in the one-, two-, three, and four-loop level would yield 18, 186, 2264, and 31860 different diagrams, respectively). The calculation of all these diagrams is done mostly automatically. Of course, s.t. in practice, one often does not see the specific diagram anymore. It is, however, still necessary for particle physicists to have a deep understanding of the mechanics of the Feynman calculus to be able to solve problems which may occur in automatic higher-order calculations.

It would undoubtedly go too far for upper secondary students to introduce programs with which Feynman diagram calculations can be automated. Still, it would be desirable to convey the connection between the drawing of simple diagrams and extensive, automated calculations to convey aspects from the Nature of Science that creativity is still needed, even though many things are automated mainly.

8.2 Use of Feynman diagrams to interpret representations from particle physics

As pointed out by Harré (1988) and Kaiser (2005, p. 370 ff.), Feynman diagrams owe their success to their resemblance with the tracks particles leave in their detecting apparatus. It is, however, essential to note that Feynman diagrams are by no means some simplification of a process observed in particle detectors. Feynman diagrams can be used to interpret these observations.

The story of the discovery of the Z-boson sketched out in Section 6.2.4, offers the opportunity to show the connections between the diagrams and disciplinary depictions, which illustrate the different steps of the story. I will revisit this story and analyse the different disciplinary representations in the following.

8.2.1 Use of Feynman diagrams to interpret bubble chamber pictures

Figure 8.1 shows a Feynman diagram of a scattering process between a muon neutrino and an electron, which can only be mediated by a Z boson (because the interaction particle has to have zero electric charge, and neutrinos do not take part in the strong or electromagnetic interaction, therefore it cannot be a gluon or a photon). A process in which a muon neutrino interacts with an electron would, therefore, be evidence for the existence of a Z boson. This theoretical consideration, depicted by the Feynman diagram in Figure 8.1, led to the implementation of an experiment that sent neutrinos into the bubble chamber experiment *Gargamelle*. A bubble chamber is an early form of a particle detector in which ionising particles, such as electrons, can be detected. Bubble chambers contain a superheated liquid in which ions serve as vaporization cores around which tiny gas bubbles form. These gas bubbles then indicate where an ionising particle flew through. The picture on the right side of Figure 8.1 shows an event recorded in 1973, in which a neutrino beam entered from the right side and interacted with an electron, which eventually scattered at other particles (it is depicted in red in the picture).

By contrasting these two representations, it is clear that they don't necessarily look similar. Some interpretative work is needed to recognize the process depicted in the Feynman diagram within the bubble chamber picture. Woithe et al. (2019) have developed three educational activities in which students use these bubble chamber pictures[1] to first determine the charges of particles by analysing a track's curvature, then identify signatures of different particles, and finally analyse the transformations of particles. Feynman diagrams could accompany the latter activity so learners can match a diagram and a corresponding bubble chamber picture in which the process is observed.

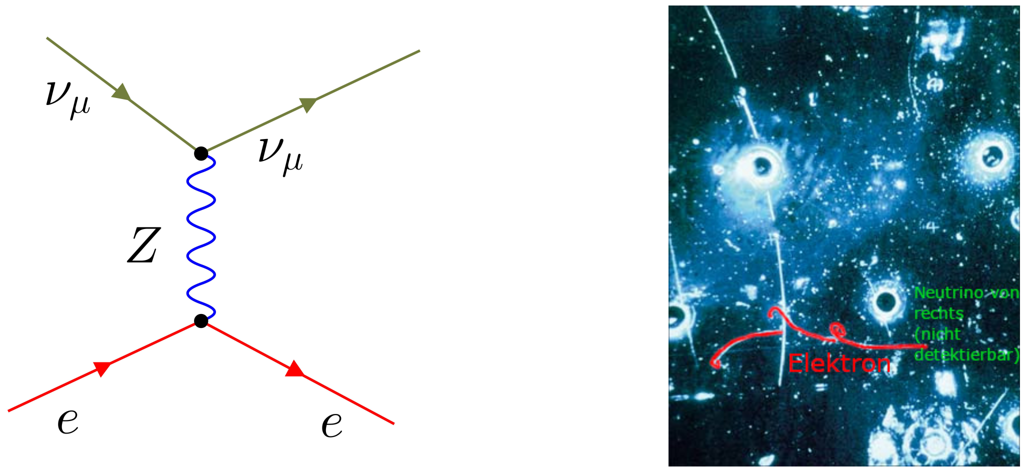


FIGURE 8.1: Representations of the first evidence of neutral currents found in the *Gargamelle* bubble chamber. The Feynman diagram on the left represents the electron neutrino scattering. The bubble chamber picture on the right shows how such a process shows up in a bubble chamber with a neutrino beam coming from the right, which is invisible to particle detectors interacting with an electron, which shows up as a curly line (highlighted in red in the picture). The right picture is taken from Cundy and Sutton (2009) with my annotations.

8.2.2 Use of Feynman diagrams to interpret event displays

Since bubble chambers are not used in usual modern-day particle physics experiments, bubble chamber pictures no longer represent current research. Instead, a typical representation of data from particle physics experiments is the so-called "event display", which is a visualisation of the data taken by a particle detector in a single collision. An example is shown on the left side of Figure 8.2. It depicts an event

[1]They use pictures from *Gargamelle*'s predecessor, the 2m Bubble Chamber. The pictures can be found in Bubble Chamber (1972).

in which an electron-positron pair is created in the UA1 detector at the $Sp\overline{p}S$ particle accelerator in 1983, where protons and antiprotons are accelerated and collided. The Feynman diagram on that figure's left side represents a process that could lead to such an event. One event, however, could have also been pure chance, as other processes exist in which an electron-positron pair is created (cf. the discussion in Section 2.1.3 on page 13 about the "background"). Therefore, it is only called a "discovery" if a statistical analysis of all the events of this kind concludes that the probability of the events being pure chance is lower than a certain threshold.

Event displays are commonly used in so-called particle physics "masterclasses" (Bilow & Cecire, 2021, 2022; Cecire et al., 2014; Johansson et al., 2007), in which students get to work hands-on with data from experiments to identify particles. Feynman diagrams accompany these activities in providing the theoretical background on which particle transformations to look out for.

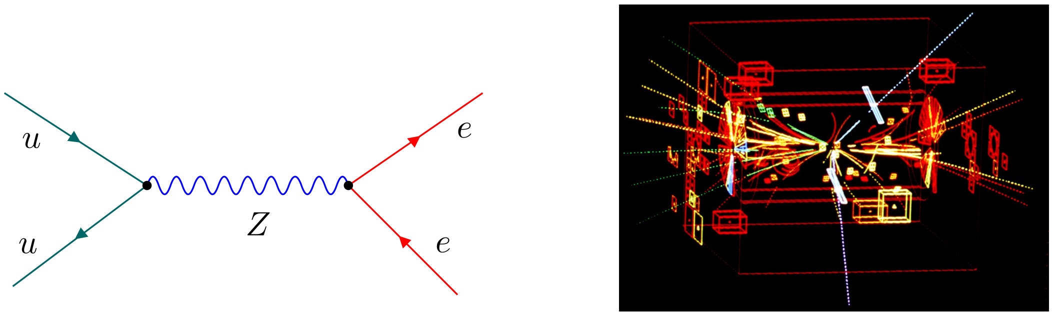


FIGURE 8.2: Representations of the first discovery of the Z boson at the UA1 experiment of the $Sp\overline{p}S$ collider. The Feynman diagram on the left shows the production mechanism in the proton-antiproton collider and the transformation from a quark-antiquark pair (in this example, an up and an anti-up, but it could be any quark) into an electron-positron pair. The event display from the UA1 experiment on the right shows how this transformation is in the experiment (the two white lines represent an electron and a positron). The right image is taken from CERN (1983).

8.2.3 Use of Feynman diagrams to interpret resonances

Though the Z boson was discovered in 1983, particle physicists continued measuring its properties, like its precise mass or couplings, to other particles. This was done, among others, in the experiments of the *LEP* particle accelerator, in which electrons and positrons were accelerated and collided. Z bosons were also produced in these collisions, which were subsequently transformed into different particles. On the top of Figure 8.3, three of such processes are depicted by their respective Feynman diagrams. The respective event displays of such events are shown on the bottom left of that figure. The plot on the bottom right side is especially interesting. This plot shows, to put it very simply, how equation 2.4 shows up in the data. Since the mass of the interaction particle is in the denominator of the probability amplitude, it diverges at the point where the total energy of the initial particles is equivalent to this mass. Here, a good connection can be made between the Feynman diagram as a picture, its translation into mathematical formulas, and actual data from particle physics experiments. A connection could also be made to resonance experiments, where a similar term shows up in the formulas to calculate the resonance frequency of a system. This was mentioned in the expert interview study in Section 6.1.

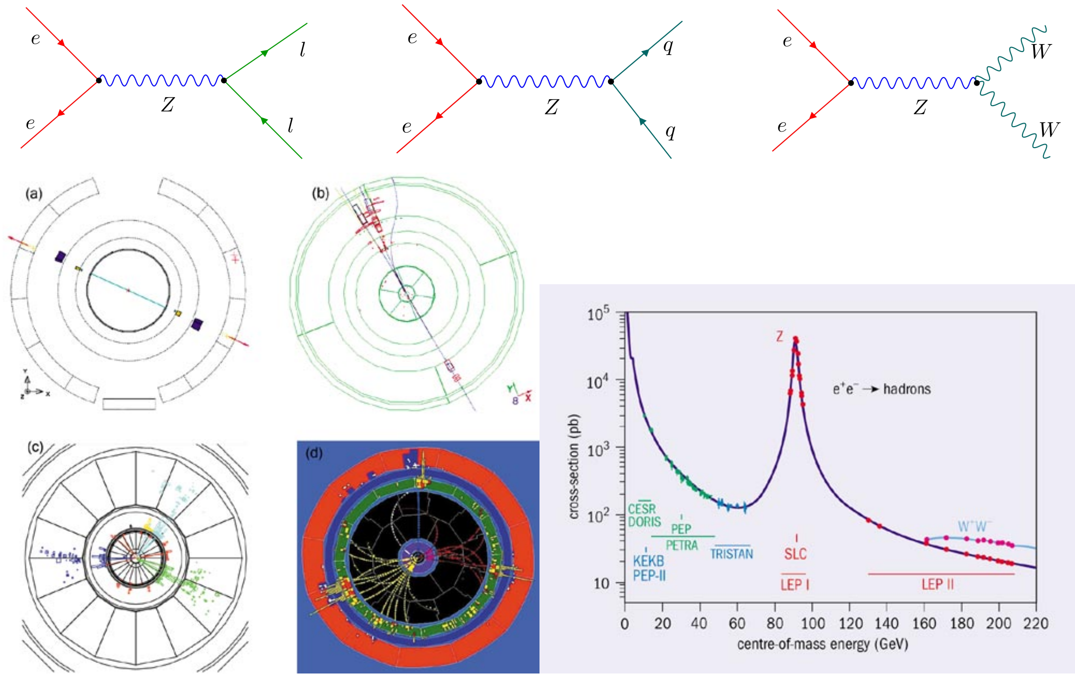


FIGURE 8.3: Representations for measuring the Z resonance at the LEP collider. The Feynman diagrams on top show possible diagrams for the Z transformation after being produced in an electron-positron collider. The bottom left image shows event displays of different experiments at LEP, showing how these transformations appear. The bottom right image shows a graph with the cross-section of the event $e^+e^- \rightarrow q\bar{q}$ depending on the energy with the peak showing the so-called Z resonance. The bottom images are taken from Sutton and Zerwas (2004).

8.3 Further examples

In this section, I mention some examples of particle physics lessons, including Feynman diagrams, for highly motivated and highly interested students who easily achieved all four learning goals presented in this work. If they want to go further, they could learn about the many intricacies of Feynman diagrams even before studying physics.

8.3.1 A diagram with high disciplinary affordance: The penguin

If students understand all the diagrams in the progression shown in Figure 7.2 and also the second-order diagrams as in Figure 7.3, the next step is not far to also be introduced to a diagram with a very high disciplinary affordance, as presented in Section 5.1.2, where the so-called *penguin diagram* (Figure 5.1) was introduced. The origin story of the name "penguin diagram" (Shifman, 1995) could also serve as a hook to get students' attention. However, the intricacies of the process represented by the diagram seem to be accessible only to the very highly motivated students.[2] The basic process, however, that a quark transforms into another type of quark, mediated by a W boson and a loop of another quark, could also be understood by students who had a previous lesson based on the learning material presented in this work.

8.3.2 The different stages of the Higgs boson discovery

One of the latest significant milestones in the development of the Standard Model of Particle Physics was the discovery of the Higgs particle at the *Large Hadron Collider* (LHC) announced on the 4th of July 2012 by Incandela et al. (2012), the (preliminary) highlight of a search started in the early 1960s. It culminated in the Nobel Prize in Physics for two of the physicists who invented the theoretical mechanism that predicted the particle that came to be known as the Higgs particle, François Englert and Peter Higgs (Nobel Prize Outreach AB, 2013). The story from the invention of the mechanism to the Nobel Prize has been told multiple times (a comprehensive version of the story can be found in Castillo (2014)) and is part of popular culture (see, e.g., Cendrowski (2012)). Connected to this story are many NoS aspects collected in a paper on the tenth anniversary of its discovery (Woithe et al., 2022). Closely linked to the story of its discovery are different Feynman diagrams, which can help in telling the story as they represent processes through which the Higgs boson is produced in the LHC and processes through which the Higgs boson transforms into other particles, which can then be detected, as the Higgs boson itself is too short-lived to be detected directly. As discussed in Section 2.1.3, Passon (2019) uses the discovery of the Higgs boson as a case study to argue that in particle detectors, it is never one Feynman diagram that is observed, but all Feynman diagrams which represent a process together. This discussion is part of the third learning goal.

8.3.3 Limitations of Feynman diagrams

An important NoS aspect is the significance of a model in physics and, with that, the limits of models. As Feynman diagrams are an essential part of the model primarily used in particle physics nowadays, it is worth exploring its limits. As is apparent from equation 2.3 (and also mentioned in Section 2.1.2 on page 11), the Feynman diagram method of calculating diagrams of ever higher-order could only converge if every vertex contributes a number smaller than one. This is given for the electromagnetic and the weak interactions but not necessarily for the strong interaction described by quantum chromodynamics (QCD). As also briefly mentioned in Section

[2] Tanedo (2012) tried to explain it for the highly interested and highly motivated reader without a particle physics background while he still builds on advanced concepts beyond this work's scope.

2.1.2, renormalisation leads to so-called "running coupling" for the strong interaction, which means that for high energies, the vertex factor of the strong interaction is indeed small enough. Hence, it works again (Griffiths, 2008, p. 301). When computing QCD problems at low energies, though, the perturbative method of Feynman diagrams breaks down, and another computation method has to be used. One standard method is the lattice gauge theory, which will not be further discussed here. However, the discussion around the third learning goal should address the principle of the limitations of Feynman diagrams.

Chapter 9

Conclusion

Within this PhD project, I designed and evaluated learning material focused on Feynman diagrams in the design-based research framework. Overall, this project highlighted the opportunities, investigated the challenges, and suggested some practices for using Feynman diagrams as educational tools in physics education. Though the results still leave room for further research, my conclusion of this project is a clear *yes* to the guiding question which was stated in Chapter 3: Feynman diagrams can be used for teaching particle physics concepts to upper secondary students, if the challenges that come with it are adequately addressed.

However, having asked whether the diagrams *could* be used, one question remains unanswered: *Should* it be done? Are the challenges worth the effort? Can the learning goals also be reached by other means?

Therefore, in this Chapter, I will revisit the opportunities and challenges discussed in Section 7.1. Yet, this time, considering the question posed above: *Should* we use Feynman diagrams for teaching particle physics concepts to upper secondary students?

9.1 Opportunities

First, I will revisit the educational purposes stated in Section 6.2 in the form of learning goals and argue why Feynman diagrams should be used to achieve them.

The first educational purpose is that Feynman diagrams can be a tool to convey the central concept of charge conservation. This purpose is well connected to the physics curriculum in topics other than particle physics, thus allowing teachers to connect particle physics with classical physics topics. Conservation laws are essential in physics and are introduced in different forms. Feynman diagrams offer an intuitive way to visualise the various particles which take part in a specific process and, with them, the other properties, in this case, charges. The concept could also be conveyed using simple formulas with the charges annotated as superscripts at different positions like common for isotope symbols (where it is the numbers of protons or neutrons in the nucleus, for example), in the case of the muon transformation for example as ${}^{-1/2}\mu^{-1} \rightarrow {}^{1/2}\nu_{\mu}^{0}\,{}^{-1/2}\overline{\nu_e}^{0}\,{}^{-1/2}e^{-1}$, but in this example, the charge annotations seem to be quite confusing and the intermediate step would be missing. Within this work, I have shown that the diagrams make it relatively easy to visualise even complex processes, and charge annotations can be placed if needed.

The second purpose reaches further into modern physics and presents Feynman diagrams as a tool to visualise the concept of interaction particles as a central concept in particle physics. The representative function of Feynman diagrams can be questioned in an epistemological discourse. However, in an educational sense, Feynman diagrams offer the opportunity to convey a modern concept of interactions in contrast to the Newtonian concept, where interactions are non-local. This would even make sense if particle physics is not explicitly part of the curriculum, as it shows how a concept as fundamental as the concept of an "interaction" evolves within history. If particle physics is taught as a topic, the concept of interaction particles is crucial. And even though technically, Feynman diagrams are not needed to introduce this concept (after all, the concept was introduced over a decade before the first Feynman diagram was published), it is hard to imagine nowadays any visualisation of the concept which does not, at least remotely, resemble a Feynman diagram.

The third purpose touches on the quantum mechanical nature of particle physics. Feynman diagrams can convey the importance of superposition and the concept of approximations by adding up ever-smaller contributions when dealing with particle physics, thus painting an accurate picture of modern science. There are less complicated examples of how the superposition in quantum mechanics can be explained. The most common example is the double-slit experiment, part of almost every quantum mechanics curriculum. In contrast, one of the most popular is undoubtedly the thought experiment of "Schrödinger's cat". These are critical examples and should be addressed in a quantum mechanics curriculum. Feynman diagrams, however, offer an opportunity to explain the quantum nature of particle physics. Particle physics is often "trivialised" in that particles are depicted as small balls without any quantum properties. Even though they are seemingly intuitive images, the method of calculations with Feynman diagrams offers the opportunity to paint a more accurate picture by introducing this quantum nature, thus connecting the public image of particle physics to the reality of modern science. The same can be said about using Feynman diagrams as approximations. A perturbative series using Feynman diagrams is among the most complicated examples of an infinite series (even without starting with renormalisation). Still, they offer an intuitive example of how this approximation principle is used in modern science. Thus, they offer a counter-narrative to the public image of science as being exact and unchangeable.

The fourth and final purpose synthesises all the abovementioned purposes and guides students to transfer the approach of Feynman diagrams to connect theoretical and experimental particle physics. Through that, students can learn about how modern particle physics and, thus, modern science works and obtain knowledge about (particle) physics and the Nature of Science. As pointed out throughout this work, NoS is a crucial part of science literacy, which should be addressed explicitly in the science classroom. While the method of Feynman diagrams cannot be addressed in its full mathematical detail, it does offer an opportunity to address the theory-laden NoS, the difference between empirical and inferential NoS, the tentative NoS, or the necessity of creativity in science explicitly.

9.2 Challenges and Practices

Even the most significant opportunities come with challenges. In the case of Feynman diagrams, some challenges are related to the terms and definitions, while others originate in the diagram's intricacies. This project uncovered some challenges discussed in Section 7.1.2 and suggested some practices discussed in Section 7.2. In this section, I will revisit them and argue why using the practices is worth overcoming the challenges.

The first challenge was a confusion of the concept of charge conservation. Students might not understand the concept of charge or confuse charge conservation with energy conservation. This might be an exciting learning opportunity for students as they learn that several quantities are conserved in a process. Even though the deep connection between conservation laws, symmetries and fundamental interactions via Noether's theorem could not be unravelled with most students, a sense of the importance of conservation laws should be conveyed as it lies at the core of physics. Therefore, in overcoming this challenge lies a key to the heart of physics.

The second challenge addresses the fact that the concept of interaction particles is a very subtle one. As mentioned in the previous section, however, it is worth conveying the idea as it is critical to a modern understanding of physics, which relies on local action to explain interactions as a model using "action at a distance", like that of Coulomb force is not following special relativity. Students do not need to understand all the intricacies of interaction particles, certainly not the mathematical ones. It might be worth it, however, to convey some key messages, like the existence of different interactions and the fact that these have different interaction particles that define the interactions' properties. The students' statements in Section 6.6.8 show that it is insufficient to show Feynman diagrams and mention the interaction particle to convey this message. It needs explanations and examples of the different interactions and practice over time to get familiar with a new concept.

The third challenge is stated as a category and concerns the diagram's elements, particularly the issues of non-sequentially ordered vertices and arrows. The problem of arrows was addressed in the first version of the learning material, in which only one direction or no arrows were used. This, however, was eventually abandoned in the update of the learning material, as the diagrams needed to be connectable to later types of diagrams. A concluding answer for the issue of how arrows in Feynman diagrams should be addressed cannot be given within this project's scope. As for intricate diagrams, the results of this project suggest a learning progression of Feynman diagrams presented in Figures 7.2 and 7.3. Also, they have shown that varying strategies when examining the diagrams are beneficial. These suggestions benefit a better understanding of Feynman diagrams and better representational competencies in general. Training with Feynman diagrams trains the perceptual fluency needed to grasp information quickly from a diagram, which might also benefit science learning in general.

The last challenge discussed in this work is the issue of many new terms that get easily confused (cf. Section 6.6.8). A possible solution for this issue would be pre-training the concepts as suggested by Mayer (cf. Section 4.3). As before, a deeper understanding of the topics comes only with practice. However, it is hard to argue why students should be taught terms that are distinctive to particle physics rather than just to understand particle physics. Here, I am back to the beginning of this Chapter. Students who are interested in particle physics will very likely come across terms like "anti-particle", "photon", or "W-boson" and hence want to know what that is. Learning about them can be a spark of curiosity, especially if it is combined with interesting facts (for example, "CERN is the only place in the universe where anti-particles are combined to form antimatter" or "the high mass of the W-boson is the reason why the sun still shines").

9.3 Final Remarks

This was my trial to explain Feynman diagrams. I hope that I could convince you, dear reader, of the usefulness of Feynman diagrams for (particle) physics education. Even though Feynman diagrams are highly specific to particle physics, there are many exciting and important aspects related to them that are worth addressing in the physics classroom, if not as part of the standard curriculum, as extracurricular learning material for interested students.

Appendix A

Declaration on the use of generative AI

In this thesis, I have used `ChatGPT 3.5` and `Google Gemini` as follows.:

- To create a first draft of individual passages, namely `ChatGPT` and `Gemini`:
  - to create a first draft of the abstract based on input by Chapter 1, 3, Section 5.1.1, and Chapter 9,
  - to create a first draft for the list items in Section 2.2.3 based on input of the literature which is cited in that section,
  - to create a first draft for the explanations of definitions introduced in Section 4.3 based on input of literature which is cited in that section, and
  - to create a draft for the definitions in Section 5.4.2 by asking `Gemini` to create a summary of the named concepts.

  All created texts were thoroughly reviewed and majorly altered by the author.
- For optimising or restructuring software source texts, namely `ChatGPT` to draft single lines of code to help with structuring and analysing data.
- For proofreading or optimising the text of the whole thesis, I used the suggestions by `Grammarly`.

I hereby declare that I have stated all uses completely.

Appendix B

Questionnaires

In this appendix, the questionnaires used in the three student studies and described in Section 5.3 are presented.

B.1 Prior Knowledge

Question	Question in TS	Rationale for including the question	Coding scheme
Which kinds of fundamental particles are there?	Q2c	The learning material covers different particles, such as electrons, but also muons or neutrinos, which are not part of the standard curriculum.	1: Mention of more than two categories of particles 0.5: Mention of one or two categories (like quarks, electrons, neutrinos)
What are the fundamental forces in nature?	Q1b	The learning material covers the electromagnetic and weak interaction, of which the latter is not necessarily part of the standard curriculum.	1: Gravity, electromagnetic, strong, and weak 0.5: Parts of the above (or synonyms)
Which kind of interaction occurs between the electron and the atomic nucleus?	Q1c	The learning material covers the electromagnetic interaction.	1: Electric or electromagnetic 0: any other (like simply "attraction")
Which kind of interaction occurs when a nucleus transforms into each other?	Q2b	The learning material covers the weak interaction.	1: Mention of weak interaction/force 0.5: mention of radioactive decay
What is a photon?	Q1d	The learning material introduces the photon as the interaction particle of the electromagnetic interaction.	1: "force carrier" of electromagnetic interaction or similar 0.5: particle/wave of light or similar
What are the differences between matter and antimatter?	Q3c	The learning material covers antiparticles.	1: matter with opposite charges 0.5: simply mentioning "opposite" w/o charges
What happens in a particle collision?	Q3e	The learning material is motivated by processes in a particle accelerator.	1: creation of particles due to release of energy 0.5: only one part of the answer above 0: answers like "splitting or fusion" of particles
Which path does a particle take between two points in space?	None	Assessing students' familiarity with the superposition principle, as the learning material covers this concept.	1: mentioning of any path or superposition 0: any other answer, e.g., direct path
What is CERN?	Q3f	The learning material is put in the context of CERN's research.	1: mention of research center and accelerator 0.5: only mention of accelerator

TABLE B.1: Adaption of the prior knowledge questionnaire from Tuzón and Solbes (2016) (TS). The second column notes which question from their questionnaire the question refers to. The last column describes the coding of the answers, based on the coding in the original study.

Question	Answer options	Rationale for including the question
What types of particles have you heard of?	MC: protons; neutrons; electrons; muons; quarks; neutrinos; Higgs-particle; photons	The learning material covers different types of particles.
Which of the statements about energy is true?	SC: Energy is a fuel that enables movement.; **Energy is always conserved in a closed system.**; Energy is always lost in a mechanical process. Electrical devices consume energy.	The learning material covers conservation laws; students might already know about energy conservation.
Which of the following particles is an interaction particle?	SC: **photon**; quark; muon; neutrino; I don't know what interaction particles are.	The learning material covers the concept of interaction particles. The question assesses whether students have heard about that concept before.
Which of the following particles has the smallest mass?	SC: electron; neutrino; **photon**; up quark; The options don't mean anything to me, and I would have to guess.	The question assesses whether students know more about the different particles than just their names.
What is Schrödinger's cat?	SC: A real cat that was used in an experiment.; **A thought experiment to explain the superposition principle.**; A theory about the behaviour of animals in quantum mechanics.; A particle in particle physics.; I don't know the term, and I would have to guess.	As the learning material covers the superposition principle in quantum mechanics, the question assesses whether students have heard about a famous thought experiment that illustrates this principle.
What is CERN?	(OC)	The learning material is put in the context of CERN's research.

TABLE B.2: Development of a new assessment instrument to assess prior knowledge in the third student study. If a question was given as a multiple choice question, this is indicated as (MC: [answer options]), if it was a single choice question, it is indicated as (SC: [answer options]) with the correct option in bold, if it was given as an open question, it was indicated as (OC). In the calculation of the final score, the first question only gives one point if all options are chosen, $2/3$ if 4-7 options are chosen, $1/3$ if 1-3 options are chosen, and 0 points if less than three are chosen. Questions 2-5 give 1 point for the correct answer and 0 points for the incorrect or no answer, the last question is not taken into account.

B.2 Cognititve Load

Question in first student study	Question in second student study	Correspondence
The topic of charge conservation is very complex.	The topics covered in the activity were very complex.	1, IL
I perceived the diagrams as very complex.	The activity covered diagrams that I perceived as very complex.	2, IL
I perceived the concepts and definitions (particle names, charges) as very complex.	The activity covered concepts and definitions that I perceived as very complex.	3, IL
The explanations during the activity were very unclear.	The explanations during the activity were very unclear.	4, EL
I did not learn a lot from the explanations.	I did not learn a lot from the explanations.	5, EL
I had to read the explanations more than once to understand them.	The explanations during the activity were full of unclear language.	6, EL
I understand more about the topic after completing the activity.	The activity really enhanced my understanding of the topics covered.	7, GL
I understand more about the diagrams from answering the questions.	The activity really enhanced my understanding of the diagrams shown in the activity.	9, GL
I understand more about the concept of charge conservation from answering the questions.	The activity really enhanced my understanding of concepts and definitions.	10, GL

TABLE B.3: Cognitive load questionnaires in the first and second student studies. The third column shows which item of Leppink et al. (2013) the question refers to alongside the dimension of cognitive load – extrinsic load (EL), intrinsic load (IL), or germane load (GL) (cf. Section 4.3) – which is assessed.

German	Original	Kind of CL
Das Thema des Materials war sehr komplex.	The topics covered in the e-text were very complex.	Intrinsic
Die Diagramme habe ich als sehr komplex empfunden.	The e-text covered content that I perceived as very complex.	Intrinsic
Ich habe die Beispiele und Konzepte als sehr komplex empfunden.	The e-text included very complex case studies and concepts.	Intrinsic
Ich fand es sehr mühsam, die Komplexität des Materials zu durchdringen.	I invested a very high mental effort in the complexity of the e-text content.	Intrinsic
Die Interaktion mit dem Material (Anklicken von Knöpfen, Alpakas usw.) war sehr ablenkend.	Manipulating e-texts was very distracting.	Extrinsic
Ich fand es nicht zielführend für meinen Lernfortschritt, mit dem Material zu interagieren.	Manipulating e-texts was, in terms of learning, very ineffective.	Extrinsic
Ich fand es sehr mühsam, mit dem Material zu interagieren (Anklicken von Knöpfen, Alpakas usw.)	I invested a very high mental effort in ineffective etext manipulation.	Extrinsic
Texte auf dem Bildschirm zu lesen war nicht zielführend für meinen Lernfortschritt.	Reading e-texts off a screen was, in terms of learning, very ineffective.	Extrinsic
Ich fand es sehr mühsam, Texte auf dem Bildschirm zu lesen.	I invested a very high mental effort in reading e-texts off a screen.	Extrinsic
Das Material hat mein Verständnis des Themas verbessert.	This activity really enhanced my understanding of the content that was covered.	Germane
Das Material hat mein Verständnis der Diagramme verbessert.	This activity really enhanced my understanding of the problem/s that was/ were covered.	Germane
Das Material hat mein Verständnis der Beispiele und Konzepte verbessert.	This activity really enhanced my knowledge of the terms that were mentioned.	Germane

TABLE B.4: Adapted cognitive load questionnaire used in the third student study, based on Leppink et al. (2014) and adapted for e-learning material by Novak et al. (2018). The second column shows the corresponding item in the original questionnaire. The items refer to different kinds of cognitive load which are indicated in the third column and were answered on a 5-point Likert scale from 1 ("Completely disagree") to 5 ("Completely agree").

B.3 Instructional motivational survey

German original	Original	Dimension
Die Art und Weise, wie der Inhalt des Lernmaterials vorgestellt wurde, hat meine Aufmerksamkeit geweckt.	The way that the content of the e-text was introduced got my attention.	A
Das Lernmaterial hat mich neugierig gemacht.	The e-text made me curious.	A
Das Lernmaterial war langweilig.	The e-text was boring.	A (-)
Es war mir wichtig, die Lektüre des Lernmaterials erfolgreich zu beenden.	Finishing the e-text readings successfully was important to me.	R
Ich kann erkennen, wie der Inhalt des Lernmaterials mit Dingen zusammenhängt, über die ich bereits Bescheid weiß.	I can see how the content of the e-text is related to things I already know about.	R
Das Wissen im Lernmaterial ist für mich NICHT nützlich.	The knowledge in the e-text is NOT useful to me.	R (-)
Die Art und Weise, wie die Informationen organisiert waren, gab mir das Gefühl, dass ich mit dem Lernmaterial erfolgreich arbeiten könnte.	The way that the information was organized made me feel that I could be successful in working with the e-text.	C
Ich konnte den größten Teil des Inhalts des Lernmaterials verstehen.	I could understand most of the content of the e-text.	C
Nachdem ich eine Weile mit dem Lernmaterial gearbeitet hatte, fühlte ich mich sicher in dem, was ich tat.	After I worked with the e-text for a while, I felt confident about what I was doing.	C
Ich habe mich gefreut, dass ich das Lernmaterial erfolgreich durchgearbeitet habe.	I was happy about finishing the e-text successfully.	S
Ich fand das Lernmaterial NICHT gut gestaltet.	I did NOT think the e-text was well designed.	S (-)
Ich habe gerne mit dem Lernmaterial gearbeitet.	I liked working with the e-text.	S
Das Material war schwieriger zu verstehen, als ich es mir gewünscht hätte.	This material was more difficult to understand than I would like for it to be.	C (-)
Das Lernmaterial enthielt Dinge, die meine Neugierde anregten.	The e-text had things that stimulated my curiosity.	A
Ich konnte den Inhalt des Lernmaterials mit Dingen in Verbindung bringen, die ich in meinem eigenen Leben gesehen, getan oder darüber nachgedacht habe.	I could relate the e-text content to things I have seen, done, or thought about in my own life.	R
Es war ein Vergnügen, mit so gut gestaltetem Lernmaterial zu arbeiten.	It was a pleasure to work with such a well-designed e-text.	S

TABLE B.5: Instructional motivation questionnaire, taken from Novak et al. (2022). The first column shows the items as they were used in the third student study, and the second column shows the original items used by Novak et al. The items were answered on a 5-point Likert scale from 1 ("Completely disagree") to 5 ("Completely agree"). The items refer to the four dimensions of the ARCS model of motivation: (A)ttention, (R)elevance, (C)onfidence, and (S)atisfaction which are indicated in the third column. In four cases, the scale was inverted during analysis. These are indicated by a (-).

Appendix C

Supplemental material of the expert study

In this appendix, the supplemental material to the expert study, as presented in Section 6.1 and published in Dahlkemper et al. (2022) is presented.

C.1 Interview guidelines

The notation is as follows:

Topic of the interview section

- Question
  - (if applicable) Follow-up question

C.1.1 Interview-Guideline used with E1, E2, E3

What do Feynman diagrams represent?

- Explain in your own words what a Feynman diagram (FD) is.
- Let's assume a high school student, 17 years old, with little previous knowledge of particle physics, is completing a course unit on the topic of FD. She is familiar with the conceptual distinction between interaction particles and elementary particles. What message should be conveyed in any case?
  - How does this message change with younger students?
  - What about university students?
- What do FD show us?
  - And now from the perspective of young people (high school, 17 years, no previous knowledge of particle physics): What do they show them?

What are potential learning difficulties in the context of FD?

- What concepts do teenagers need to know in order to read a Feynman diagram?
  - What additional concepts do they need to master if they are to draw one?
- We have now talked about what a Feynman diagram is. What is a FD not?
- If you have experience of working with pupils: Are you aware of any ideas or learning difficulties about FD that learners bring with them?
  - If you have no experience: Can you think of any ideas that young people might have?
- What visual elements should be given special attention?
  - If visual cues should be placed in an instruction, where would it be?
- What terms do experts use in the linguistic description of FD?
  - Which of them can be used with students? Which ones should be rephrased, omitted or even avoided?

Which potentials do Feynman diagrams offer for the classroom?

- Should Feynman diagrams be used in physics lessons?
  - If so, why? In what way should they be introduced?
  - If no, what stands in the way?
- When should FD be introduced as part of a learning unit on particle physics?
- Suppose the young person described earlier is to check whether a FD is theoretically possible or not. What prior knowledge should be provided and what would be the steps of an instruction for this task?
  - How does it differ for students and experts?
- Can students benefit from the FD instruction beyond the particle physics context?
- How would you explain this specific FD to the already known high school student (17 years old, little prior knowledge in particle physics)?

C.1.2 Interview guideline used with E4

Clarification of terms

Before we start with the actual questions, I would like to clarify a few terms in order to come to a common language. There are different definitions for many terms and it is about your understanding of them, just so that we don't talk past each other.

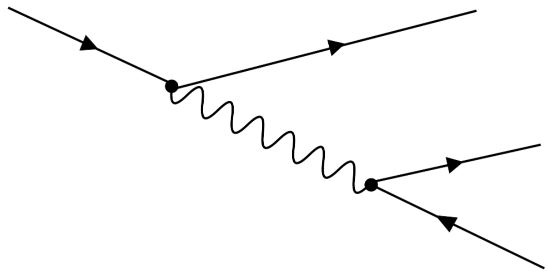


FIGURE C.1: Feynman diagram shown to experts E1-E3.

- What is your understanding of a model?
- What is your understanding of a theory?
- What is your understanding of reality?
- What is the standard model of particle physics?

The Connection between models and reality

In the following, I would like to ask some questions about how models represent reality and what consequences this has for the representation of physics.

- To what extent do models represent reality?
  - In the sense of: Models are never an exact representation of reality, but how close can we get to reality?
- More specifically, to what extent does the Standard Model of particle physics describe reality?
  - What are its limitations?
  - To what extent can it not describe reality in principle?
- What role do mathematical descriptions of a model play?
  - Are they an extension of our language for describing the models?
  - Is it possible to describe physical models without mathematics?
- What role do graphical approximations play in models? To give two examples of what I mean by this term: Field lines, for example, are graphical approximations of fields. Feynman diagrams are graphical approximations of particle processes.
  - If we describe a model with the help of a graphical approximation, to what extent does this graphical approximation describe reality?

Feynman diagrams

I would now like to ask some questions specifically about Feynman diagrams, knowing full well that this is not the core area of your expertise.

- What is a Feynman diagram?
- What do FDs show us?
- What is an FD not?
- What concepts do you need to know to read an FD?
  - What are the possible misconceptions involved?
- Should FDs be used in physics lessons?
  - Why or why not?

Modern Physics in school

Finally, I have some questions about whether and how modern physics and epistemological aspects should be taught in physics classes.

- Which aspects of particle physics should be taught in school and why?
  - What are possible opportunities?
- Should philosophical aspects of physics (e.g. epistemology) be addressed in physics lessons?
  - What are the opportunities?
- What are the difficulties involved?
  - Which aspects of the Nature of Science do you consider crucial for teaching in physics classes and why?

C.2 Coding manual

TABLE C.1: Coding manual of the expert study

Code	Code description	Instruction	Suggestions
C	Challenges		
C.1	FDs elicit and perpetuate inadequate conceptions	The segment mentions points where the use of Feynman diagrams could inhibit a deeper understanding by fostering wrong or misleading conceptions.	
C.1.1	Types of inadequate conceptions connected to FD		
C.1.1.1	Particle processes are embedded in spacetime	The segment mentions the space-time embedding of a FD. Use this code if notions of space and time are mentioned, e.g. by describing something as physical trajectories and also if the FD is described as a space-time diagram.	If the axes of the diagram (space and/or time) are mentioned explicitly → C.1.2.2
C.1.1.2	FDs show observable processes	The segment mentions the (mis)conception of one diagram as a depiction of a physical process. Use this code if the connection between one diagram and a physical process is mentioned (i.e. either as the misconception or an explicit caution that a diagram is not depicted by one single diagram).	If it is mentioned as an opportunity that FDs connect to quantum mechanics via different processes → O.1.2

continued on the next page

TABLE C.1: Coding manual of the expert study

Code	Code description	Instruction	Suggestions
C.1.1.3	Particles are small balls	The segment mentions naïve particle conceptions, like particles as small balls or the exchange of particles as physical objects.	If exchange particles are mentioned without the explicit mentioning of exchanging → O.2.2
C.1.1.4	Focus on the concept of "building blocks" and neglect of the concept of "interaction"	The segment mentions particles and their superficial treatment or the reductionist approach to particle physics.	If the concrete conception is mentioned that they are small balls or if the conception is mentioned explicitly together with FDs → C.1.1.3
C.1.2	Potential sources of inadequate conceptions		
C.1.2.1	Use of scientific language is a source of misconceptions	The Diagram or a part of it is explicitly mentioned as part of the scientific language or the scientific jargon.	
C.1.2.2	Axes of the FDs are misleading	The segment mentions axes of the diagram as a source of misunderstanding.	If the segment mentions the FD as a space-time-diagram but not axes explicitly → C.1.1.1
C.2	Particle physics can only be treated superficially	The segment mentions points where a superficial treatment inhibits a deeper understanding.	
C.2.1	Limitations by educational setting		
C.2.1.1	The time used for pp could be used otherwise	The segment mentions the time use in physics class.	
C.2.1.2	Necessary prior knowledge is missing in school	The segment mentions that it is difficult to speak about Feynman diagrams because the necessary prior knowledge is missing.	

continued on the next page

TABLE C.1: Coding manual of the expert study

Code	Code description	Instruction	Suggestions
C.2.1.3	It is a challenge for teachers to teach modern physics	The segment mentions how teachers handle concepts of modern physics or modern physics in general in school.	If the particular problem of time usage is mentioned → C.2.1.1 If missing prior knowledge is mentioned → C.2.1.2
C.2.2	The disciplinary handling of FD is not taught at schools		
C.2.2.1	Calculations might be too difficult	The segment mentions calculations as too difficult for doing in school.	
C.2.2.2	Some connected concepts too difficult for school	The segment explicitly mentions concepts which can not be discussed in school.	
C.2.2.3	Drawing FDs	The segment mentions that drawing might be more difficult than interpreting a FD	
O	Opportunities		
O.1	FDs offer a link between pp and high school teaching	The segment mentions how Feynman diagrams could connect particle physics to other topics in high school physics.	
O.1.1	FDs are suited to teach about conservation laws.	Referring to opportunities where either conservation laws could be taught through FD or FD could be explained by conservation laws.	
O.1.2	FDs link pp and quantum mechanics	The segment mentions that FD can connect to quantum mechanics.	

continued on the next page

TABLE C.1: Coding manual of the expert study

Code	Code description	Instruction	Suggestions
O.1.3	FDs offer an insight into the use of structurally equivalent representations	The segment mentions electrical circuits as an analogy or a depiction with similarities to Feynman diagrams. Also: the segment mentions concepts from electricity as an analogy.	
O.1.4	Analogy between FDs and resonance phenomena in classical oscillations	The segment mentions forced oscillations as an analogy or similarities between this concept and FD.	
O.2	FD offer an opportunity for different pp topics to be taught		
O.2.1	Outer and inner lines / Virtual particles	The segment mentions outer and inner lines and/or virtual particles as opportunity. Also code here if the segment mentions the time ordering of inner vertices	
O.2.2	Introduction of Interaction particles	The segment mentions FDs as opportunity to introduce interaction particles.	
O.2.3	Suggestions to practice the handling of FDs	The segment mentions tools or suggestions with or how Feynman diagrams can be handled by high-school students. Also code here for concrete examples of how calculations work for students	If calculations mentioned as being too difficult → C.2.2.1

continued on the next page

TABLE C.1: Coding manual of the expert study

Code	Code description	Instruction	Suggestions
O.2.4	Particle types	The segment mentions different types of particles.	If FDs are mentioned as a motivation for interaction particles → O.2.2
O.2.5	Introduction of pair production and annihilation	The segment mentions FD as explicit motivation for pair production and/or annihilation.	
O.3	FDs offer a connection to current research	The segment mentions how Feynman diagrams can offer a perspective	
O.3.1	FDs help scientists discussing particle processes	The segment mentions how Feynman diagrams are helpful for scientists.	
O.3.2	PP is a showcase for modern science	The segment mentions that particle physics can be used to show how modern science works in general.	If the daily work of scientists is mentioned → O.3.1
O.3.3	Students find FD in popular scientific representations	The segment mentions popular scientific representations.	

Appendix D

Supplemental material for the first and second student study

In this appendix chapter, I present the final draft of the learning material used in the first and second student studies. The material was altered during the first study, but it was essentially used like this at the end of the first study. The details of the design process are described in Sections 5.1.2, 6.4.2, and 6.5.1. The versions of this learning material used at the end of the first student study and in the second student study are available via this link: https://doi.org/10.5281/zenodo.11501410.

Furthermore, the answers to questionnaires as well as additional heat maps are available via the following link: https://doi.org/10.5281/zenodo.11505981. The questionnaire answers were partly analysed using the `R` programming language using scripts which is available under this link: https://doi.org/10.5281/zenodo.11519651.

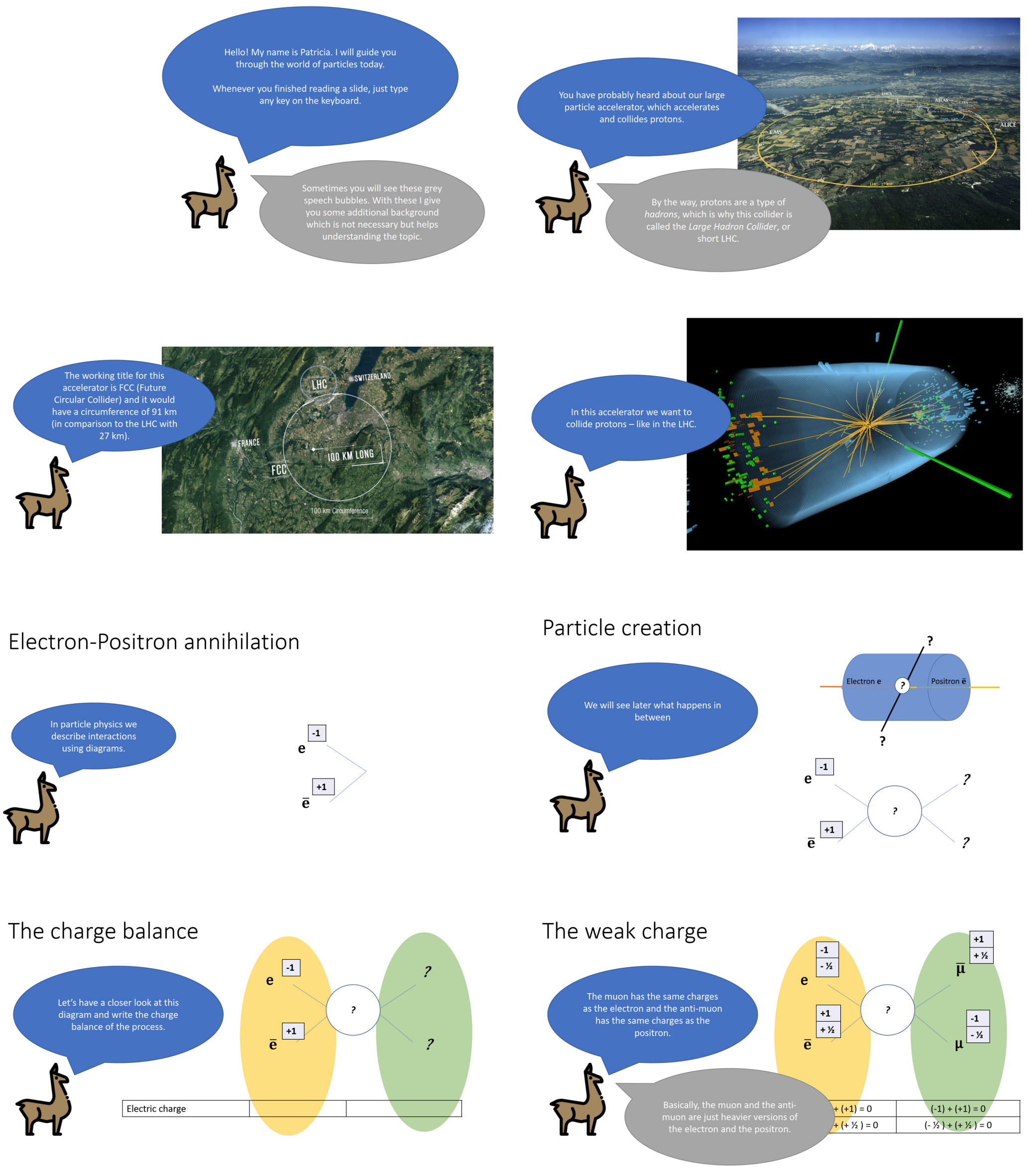


FIGURE D.1: Example slides from the introduction and explanation of charge conservation in the version of the learning material used in the first student study.

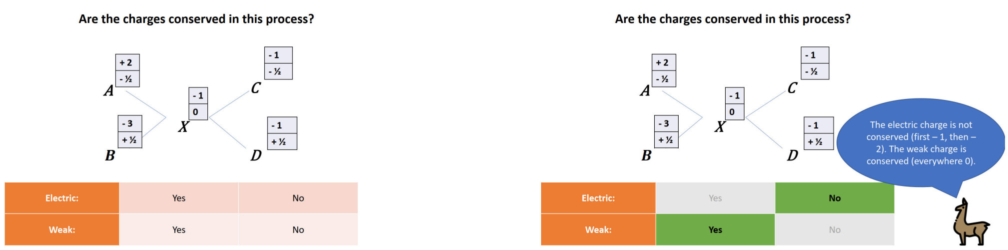
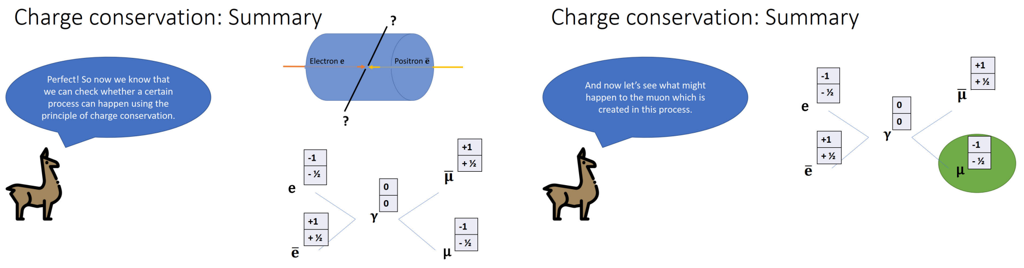


FIGURE D.2: First task and its solution from the first and second student study.



First summary slides of charge conservation.

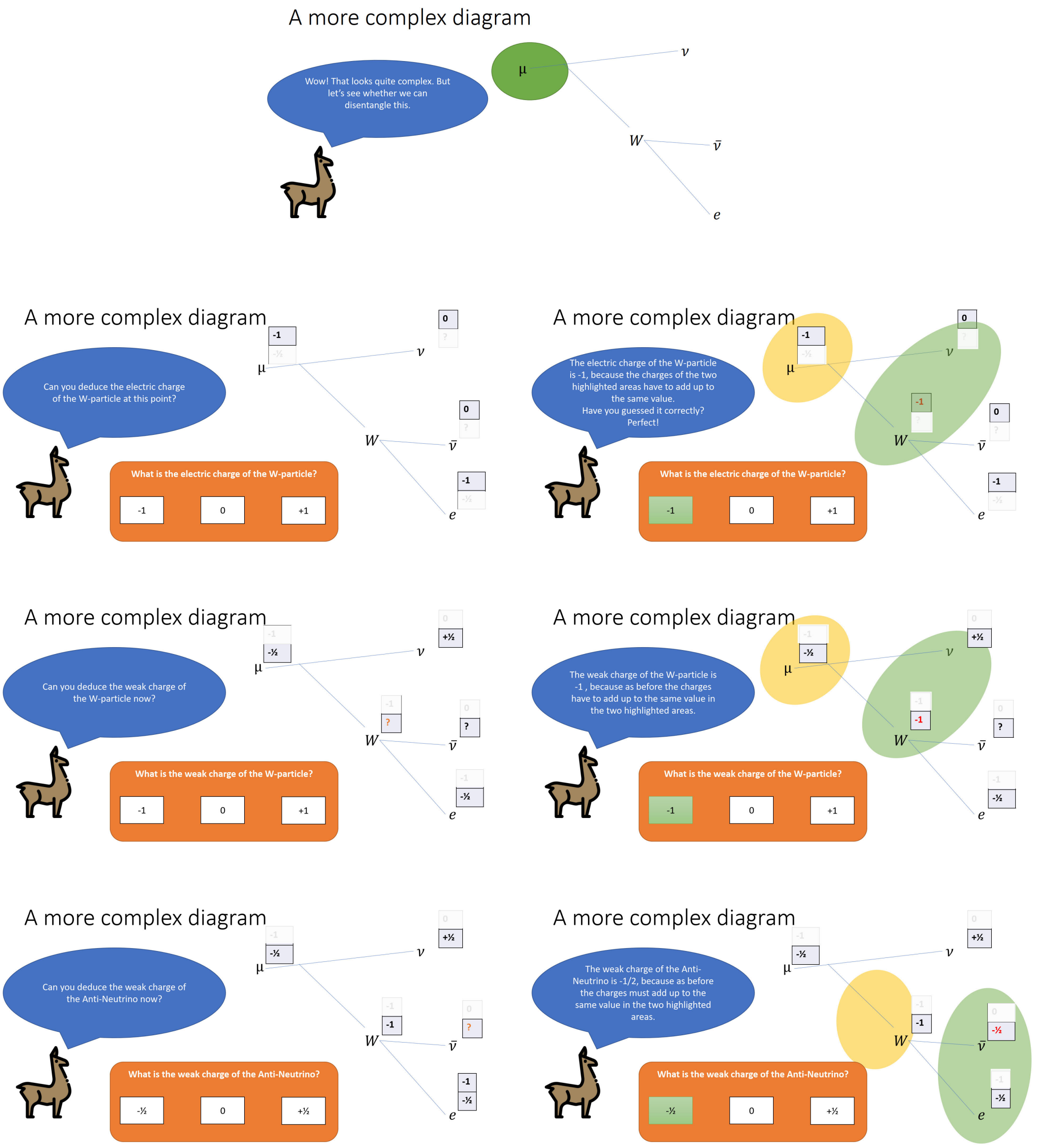
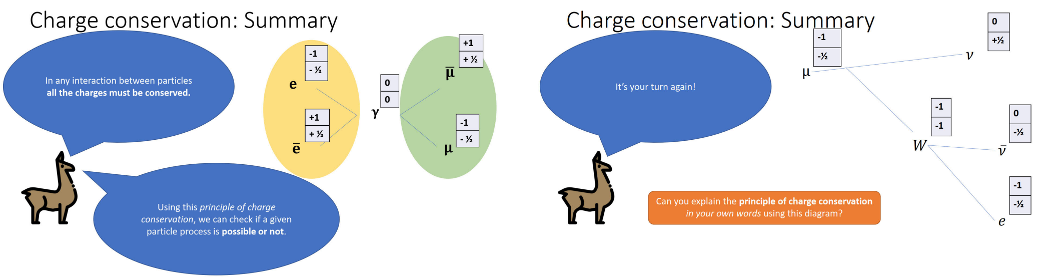


FIGURE D.3: Tasks used in the first and second student study about the muon transformation.



Second summary slides of charge conservation.

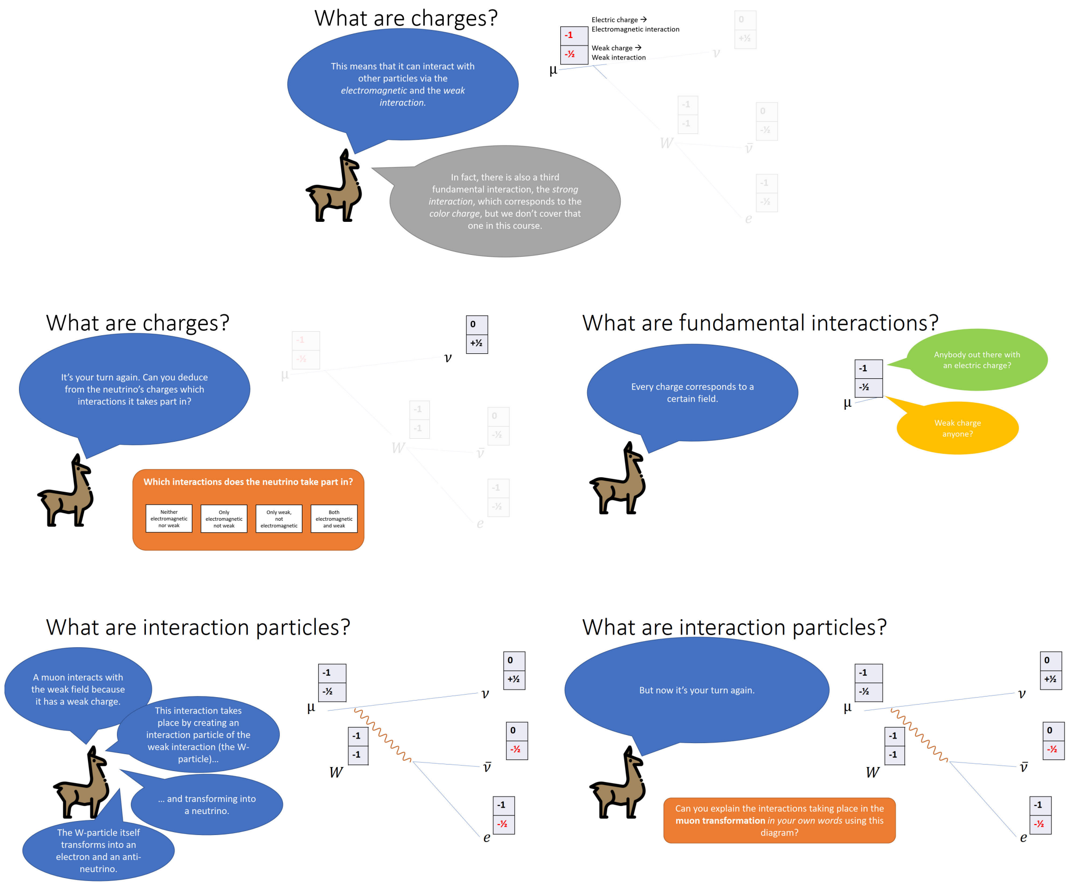
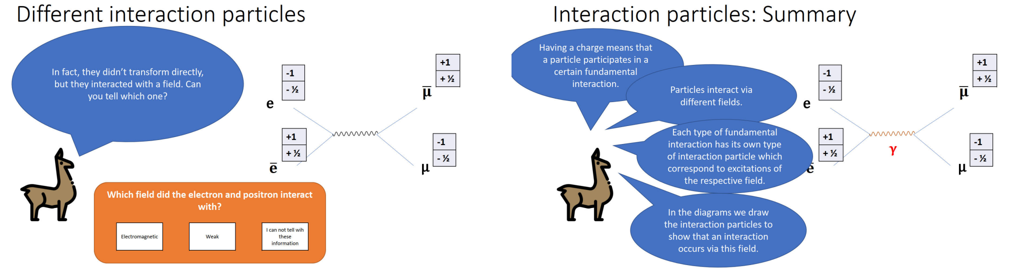
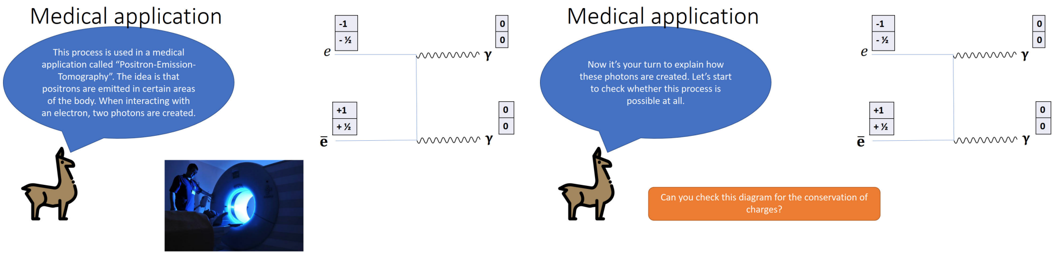


FIGURE D.4: Example slides from the learning material in the second student study to include the second learning goal into the material.



Example slides from the summary of the second learning goal



Example slides from the transfer task using an additional example from a medical application (PET scan)

Appendix E

Supplemental material of the third student study

In this chapter I present the learning material used in the third student study as presented in Section 6.6.3. The first part about charge conservation is presented as it was used in the study, while the second part was interactive and thus cannot be presented in full. It is available as open access material under this link: https://doi.org/10.5281/zenodo.11486241.

The eye tracking data and questionnaire answers obtained in this study are published under the following link: https://doi.org/10.5281/zenodo.11503036. The data were analysed using the R programming language. The scripts are available via the following link: https://doi.org/10.5281/zenodo.11519651.


Alpacarticle Physics mit Feynman-Diagrammen
Wie Teilcheninteraktionen die Welt beherrschen von Tieren bis zu Teilchenkollisionen.
START
Introduction
Hier kannst du jederzeit zur vorhergehenden Folie springen.
Einleitung
Mit einem Klick auf den Knopf in der Ecke findest du alle interaktiven Elemente der Folie.
Am Ende dieses Kurses solltest du ein grundlegendes Verständnis davon haben, wie Teilchen miteinander wechselwirken und wie Teilchenphysiker Vorhersagen machen können.
Klick mich an!
Bewege die Maus über mich!
Presentation
Du kommst von jeder Folie hierher zurück mit einem Klick auf diesen Knopf.
Diese Knöpfe leiten dich direkt zum jeweiligen Kapitel.
Inhalte und Ziele
Dieser Kurs enthält vier Kapitel, jedes mit seinem jeweiligen Ziel.
Wir empfehlen, ihn Schritt für Schritt durchzugehen. Aber du kannst jederzeit zurückkommen und dir ein vorhergehendes Kapitel anschauen.
Wann immer du einen Begriff oder eine Formulierung nicht verstehst, frag einfach nach.
1 Fundamentale Regeln für Teilchenprozesse
Du weißt welche fundamentalen Regeln Teilchenprozesse gehorchen und wie man diese in einfache Diagramme übersetzt.
2 Teilchenwechselwirkungen verstehen
Du verstehst, was es heißt, dass "Teilchen miteinander wechselwirken" und lernst einige Beispiele kennen.
3 Berechnungen mit Diagrammen verstehen
Du lernst, wie die Diagramme genutzt werden, um damit Berechnungen in der Teilchenphysik anzustellen.
4 Teilchenentdeckungen verstehen
Du verstehst, wie die Teilchen verwendet werden, um neue Teilchen zu entdecken und zu untersuchen.
Fundamentale Regeln für Teilchenwechselwirkungen
Vielleicht hast du schon einmal davon gehört, dass alles um uns herum - und auch du selbst - aus Teilchen besteht. Aber Teilchen alleine sind nicht besonders spannend. Spannend wird es, wenn wir uns anschauen, wie sie miteinander wechselwirken.
Diese Teilchen-Wechselwirkungen werden beschrieben durch eine Theorie, die wir das Standardmodell der Elementarteilchenphysik nennen.
Diese Theorie ist sehr mathematisch, aber es gibt ein bestimmte Art von Bildchen, die uns helfen können, diese Theorie zu verstehen. Diese Bildchen nennen wir Feynman-Diagramme.
Die Diagramme sind eigentlich Abkürzungen für mathematische Ausdrücke.
Feynman-Diagramme sind eine Kombination von unterschiedlichen Linien und deren Kreuzungspunkten, den sogenannten Vertices. Ein Feynman-Diagramm ist immer eine Kombination aus zwei oder mehr von diesen Vertices, wie hier dargestellt:
Linker Vertex
Rechter Vertex
Feynman diagram
Es gibt aber ein paar Regeln für diese Vertices, damit sie gültig sind. Diese lernst du nun kennen.

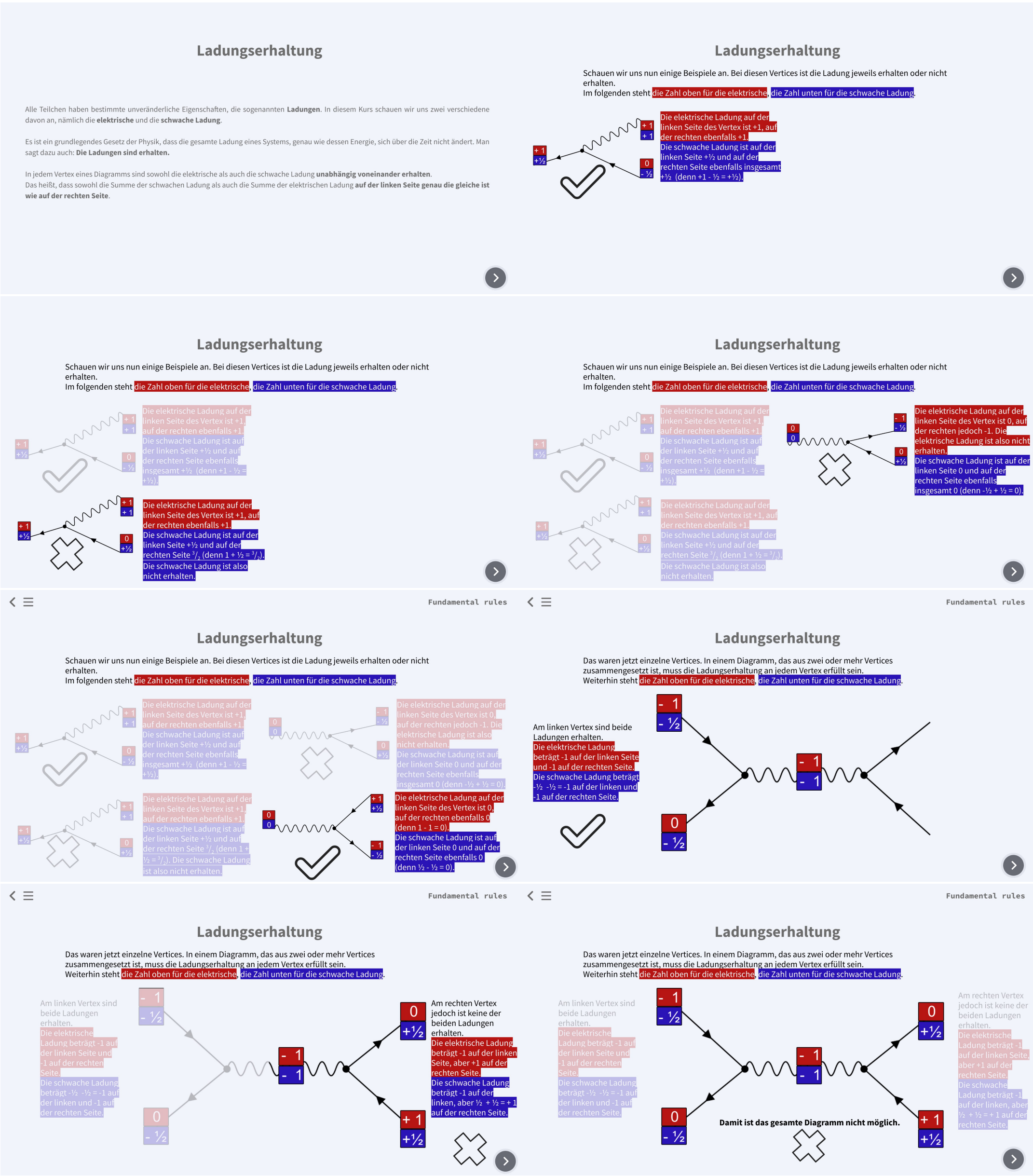


FIGURE E.1: Example slides from the third student study

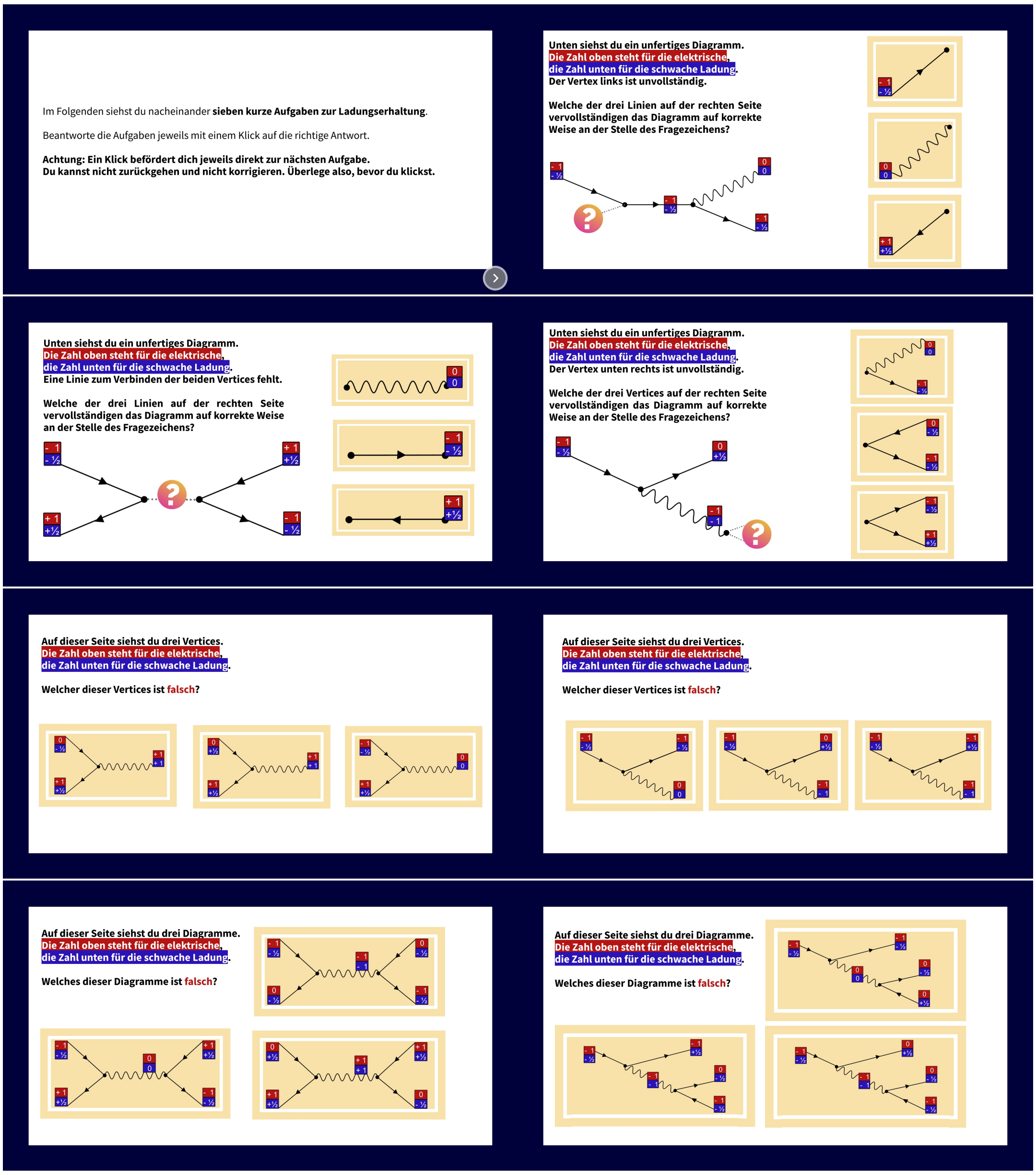


FIGURE E.2: Tasks in the third student study.

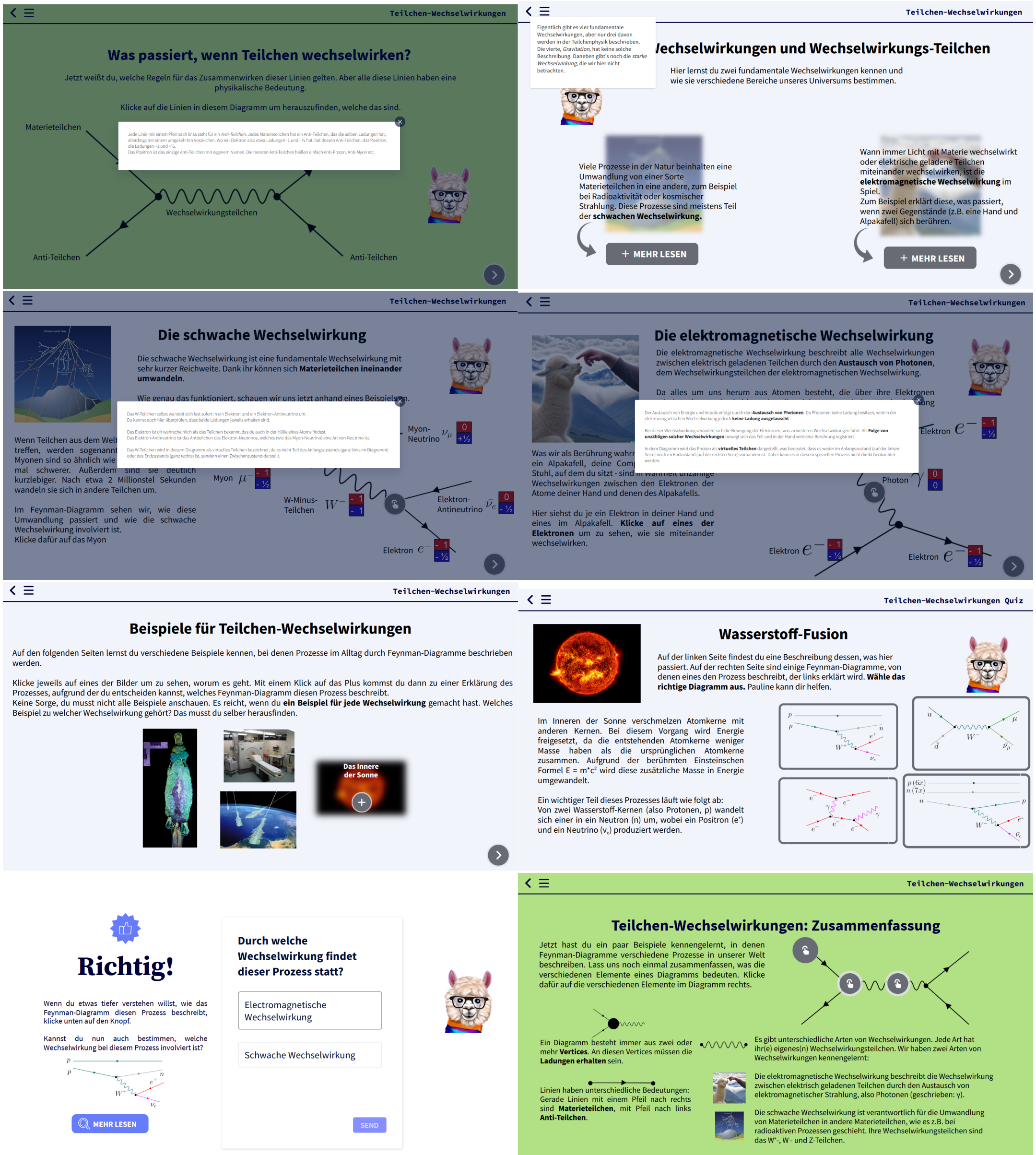


FIGURE E.3: Screenshots of the interactive part of the learning material in the third student study.

Bibliography

Abd-El-Khalick, F. (2012). Examining the sources for our understandings about science: Enduring conflations and critical issues in research on nature of science in science education [ISBN: 0950-0693 Publisher: Taylor & Francis]. *International Journal of Science Education*, *34*(3), 353–374. https://doi.org/10.1080/09500693.2011.629013

Adams, W. K., & Wieman, C. E. (2011). Development and Validation of Instruments to Measure Learning of Expert-Like Thinking [Publisher: Routledge _eprint: https://doi.org/10.1080/09500693.2010.512369]. *International Journal of Science Education*, *33*(9), 1289–1312. https://doi.org/10.1080/09500693.2010.512369

Ainsworth, S. (1999). The functions of multiple representations. *Computers & Education*, *33*(2), 131–152. https://doi.org/10.1016/S0360-1315(99)00029-9

Ainsworth, S. (2006). DeFT: A conceptual framework for considering learning with multiple representations [Place: Oxford Publisher: Pergamon-Elsevier Science Ltd WOS:000238326600001]. *Learning and Instruction*, *16*(3), 183–198. https://doi.org/10.1016/j.learninstruc.2006.03.001

Ainsworth, S., Prain, V., & Tytler, R. (2011). Drawing to Learn in Science. *Science*, *333*(6046), 1096–1097. https://doi.org/10.1126/science.1204153

Airey, J. Social Semiotics in Higher Education: Examples from teaching and learning in undergraduate physics. eng. In: Swedish Foundation for International Cooperation in Research in Higher Education (STINT), 2015, 103. Retrieved May 30, 2021, from http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-266049

Airey, J., & Linder, C. Social Semiotics in University Physics Education (D. F. Treagust, R. Duit, & H. E. Fischer, Eds.). en. In: *Multiple Representations in Physics Education* (D. F. Treagust, R. Duit, & H. E. Fischer, Eds.). Ed. by Treagust, D. F., Duit, R., & Fischer, H. E. Models and Modeling in Science Education. Cham: Springer International Publishing, 2017, pp. 95–122. ISBN: 978-3-319-58914-5. https://doi.org/10.1007/978-3-319-58914-5_5.

Alemdag, E., & Cagiltay, K. (2018). A systematic review of eye tracking research on multimedia learning. *Computers & Education*, *125*, 413–428. https://doi.org/10.1016/j.compedu.2018.06.023

Alexopoulos, A., Pavlidou, M., & Cherouvis, S. (2018). 'Playing with Protons': A training course for primary school teachers at CERN [Publisher: IOP Publishing]. *Physics Education*, *54*(1), 015013. https://doi.org/10.1088/1361-6552/aae7a4

Allday, J. (1997). The nature of force in particle physics [Publisher: IOP Publishing]. *Physics Education*, *32*(5), 327. https://doi.org/10.1088/0031-9120/32/5/016

American Association of Physics Teachers. (2024). Assessments. Retrieved February 24, 2024, from https://www.physport.org/assessments/

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive Tutors: Lessons Learned. *Journal of the Learning Sciences*, *4*(2), 167–207. https://doi.org/10.1207/s15327809jls0402_2

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing : A revision of Bloom's taxonomy of educational objectives* [Accepted: 2015-11-06T15:44:27Z Artwork Medium: Ressource physique Interview Medium: Ressource physique]. Longman. Retrieved February 2, 2022, from https://eduq.info/xmlui/handle/11515/18345

Anderson, T., & Shattuck, J. (2012). Design-Based Research: A Decade of Progress in Education Research? *Educational Researcher*, *41*(1), 16–25. https://doi.org/10.3102/0013189x11428813

Andrews, I., & Nikolopoulos, K. (2018). Introducing particle physics concepts through visual art [Publisher: IOP Publishing]. *Physics Education*, *53*(5), 054001. https://doi.org/10.1088/1361-6552/aad276

ATLAS collaboration. (2023). Improved W boson Mass Measurement using 7 TeV Proton-Proton Collisions with the ATLAS Detector [Place: Geneva]. Retrieved February 6, 2024, from https://cds.cern.ch/record/2853290

Aubrecht, G. J. (1986). Report on the conference on the teaching of modern physics. *The Physics Teacher*, *24*(9), 540–547. https://doi.org/10.1119/1.2342122

Baddeley, A. (1992). Working Memory. *Science*, *255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Barlow, R. (1992). Particle physics: From school to university [Publisher: IOP Publishing]. *Physics Education*, *27*(2), 92–95. https://doi.org/10.1088/0031-9120/27/2/007

BBC. (1981). The Pleasure of Finding Things Out. Retrieved February 25, 2024, from https://videos.cern.ch/record/1055846

Berg, E. v. d., & Hoekzema, D. (2006). Teaching conservation laws, symmetries and elementary particles with fast feedback. *Physics Education*, *41*(1), 47–56. https://doi.org/10.1088/0031-9120/41/1/004

Bertozzi, E. (2013). What is what we call the 'quantum field'? Answering from a teaching perspective by taking the foundations into account. *European Journal of Physics*, *34*(3), 603–611. https://doi.org/10.1088/0143-0807/34/3/603

Bilow, U., & Cecire, K. Current Status of International Particle Physics Masterclasses [Conference Name: 40th International Conference on High Energy physics].

en. In: In *Proceedings of 40th International Conference on High Energy physics — PoS(ICHEP2020)*. Vol. 390. Conference Name: 40th International Conference on High Energy physics. SISSA Medialab, 2021, April, p. 937. https://doi.org/10.22323/1.390.0937.

Bilow, U., & Cecire, K. (2022, May). Physics Masterclasses in Africa and the World [arXiv:2206.07004 [physics]]. https://doi.org/10.48550/arXiv.2206.07004

Bortz, J., & Doering, N. (2016). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler: Limitierte Sonderausgabe* (5th ed.). Springer-Verlag. https://doi.org/10.1007/978-3-540-33306-7

Bubble Chamber, m. B. (1972). Bubble Chamber pictures for education. Retrieved March 1, 2024, from https://cds.cern.ch/record/2307419

Burde, J.-P., & Wilhelm, T. Results of a Design-Based-Research Study to Improve Students' Understanding of Simple Electric Circuits (J. Guisasola & K. Zuza, Eds.). en. In: *Research and Innovation in Physics Education: Two Sides of the Same Coin* (J. Guisasola & K. Zuza, Eds.). Ed. by Guisasola, J., & Zuza, K. Challenges in Physics Education. Cham: Springer International Publishing, 2020, pp. 119–133. ISBN: 978-3-030-51182-1. https://doi.org/10.1007/978-3-030-51182-1_10.

Burde, J.-P., & Wilhelm, T. (2020b). Teaching electric circuits with a focus on potential differences [Publisher: American Physical Society]. *Physical Review Physics Education Research*, *16*(2), 020153. https://doi.org/10.1103/PhysRevPhysEducRes.16.020153

Cao, T. Y. The Quantum Field Programme for Fundamental Interactions (2nd ed.). In: In *Conceptual Developments of 20th Century Field Theories* (2nd ed.). 2nd ed. Cambridge University Press, 2019, pp. 107–234. https://doi.org/10.1017/9781108566926.009.

Castillo, L. (2014). *The Search and Discovery of the Higgs Boson*. Institute of Physics. http://dx.doi.org/10.1088/978-1-6817-4078-2ch1

CDF Collaboration, Aaltonen, T., Amerio, S., Amidei, D., Anastassov, A., Annovi, A., Antos, J., Apollinari, G., Appel, J. A., Arisawa, T., Artikov, A., Asaadi, J., Ashmanskas, W., Auerbach, B., Aurisano, A., Azfar, F., Badgett, W., Bae, T., Barbaro-Galtieri, A., ... Zucchelli, S. (2022). High-precision measurement of the W boson mass with the CDF II detector. *Science*, *376*(6589), 170–176. https://doi.org/10.1126/science.abk1781

Cecire, K., Bardeen, M., & McCauley, T. (2014). The CMS Masterclass and Particle Physics Outreach (L. Bravina, Y. Foka, & S. Kabana, Eds.). *EPJ Web of Conferences*, *71*, 00027. https://doi.org/10.1051/epjconf/20147100027

Cendrowski, M. (2008, April). The Bat Jar Conjecture [Distributor: Chuck Lorre Productions, Warner Bros. Television IMDb ID: tt1127387 event-location: United States].

Cendrowski, M. (2012, October). The Higgs Boson Observation [Distributor: Chuck Lorre Productions, Warner Bros. Television IMDb ID: tt2378184 event-location: United States].

CERN. (1983). UA1: First Z event recorded. Retrieved February 5, 2024, from http://cds.cern.ch/record/615824

CERN. (2023). Past Programmes | High-School Students Internship Programme. Retrieved February 26, 2024, from https://hssip.web.cern.ch/past-programmes

Cham, J. (2007). PHD Comics: More Quantum Gradnamics. Retrieved December 11, 2023, from https://phdcomics.com/comics.php?f=1461

Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction [ISBN: 0737-0008 Publisher: Taylor & Francis]. *Cognition and instruction*, *8*(4), 293–332.

Chase, W. G., & Simon, H. A. (1973). Perception in chess [Publisher: Academic Press]. *Cognitive Psychology*, *4*(1), 55–81. https://doi.org/10.1016/0010-0285(73)90004-2

Chatterjee, L. (2002). Introducing particle physics through story [Publisher: IOP Publishing]. *Physics Education*, *37*(3), 260–262. https://doi.org/10.1088/0031-9120/37/3/405

Chen, S.-C., & She, H.-C. (2020). Effects of Analogical Learning Approaches and Presentation Modalities on Ninth Graders' Learning Outcome and Eye Movements: A Preliminary Study. *Journal of Science Education and Technology*, *29*(4), 547–560. https://doi.org/10.1007/s10956-020-09835-7

Chiou, G.-L., Hsu, C.-Y., & Tsai, M.-J. (2022). Exploring how students interact with guidance in a physics simulation: Evidence from eye-movement and log data analyses. *Interactive Learning Environments*, *30*(3), 484–497. https://doi.org/10.1080/10494820.2019.1664596

Cundy, D., & Sutton, C. (2009). Gargamelle: The tale of a giant discovery. *CERN Courier*. Retrieved February 5, 2024, from https://cerncourier.com/a/gargamelle-the-tale-of-a-giant-discovery/

Dahlkemper, M., Wiener, J., Müller, A., Schmeling, S., & Klein, P. (2021). Untersuchung der Wahrnehmung von Feynman-Diagrammen mittels Eye Tracking [Number: 0]. *PhyDid B - Didaktik der Physik - Beiträge zur DPG-Frühjahrstagung*, *1*(0). Retrieved October 13, 2021, from http://www.phydid.de/index.php/phydid-b/article/view/1153

Dahlkemper, M. N., Klein, P., Müller, A., Schmeling, S. M., & Wiener, J. (2022). Opportunities and Challenges of Using Feynman Diagrams with Upper Secondary Students [Number: 4 Publisher: Multidisciplinary Digital Publishing Institute]. *Physics*, *4*(4), 1331–1347. https://doi.org/10.3390/physics4040085

Daniel, M. (2006). Particles, Feynman diagrams and all that [Publisher: IOP Publishing]. *Physics Education*, *41*(2), 119–129. https://doi.org/10.1088/0031-9120/41/2/001

Darrigol, O. (2019). The magic of Feynman's QED: From field-less electrodynamics to the Feynman diagrams. *The European Physical Journal H*, *44*(4-5), 349–369. https://doi.org/10.1140/epjh/e2019-100025-2

Day, J., Liao, T. Y.-H., Hoyt, C., & Can, O. (2022). Making the Quantum Scattering of Fundamental Particles Accessible. *The Physics Teacher*, *60*(2), 97–99. https://doi.org/10.1119/5.0021481

de Koning, B. B., Tabbers, H. K., Rikers, R. M. J. P., & Paas, F. (2009). Towards a Framework for Attention Cueing in Instructional Animations: Guidelines for Research and Design. *Educational Psychology Review*, *21*(2), 113–140. https://doi.org/10.1007/s10648-009-9098-7

De Cock, M. (2012). Representation use and strategy choice in physics problem solving [Publisher: American Physical Society]. *Physical Review Special Topics - Physics Education Research*, *8*(2), 020117. https://doi.org/10.1103/PhysRevSTPER.8.020117

Delabarre, E. B. (1898). A Method of Recording Eye-Movements [Publisher: University of Illinois Press]. *The American Journal of Psychology*, *9*(4), 572–574. https://doi.org/10.2307/1412191

Design-Based Research Collective. (2003). Design-Based Research: An Emerging Paradigm for Educational Inquiry. *Educational Researcher*, *32*(1), 5–8. https://doi.org/10.3102/0013189x032001005

Diehl, B., Erb, R., & Hesse, H. (2011). *Physik Oberstufe Gesamtband. Physik Oberstufe Gesamtband* (1. Aufl., 4. Druck). Cornelsen Verlag.

Ding, L., & Beichner, R. (2009). Approaches to data analysis of multiple-choice questions [Publisher: American Physical Society]. *Physical Review Special Topics - Physics Education Research*, *5*(2), 020103. https://doi.org/10.1103/PhysRevSTPER.5.020103

Dirac, P. A. M. (1963). The Evolution of the Physicist's Picture of Nature [Publisher: Scientific American, a division of Nature America, Inc.]. *Scientific American*, *208*(5), 45–53. Retrieved February 8, 2024, from https://www.jstor.org/stable/24936146

Dirac, P. A. M. (1927). The quantum theory of the emission and absorption of radiation [Publisher: Royal Society]. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *114*(767), 243–265. https://doi.org/10.1098/rspa.1927.0039

Dirac, P. A. M. (1930). A theory of electrons and protons (R. H. Fowler, Ed.) [Publisher: Royal Society]. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, *126*(801), 360–365. https://doi.org/10.1098/rspa.1930.0013

Dorris, M. R., & Rau, M. A. (2022). Conceptual Challenges Exhibited by Naıve Undergraduate Students in the Context of Atomic Orbital Energy Diagrams [Publisher: American Chemical Society]. *Journal of Chemical Education*, *99*(8), 2777–2786. https://doi.org/10.1021/acs.jchemed.1c01135

Duchowski, A. (2007). *Eye Tracking Methodology: Theory and Practice* (2nd ed.). Springer-Verlag. https://doi.org/10.1007/978-1-84628-609-4

Duit, R. (1996). The constructivist view in science education–what it has to offer and what should not be expected from it [ISBN: 1518-8795]. *Investigaçôes em ensino de ciências*, *1*(1), 40–75.

Duit, R., Gropengießer, H., Kattmann, U., Komorek, M., & Parchmann, I. The Model of Educational Reconstruction – a Framework for Improving Teaching and Learning Science (D. Jorde & J. Dillon, Eds.). en. In: *Science Education Research and Practice in Europe: Retrosspective and Prospecctive* (D. Jorde & J. Dillon, Eds.). Ed. by Jorde, D., & Dillon, J. Cultural Perpectives in Science Education. Rotterdam, The Netherlands: SensePublishers, 2012, pp. 13–37. ISBN: 978-94-6091-900-8. https://doi.org/10.1007/978-94-6091-900-8_2.

Duit, R., & Komorek, M. (1997). Understanding the basic ideas of chaos-theory in a study of limited predictability. *International Journal of Science Education*, *19*(3), 247–264. https://doi.org/10.1080/0950069970190301

Dunne, P. (2001). Looking for consistency in the construction and use of Feynman diagrams [Publisher: IOP Publishing]. *Physics Education*, *36*(5), 366–374. https://doi.org/10.1088/0031-9120/36/5/301

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random house.

Dyson, F. J. (1949a). The Radiation Theories of Tomonaga, Schwinger, and Feynman. *Physical Review*, *75*(3), 486–502. https://doi.org/10.1103/PhysRev.75.486

Dyson, F. J. (1949b). The S Matrix in Quantum Electrodynamics [Publisher: American Physical Society]. *Physical Review*, *75*(11), 1736–1755. https://doi.org/10.1103/PhysRev.75.1736

Ellis, J., Gaillard, M. K., Nanopoulos, D. V., & Rudaz, S. (1977). The phenomenology of the next left-handed quarks. *Nuclear Physics B*, *131*(2), 285–307. https://doi.org/10.1016/0550-3213(77)90374-1

Fan, X., Myers, T., Sukra, B., & Gabrielse, G. (2023). Measurement of the Electron Magnetic Moment [Publisher: American Physical Society]. *Physical Review Letters*, *130*(7), 071801. https://doi.org/10.1103/PhysRevLett.130.071801

Fenkart. (2023). CERN's neutrino odyssey. *CERN Courier*. Retrieved February 6, 2024, from https://cerncourier.com/a/cerns-neutrino-odyssey/

Feynman, R. P. (1949a). Space-Time Approach to Quantum Electrodynamics. *Physical Review*, *76*(6), 769–789. https://doi.org/10.1103/PhysRev.76.769

Feynman, R. P. (1949b). The Theory of Positrons [Publisher: American Physical Society]. *Physical Review*, *76*(6), 749–759. https://doi.org/10.1103/PhysRev.76.749

Feynman, R. P. (1950). Mathematical Formulation of the Quantum Theory of Electromagnetic Interaction [Publisher: American Physical Society]. *Physical Review*, *80*(3), 440–457. https://doi.org/10.1103/PhysRev.80.440

Feynman, R. P. (1966). The Development of the Space-Time View of Quantum Electrodynamics [Publisher: American Association for the Advancement of Science]. *Science*, *153*(3737), 699–708. Retrieved May 26, 2021, from http://www.jstor.org/stable/1719094

Feynman, R. P. (1985). *QED: The Strange Theory of Light and Matter*. Princeton University Press.

Finta, Z., Jenei, P., & Schramek, A. (2021). The subject of waves in a new approach introductory steps of a Design Based Research (DBR) [Publisher: IOP Publishing]. *Journal of Physics: Conference Series*, *1929*(1), 012018. https://doi.org/10.1088/1742-6596/1929/1/012018

Fredlund, T., Airey, J., & Linder, C. (2012). Exploring the role of physics representations: An illustrative example from students sharing knowledge about refraction [Publisher: IOP Publishing]. *European Journal of Physics*, *33*(3), 657–666. https://doi.org/10.1088/0143-0807/33/3/657

Fredlund, T., Linder, C., Airey, J., & Linder, A. (2014). Unpacking physics representations: Towards an appreciation of disciplinary affordance. *Physical Review Special Topics - Physics Education Research*, *10*(2), 020129. https://doi.org/10.1103/PhysRevSTPER.10.020129

Gegenfurtner, A., Lehtinen, E., & Säljö, R. (2011). Expertise Differences in the Comprehension of Visualizations: A Meta-Analysis of Eye-Tracking Research in Professional Domains. *Educational Psychology Review*, *23*(4), 523–552. https://doi.org/10.1007/s10648-011-9174-7

Gobet, F., Lane, P. C. R., Croker, S., Cheng, P. C.-H., Jones, G., Oliver, I., & Pine, J. M. (2001). Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, *5*(6), 236–243. https://doi.org/10.1016/S1364-6613(00)01662-4

Gourlay, H. (2016). Learning about A level physics students' understandings of particle physics using concept mapping [Publisher: IOP Publishing]. *Physics Education*, *52*(1), 014001. https://doi.org/10.1088/1361-6552/52/1/014001

Grehn, J., & Krause, J. (Eds.). (2011). *Metzler Physik. Schülerbd.* (3. Aufl., Dr. A, 10). Metzler.

Griffiths, D. (2008). *Introduction to Elementary Particles* (2nd, Revised Edition). Wiley-VCH.

Grünewald, M. (2005). Experiments finally unveil a precise portrait of the Z. *CERN Courier*. Retrieved February 5, 2024, from https://cerncourier.com/a/experiments-finally-unveil-a-precise-portrait-of-the-z/

Haagen-Schützenhöfer, C., & Hopf, M. (2020). Design-based research as a model for systematic curriculum development: The example of a curriculum for introductory optics. *Physical Review Physics Education Research*, *16*(2), 020152. https://doi.org/10.1103/PhysRevPhysEducRes.16.020152

Hahn, L., & Klein, P. (2022). Eye tracking in physics education research: A systematic literature review. *Physical Review Physics Education Research*, *18*(1), 013102. https://doi.org/10.1103/PhysRevPhysEducRes.18.013102

Hahn, L., & Klein, P. (2023). Clustering eye-movement data uncovers students' strategies for coordinating equations and diagrams of vector fields. *Educational Studies in Mathematics*. https://doi.org/10.1007/s10649-023-10243-y

Han, J., Chen, L., Fu, Z., Fritchman, J., & Bao, L. (2017). Eye-tracking of visual attention in web-based assessment using the Force Concept Inventory [Publisher: IOP Publishing]. *European Journal of Physics*, *38*(4), 045702. https://doi.org/10.1088/1361-6404/aa6c49

Hansen, D. W., & Ji, Q. (2010). In the Eye of the Beholder: A Survey of Models for Eyes and Gaze [Conference Name: IEEE Transactions on Pattern Analysis and Machine Intelligence]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *32*(3), 478–500. https://doi.org/10.1109/TPAMI.2009.30

Harlander, R. V., Klein, S. Y., & Lipp, M. (2020). FeynGame. *Computer Physics Communications*, *256*, 107465. https://doi.org/10.1016/j.cpc.2020.107465

Harlander, R. (2021). Feynman diagrams. *Synthese*. https://doi.org/10.1007/s11229-021-03387-y

Harrison, A. G., & Treagust, D. F. The Particulate Nature of Matter: Challenges in Understanding the Submicroscopic World (J. K. Gilbert, O. De Jong, R. Justi, D. F. Treagust, & J. H. Van Driel, Eds.). en. In: *Chemical Education: Towards Research-based Practice* (J. K. Gilbert, O. De Jong, R. Justi, D. F. Treagust, & J. H. Van Driel, Eds.). Ed. by Gilbert, J. K., De Jong, O., Justi, R., Treagust, D. F., & Van Driel, J. H. Science & Technology Education Library. Dordrecht: Springer Netherlands, 2003, pp. 189–212. ISBN: 978-0-306-47977-9. Retrieved July 1, 2022, from https://doi.org/10.1007/0-306-47977-X_9

Harré, R. Parsing the amplitudes (H. R. Brown & R. Harré, Eds.). In: *Philosophical Foundations of Quantum Field Theory* (H. R. Brown & R. Harré, Eds.). Ed. by Brown, H. R., & Harré, R. New York: Oxford University Press, 1988, pp. 59–71. ISBN: 0-19-824975-6.

Hasert, F. J., Faissner, H., Krenz, W., Von Krogh, J., Lanske, D., Morfin, J., Schultze, K., Weerts, H., Bertrand-Coremans, G. H., Lemonne, J., Sacton, J., Van Doninck, W., Vilain, P., Baltay, C., Cundy, D. C., Haidt, D., Jaffre, M., Musset, P., Pullia, A., … Scott, W. G. (1973). Search for elastic muon-neutrino electron scattering. *Physics Letters B*, *46*(1), 121–124. https://doi.org/10.1016/0370-2693(73)90494-2

Heisenberg, W., & Pauli, W. (1929). Zur Quantendynamik der Wellenfelder. *Zeitschrift für Physik*, *56*(1), 1–61. https://doi.org/10.1007/BF01340129

Hobson, A. (2005). Electrons as field quanta: A better way to teach quantum physics in introductory general physics courses. *American Journal of Physics*, *73*(7), 630–634. https://doi.org/10.1119/1.1900097

Hobson, A. (2010). Teaching Elementary Particle Physics, Part I [Publisher: American Association of Physics TeachersAAPT]. *The Physics Teacher*, *49*(1), 12–15. https://doi.org/10.1119/1.3527746

Hobson, A. (2011). Teaching Elementary Particle Physics: Part II [Publisher: American Association of Physics TeachersAAPT]. *The Physics Teacher*, *49*(3), 136–138. https://doi.org/10.1119/1.3555495

Hoekzema, D., Schooten, G., van den Berg, E., & Lijnse, P. (2005). Conservation Laws, Symmetries, and Elementary Particles [Publisher: American Association of Physics Teachers]. *The Physics Teacher*, *43*(5), 266–271. https://doi.org/10.1119/1.1903808

Holmqvist, K., Andrà, C., Lindström, P., Arzarello, F., Ferrara, F., Robutti, O., & Sabena, C. (2011). A method for quantifying focused versus overview behavior in AOI sequences. *Behavior Research Methods*, *43*(4), 987–998. https://doi.org/10.3758/s13428-011-0104-x

Hoyer, C., & Girwidz, R. (2020). Animation and interactivity in computer-based physics experiments to support the documentation of measured vector quantities in diagrams: An eye tracking study [Publisher: American Physical Society]. *Physical Review Physics Education Research*, *16*(2), 020124. https://doi.org/10.1103/PhysRevPhysEducRes.16.020124

Humphrey, G. (1924). The psychology of the gestalt. [ISBN: 1939-2176 Publisher: Warwick & York]. *Journal of Educational Psychology*, *15*(7), 401.

Hyönä, J. (2010). The use of eye movements in the study of multimedia learning. *Learning and Instruction*, *20*(2), 172–176. https://doi.org/10.1016/j.learninstruc.2009.02.013

Höttecke, D., & Schecker, H. Unterrichtskonzeptionen für Nature of Science (NOS) (T. Wilhelm, H. Schecker, & M. Hopf, Eds.). de. In: *Unterrichtskonzeptionen für den Physikunterricht* (T. Wilhelm, H. Schecker, & M. Hopf, Eds.). Ed. by Wilhelm, T., Schecker, H., & Hopf, M. Berlin, Heidelberg: Springer Berlin Heidelberg, 2021, pp. 401–433. ISBN: 978-3-662-63052-5 978-3-662-63053-2. https://doi.org/10.1007/978-3-662-63053-2_13.

Ibrahim, B., & Ding, L. (2021). Sequential and simultaneous synthesis problem solving: A comparison of students' gaze transitions [Publisher: American Physical Society]. *Physical Review Physics Education Research*, *17*(1), 010126. https://doi.org/10.1103/PhysRevPhysEducRes.17.010126

Incandela, J., Gianotti, F., & Heuer, R. (2012, July). Latest update in the search for the Higgs boson. Retrieved May 4, 2022, from https://indico.cern.ch/event/197461/

Jarodzka, H., Holmqvist, K., & Gruber, H. (2017). Eye tracking in Educational Science: Theoretical frameworks and research agendas. *Journal of Eye Movement Research*, *10*(1). https://doi.org/10.16910/jemr.10.1.3

Jauch, J. M., & Rohrlich, F. (1955). *The Theory of Photons and Electrons* [Google-Books-ID: Epn6CAAAQBAJ]. Addison-Wesley Publishing. https://pubs.aip.org/physicstoday/article-abstract/9/8/34/416995/The-Theory-of-Photons-and-Electrons

Jepsen, K. (2014, May). Saving the Feynman van | symmetry magazine. Retrieved December 12, 2023, from https://www.symmetrymagazine.org/article/may-2014/saving-the-feynman-van?language_content_entity=und

Johansson, K. E., Kobel, M., Hillebrandt, D., Engeln, K., & Euler, M. (2007). European particle physics masterclasses make students into scientists for a day [Publisher: IOP Publishing]. *Physics Education*, *42*(6), 636–644. https://doi.org/10.1088/0031-9120/42/6/012

Johansson, K. E., & Watkins, P. M. (2012). Exploring the standard model of particles [Publisher: IOP Publishing]. *Physics Education*, *48*(1), 105–114. https://doi.org/10.1088/0031-9120/48/1/105

Jones, G. T. (2002). The uncertainty principle, virtual particles and real forces [Publisher: IOP Publishing]. *Physics Education*, *37*(3), 223–233. https://doi.org/10.1088/0031-9120/37/3/306

Jordan, J., & Slater, M. (2009). An Analysis of Eye Scanpath Entropy in a Progressively Forming Virtual Environment. *Presence: Teleoperators and Virtual Environments*, *18*(3), 185–199. https://doi.org/10.1162/pres.18.3.185

Jordan, P., & Pauli, W. (1928). Zur Quantenelektrodynamik ladungsfreier Felder. *Zeitschrift für Physik*, *47*(3), 151–173. https://doi.org/10.1007/BF02055793

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension [Place: US Publisher: American Psychological Association]. *Psychological Review*, *87*(4), 329–354. https://doi.org/10.1037/0033-295X.87.4.329

Kaiser, D. (2005). *Drawing Theories Apart: The Dispersion of Feynman Diagrams in Postwar Physics* [Google-Books-ID: _wz9JZvBt78C]. University of Chicago Press.

Kamphorst, F., Vollebregt, M. J., Savelsbergh, E. R., & van Joolingen, W. R. (2021). An Educational Reconstruction of Special Relativity Theory for Secondary Education. *Science & Education*. https://doi.org/10.1007/s11191-021-00283-2

Kattmann, U., Duit, R., Gropengießer, H., & Komorek, M. Educational reconstruction–bringing together issues of scientific clarification and students' conceptions. In: In *Annual meeting of the national association of research in science teaching (NARST), St. Louis*. 1996.

Kattmann, U., Duit, R., Gropengießer, H., & Komorek, M. (1997). Das Modell der Didaktischen Rekonstruktion [Publisher: Schmidt & Klaunig Kiel]. *Zeitschrift für Didaktik der Naturwissenschaften*, *3*(3), 3–18.

Kekule, M., & Viiri, J. (2018). Students' approaches to solving R-FCI tasks observed by eye-tracking method [Number: 2]. *Scientia in educatione*, *9*(2), 117–130. https://doi.org/10.14712/18047106.1010

Keller, J. M. (1987). Development and use of the ARCS model of instructional design. *Journal of instructional development*, *10*(3), 2–10. https://doi.org/10.1007/BF02905780

Keller, J. M. (2010). *Motivational Design for Learning and Performance: The ARCS Model Approach*. Springer US. https://doi.org/10.1007/978-1-4419-1250-3

Kersting, M. (2019a). Free fall in curved spacetime—how to visualise gravity in general relativity [Publisher: IOP Publishing]. *Physics Education*, *54*(3), 035008. https://doi.org/10.1088/1361-6552/ab08f5

Kersting, M. (2019b). *General Relativity in Secondary School: Research-Based Development of Learning Resources and Analyses of Students' Conceptual Understanding Using the Model of Educational Reconstruction* (PhD) [Accepted: 2019-12-20T12:35:57Z]. University of Oslo. Retrieved August 5, 2021, from https://www.duo.uio.no/handle/10852/71817

Kersting, M., Blair, D., Sandrelli, S., Sherson, J., & Woithe, J. (2023). Making an IMPRESSion: Mapping out future directions in modern physics education [Publisher: IOP Publishing]. *Physics Education*, *59*(1), 015501. https://doi.org/10.1088/1361-6552/ad11e8

Kersting, M., Henriksen, E. K., Bøe, M. V., & Angell, C. (2018). General relativity in upper secondary school: Design and evaluation of an online learning environment using the model of educational reconstruction [Publisher: American Physical Society]. *Physical Review Physics Education Research*, *14*(1), 010130. https://doi.org/10.1103/PhysRevPhysEducRes.14.010130

Kircher, E., Girwidz, R., & Häußler, P. (Eds.). (2015). *Physikdidaktik: Theorie und Praxis* (3rd ed.). Springer Spektrum. https://doi.org/10.1007/978-3-642-41745-0

Klein, P., Becker, S., Küchemann, S., & Kuhn, J. (2021a). Test of understanding graphs in kinematics: Item objectives confirmed by clustering eye movement transitions [Publisher: American Physical Society]. *Physical Review Physics Education Research*, *17*(1), 013102. https://doi.org/10.1103/PhysRevPhysEducRes.17.013102

Klein, P., Lichtenberger, A., Küchemann, S., Becker, S., Kekule, M., Viiri, J., Baadte, C., Vaterlaus, A., & Kuhn, J. (2020). Visual attention while solving the test of understanding graphs in kinematics: An eye-tracking analysis. *European Journal of Physics*, *41*(2), 025701. https://doi.org/10.1088/1361-6404/ab5f51

Klein, P., Viiri, J., & Kuhn, J. (2019). Visual cues improve students' understanding of divergence and curl: Evidence from eye movements during reading and problem solving [Publisher: American Physical Society]. *Physical Review Physics Education Research*, *15*(1), 010126. https://doi.org/10.1103/PhysRevPhysEducRes.15.010126

Klein, P., Viiri, J., Mozaffari, S., Dengel, A., & Kuhn, J. (2018). Instruction-based clinical eye-tracking study on the visual interpretation of divergence: How do students look at vector field plots? [Publisher: American Physical Society]. *Physical Review Physics Education Research*, *14*(1), 010116. https://doi.org/10.1103/PhysRevPhysEducRes.14.010116

Klein, P., Hahn, L., & Kuhn, J. (2021b). Einfluss visueller Hilfen und räumlicher Fähigkeiten auf die graphische Interpretation von Vektorfeldern: Eine Eye-Tracking-Untersuchung. *Zeitschrift für Didaktik der Naturwissenschaften*, *27*(1), 181–201. https://doi.org/10.1007/s40573-021-00133-2

Kobel, M. (2003). High school students' exposure to modern particle physics [Number: 3 Publisher: EDP Sciences]. *Europhysics News*, *34*(3), 108–110. https://doi.org/10.1051/epn:2003307

Kobel, M., Bilow, U., Lindenau, P., & Schorn, B. (2022). *Teilchenphysik: Ladungen, Wechselwirkungen und Teilchen* (4. Auflage). Netzwerk Teilchenwelt. https://doi.org/10.25368/2021.71

Kobel, M., Bilow, U., Lindenau, P., Schorn, B., & Bachmann, G. (2021). *Teilchenphysik: Forschungsmethoden* (3. Auflage). Netzwerk Teilchenwelt. https://doi.org/10.25368/2021.71

Kobel, M., & Lindenau, P. Eine anschlussfähige Begriffsbildung der Basiskonzepte des Standardmodells für die Schule (O. Passon, T. Zügge, & J. Grebe-Ellis, Eds.). de. In: *Kohärenz im Unterricht der Elementarteilchenphysik: Tagungsband des Symposiums zur Didaktik der Teilchenphysik, Wuppertal 2018* (O. Passon, T. Zügge, & J. Grebe-Ellis, Eds.). Ed. by Passon, O., Zügge, T., & Grebe-Ellis, J. Berlin, Heidelberg: Springer, 2020, pp. 93–106. ISBN: 978-3-662-61607-9. https://doi.org/10.1007/978-3-662-61607-9_6.

Kontokostas, G., & Kalkanis, G. (2013). Teaching Electron—Positron—Photon Interactions with Hands-on Feynman Diagrams [Publisher: American Association of Physics Teachers]. *The Physics Teacher*, *51*(4), 232–233. https://doi.org/10.1119/1.4795369

Kranjc Horvat, A. (2022). *Particle physics in high-school education : What should students and teachers learn?* (Doctoral dissertation). Universität Potsdam. https://doi.org/10.25932/publishup-56026

Kranjc Horvat, A., Wiener, J., Schmeling, S. M., & Borowski, A. (2022). What Does the Curriculum Say? Review of the Particle Physics Content in 27 High-School Physics Curricula [Number: 4 Publisher: Multidisciplinary Digital Publishing Institute]. *Physics*, *4*(4), 1278–1298. https://doi.org/10.3390/physics4040082

Krejtz, K., Duchowski, A., Szmidt, T., Krejtz, I., González Perilli, F., Pires, A., Vilaro, A., & Villalobos, N. (2015). Gaze Transition Entropy. *ACM Transactions on Applied Perception*, *13*(1), 4:1–4:20. https://doi.org/10.1145/2834121

Krejtz, K., Duchowski, A. T., Krejtz, I., Kopacz, A., & Chrząstowski-Wachtel, P. (2016). Gaze transitions when learning with multimedia [Number: 1]. *Journal of Eye Movement Research*, *9*(1). https://doi.org/10.16910/jemr.9.1.5

Krejtz, K., Szmidt, T., Duchowski, A. T., & Krejtz, I. Entropy-based statistical analysis of eye movement transitions. en. In: In *Proceedings of the Symposium on Eye Tracking Research and Applications*. Safety Harbor Florida: ACM, 2014, March, 159–166. ISBN: 978-1-4503-2751-0. https://doi.org/10.1145/2578153.2578176.

Krijtenburg-Lewerissa, K., Pol, H., Brinkman, A., & Van Joolingen, W. (2017). Insights into teaching quantum mechanics in secondary and lower undergraduate education. *Physical Review Physics Education Research*, *13*(1), 010109. https://doi.org/10.1103/PhysRevPhysEducRes.13.010109

Kuhlmann, M. Process Ontology [Google-Books-ID: Y6Oot1o0vsAC]. en. In: In *The Ultimate Constituents of the Material World: In Search of an Ontology for Fundamental Physics*. Google-Books-ID: Y6Oot1o0vsAC. Walter de Gruyter, 2010, pp. 121–134. ISBN: 978-3-11-032612-3.

Kuhlmann, M. (2023). Quantum Field Theory (E. N. Zalta & U. Nodelman, Eds.; Summer 2023). https://plato.stanford.edu/archives/sum2023/entries/quantum-field-theory/

Kulgemeyer, C. (2020). A Framework of Effective Science Explanation Videos Informed by Criteria for Instructional Explanations. *Research in Science Education*, *50*(6), 2441–2462. https://doi.org/10.1007/s11165-018-9787-7

Kulgemeyer, C., & Schecker, H. (2013). Students Explaining Science—Assessment of Science Communication Competence. *Research in Science Education*, *43*(6), 2235–2256. https://doi.org/10.1007/s11165-013-9354-1

Küchemann, S., Klein, P., Becker, S., Kumari, N., & Kuhn, J. Classification of Students' Conceptual Understanding in STEM Education using Their Visual Attention Distributions: A Comparison of Three Machine-Learning Approaches. In: 2021, May, 36–46. ISBN: 978-989-758-417-6. Retrieved May 21, 2021, from https://www.scitepress.org/PublicationsDetail.aspx?ID=RR+w7e5T9Tc=&t=1

Küchemann, S., Ubben, M., Dzsotjan, D., Mukhametov, S., Weidner, C. A., Qerimi, L., Kuhn, J., Heusler, S., & Sherson, J. F. (2023, February). The impact of an interactive visualization and simulation tool on learning quantum physics: Results of an eye-tracking study [arXiv:2302.06286 [physics]]. Retrieved August 17, 2023, from http://arxiv.org/abs/2302.06286

Lacki, J. (2017). An EPS Historic Sites Award celebrating two Genevan physicists. *SPS Communications*, (52).

Lai, M.-L., Tsai, M.-J., Yang, F.-Y., Hsu, C.-Y., Liu, T.-C., Lee, S. W.-Y., Lee, M.-H., Chiou, G.-L., Liang, J.-C., & Tsai, C.-C. (2013). A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, *10*, 90–115. https://doi.org/10.1016/j.edurev.2013.10.001

Lambourne, R. (1992). Predicting the physics of particles [Publisher: IOP Publishing]. *Physics Education*, *27*(2), 71. https://doi.org/10.1088/0031-9120/27/2/003

Langendorf, R. (2022, September). *Technologiebezogene Professionalisierung angehender Physiklehrkräfte - Entwicklung und Evaluation eines Lehrkonzepts in der Astrophysik mit digitalen Medien* (doctoralThesis) [Accepted: 2022-09-09T12:38:36Z ISBN: 9781816394231]. https://doi.org/10.53846/goediss-9439

Langendorf, R., Schneider, S., & Klein, P. (2022). Extracting information from the Hertzsprung-Russell diagram: An eye-tracking study [Publisher: American Physical Society]. *Physical Review Physics Education Research*, *18*(2), 020121. https://doi.org/10.1103/PhysRevPhysEducRes.18.020121

Launer, S. (2020). *Wechselwirkungskonzepte der Physik und ihre Propädeutik für die gymnasiale Oberstufe* (Doctoral Thesis) [Accepted: 2020-08-14T08:06:22Z Journal Abbreviation: Interaction concepts of physics and their propaedeutics for high school]. https://doi.org/10.25819/ubsi/4204

Lederman, N. G. Nature of science: Past, present, and future (1st ed.). In: In *Handbook of Research on Science Education* (1st ed.). 1st ed. Routledge, 2013, pp. 845–894. ISBN: 0-203-82469-5. https://doi.org/10.4324/9780203824696.

Leppink, J., Paas, F., van Gog, T., van der Vleuten, C. P. M., & van Merriënboer, J. J. G. (2014). Effects of pairs of problems and examples on task performance and different types of cognitive load. *Learning and Instruction*, *30*, 32–42. https://doi.org/10.1016/j.learninstruc.2013.12.001

Leppink, J., Paas, F., Van der Vleuten, C. P. M., Van Gog, T., & Van Merriënboer, J. J. G. (2013). Development of an instrument for measuring different types of cognitive load. *Behavior Research Methods*, *45*(4), 1058–1072. https://doi.org/10.3758/s13428-013-0334-1

Lindenau, P., & Kobel, M. (2019). Introducing an innovative approach of teaching the Standard Model of particle physics at high school [Publisher: IOP Publishing]. *Journal of Physics: Conference Series*, *1287*, 012045. https://doi.org/10.1088/1742-6596/1287/1/012045

Lindenau, P., & Winkler, O. (2020). Feynman-Rhombino. *Unterricht Physik*, *180*, 40–43. https://www.friedrich-verlag.de/physik/quantenphysik/feynman-rhombino-7486

MacDonald, T., & Bean, A. (2009). Quarked! — Adventures in Particle Physics Education. *The Physics Teacher*, *47*(1), 38–41. https://doi.org/10.1119/1.3049879

Madsen, A., McKagan, S. B., & Sayre, E. C. (2017). Resource Letter RBAI-1: Research-Based Assessment Instruments in Physics and Astronomy [Publisher: American Association of Physics Teachers]. *American Journal of Physics*, *85*(4), 245–264. https://doi.org/10.1119/1.4977416

Madsen, A., McKagan, S. B., Sayre, E. C., & Paul, C. A. (2019). Resource Letter RBAI-2: Research-based assessment instruments: Beyond physics topics [Publisher: American Association of Physics Teachers]. *American Journal of Physics*, *87*(5), 350–369. https://doi.org/10.1119/1.5094139

Madsen, A., Rouinfar, A., Larson, A. M., Loschky, L. C., & Rebello, N. S. (2013). Can short duration visual cues influence students' reasoning and eye movements in physics problems? [Publisher: American Physical Society]. *Physical Review Special Topics - Physics Education Research*, *9*(2), 020104. https://doi.org/10.1103/PhysRevSTPER.9.020104

Malgieri, M., Onorato, P., & De Ambrosis, A. (2017). Test on the effectiveness of the sum over paths approach in favoring the construction of an integrated knowledge of quantum physics in high school [Publisher: American Physical Society]. *Physical Review Physics Education Research*, *13*(1), 010101. https://doi.org/10.1103/PhysRevPhysEducRes.13.010101

Marcelloni, C. (2007). John Ellis, CERN theorist, working on the blackboard. John Ellis, théoricien, travaillant au tableau noir. Retrieved January 29, 2024, from https://cds.cern.ch/record/1071200

Mayer, R. E. Multimedia learning. en. In: In *Psychology of Learning and Motivation*. Vol. 41. Academic Press, 2002, January, pp. 85–139. https://doi.org/10.1016/S0079-7421(02)80005-6.

Mayer, R. E. (2002b). Rote Versus Meaningful Learning. *Theory Into Practice*, *41*(4), 226–232. https://doi.org/10.1207/s15430421tip4104_4

Mayer, R. E. (2005). Cognitive theory of multimedia learning. *The Cambridge handbook of multimedia learning*, *41*, 31–48.

Mayer, R. E. (2008). Applying the science of learning: Evidence-based principles for the design of multimedia instruction. *American Psychologist*, *63*(8), 760–769. https://doi.org/10.1037/0003-066X.63.8.760

Mayer, R. E. (2010). Unique contributions of eye-tracking research to the study of learning with graphics. *Learning and Instruction*, *20*(2), 167–171. https://doi.org/10.1016/j.learninstruc.2009.02.012

McGinness, L., Dührkoop, S., Woithe, J., & Jansky, A. (2019). 3D Printable Quark Puzzle: A Model to Build Your Own Particle Systems. *The Physics Teacher*, *57*(8), 526–528. https://doi.org/10.1119/1.5131116

McKagan, S. B., Perkins, K. K., Dubson, M., Malley, C., Reid, S., LeMaster, R., & Wieman, C. E. (2008). Developing and researching PhET simulations for teaching quantum mechanics. *American Journal of Physics*, *76*(4), 406–417. https://doi.org/10.1119/1.2885199

Meynell, L. (2008). Why Feynman Diagrams Represent. *International Studies in the Philosophy of Science*, *22*(1), 39–59. https://doi.org/10.1080/02698590802280902

Molz, A. (2016). *Verbindung von Schülerlabor und Schulunterricht-Auswirkungen auf Motivation und Kognition im Fach Physik*. Verlag Dr. Hut.

Moreno, R., & Mayer, R. (2007). Interactive Multimodal Learning Environments: Special Issue on Interactive Learning Environments: Contemporary Issues and Trends. *Educational Psychology Review*, *19*(3), 309–326. https://doi.org/10.1007/s10648-007-9047-2

Mozaffari Chanijani, S. S., Klein, P., Bukhari, S. S., Kuhn, J., & Dengel, A. Entropy based transition analysis of eye movement on physics representational competence. In: In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct*. UbiComp '16. New York, NY, USA: Association for Computing Machinery, 2016, September, 1027–1034. ISBN: 978-1-4503-4462-3. https://doi.org/10.1145/2968219.2968564.

Munroe, R. (2015a, February). Fundamental Forces. Retrieved February 28, 2024, from https://xkcd.com/1489/

Munroe, R. (2015b, December). Xkcd: Fixion. Retrieved December 11, 2023, from https://xkcd.com/1621/

Nikolopoulos, K., & Pardalaki, M. (2020). Particle dance: Particle physics in the dance studio [Publisher: IOP Publishing]. *Physics Education*, *55*(2), 025018. https://doi.org/10.1088/1361-6552/ab6952

Nobel Prize Outreach AB. (2013, October). The Nobel Prize in Physics 2013. Retrieved April 25, 2022, from https://www.nobelprize.org/prizes/physics/2013/summary/

Nobel Prize Outreach AB. (2024a). The Nobel Prize in Physics 1965. Retrieved March 12, 2024, from https://www.nobelprize.org/prizes/physics/1965/summary/

Nobel Prize Outreach AB. (2024b). The Nobel Prize in Physics 1984. Retrieved February 6, 2024, from https://www.nobelprize.org/prizes/physics/1984/summary/

Novak, E., Daday, J., & McDaniel, K. (2018). Assessing Intrinsic and Extraneous Cognitive Complexity of E-textbook Learning. *Interacting with Computers*, *30*(2), 150–161. https://doi.org/10.1093/iwc/iwy001

Novak, E., McDaniel, K., Daday, J., & Soyturk, I. (2022). Frustration in technology-rich learning environments: A scale for assessing student frustration with e-textbooks. *British Journal of Educational Technology*, *53*(2), 408–431. https://doi.org/10.1111/bjet.13172

Opfermann, M., Schmeck, A., & Fischer, H. E. Multiple Representations in Physics and Science Education – Why Should We Use Them? (D. F. Treagust, R. Duit, & H. E. Fischer, Eds.). en. In: *Multiple Representations in Physics Education* (D. F. Treagust, R. Duit, & H. E. Fischer, Eds.). Ed. by Treagust, D. F., Duit, R., & Fischer, H. E. Models and Modeling in Science Education. Cham: Springer International Publishing, 2017, pp. 1–22. ISBN: 978-3-319-58914-5. https://doi.org/10.1007/978-3-319-58914-5_1.

Oppenheimer, J. R. (1930). Note on the Theory of the Interaction of Field and Matter. *Physical Review*, *35*(5), 461–477. https://doi.org/10.1103/PhysRev.35.461

Organtini, G. (2011). Matter and interactions: A particle physics perspective [Publisher: IOP Publishing]. *Physics Education*, *46*(5), 544–550. https://doi.org/10.1088/0031-9120/46/5/005

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments [Publisher: Lawrence Erlbaum Associates, Inc.]. *Educational Psychologist*. https://doi.org/10.1207/S15326985EP3801_1

Paivio, A. (1990, September). *Mental Representations: A Dual Coding Approach* [Google-Books-ID: hLGmKkh_4K8C]. Oxford University Press.

Park, W., Yang, S., & Song, J. (2019). When Modern Physics Meets Nature of Science. *Science & Education*, *28*(9), 1055–1083. https://doi.org/10.1007/s11191-019-00075-9

Pascolini, A., & Pietroni, M. (2002). Feynman diagrams as metaphors: Borrowing the particle physicist's imagery for science communication purposes [Publisher: IOP Publishing]. *Physics Education*, *37*(4), 324. https://doi.org/10.1088/0031-9120/37/4/306

Passon, O. (2019). On the interpretation of Feynman diagrams, or, did the LHC experiments observe H $\rightarrow$? *European Journal for Philosophy of Science*, *9*(2), 20. https://doi.org/10.1007/s13194-018-0245-1

Passon, O. (2022). The Quasi-History of Early Quantum Theory [Number: 3 Publisher: Multidisciplinary Digital Publishing Institute]. *Physics*, *4*(3), 880–891. https://doi.org/10.3390/physics4030057

Passon, O., Lindenau, P., & Kobel, M. (2020). Von Feynman-Diagrammen und Stromkreisen. *Unterricht Physik*, *180*. Retrieved March 19, 2021, from https://www.friedrich-verlag.de/physik/quantenphysik/von-feynman-diagrammen-und-stromkreisen-7480

Passon, O., Zügge, T., & Grebe-Ellis, J. (2018). Pitfalls in the teaching of elementary particle physics [Publisher: IOP Publishing]. *Physics Education*, *54*(1), 015014. https://doi.org/10.1088/1361-6552/aadbc7

Pavlidou, M., & Lazzeroni, C. (2016). Particle physics for primary schools—enthusing future physicists [Publisher: IOP Publishing]. *Physics Education*, *51*(5), 054003. https://doi.org/10.1088/0031-9120/51/5/054003

Polen, C. (2019). Particle Physics: An Essential and Engaging Part of the High School Physics Program [Publisher: American Association of Physics Teachers]. *The Physics Teacher*, *57*(5), 320–322. https://doi.org/10.1119/1.5098922

Prather, E. E., & Harrington, R. R. (2001). Student understanding of ionizing radiation and radioactivity [Publisher: National Science Teachers Association]. *Journal of College Science Teaching*, *31*(2), 89.

R Core Team. (2023). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Rau, M. A. (2017). Conditions for the Effectiveness of Multiple Visual Representations in Enhancing STEM Learning. *Educational Psychology Review*, *29*(4), 717–761. https://doi.org/10.1007/s10648-016-9365-3

Riva, F., Monin, A., Sonderegger, A., & Stojkovic, L. (2022). Tutti Quantum. Retrieved December 20, 2023, from https://helvetiq.com/ch_en/tutti-quantum

Rodemer, M., Eckhard, J., Graulich, N., & Bernholt, S. (2020). Decoding Case Comparisons in Organic Chemistry: Eye-Tracking Students' Visual Behavior [Publisher: American Chemical Society]. *Journal of Chemical Education*, *97*(10), 3530–3539. https://doi.org/10.1021/acs.jchemed.0c00418

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20–27. https://doi.org/10.1016/j.tics.2010.09.003

Rosenberg, J. M., Kubsch, M., Wagenmakers, E.-J., & Dogucu, M. (2022). Making Sense of Uncertainty in the Science Classroom. *Science & Education*. https://doi.org/10.1007/s11191-022-00341-3

Rosengrant, D., Thomson, C., & Mzoughi, T. (2009). Comparing Experts and Novices in Solving Electrical Circuit Problems with the Help of Eye-Tracking [Publisher: American Institute of Physics]. *AIP Conference Proceedings*, *1179*(1), 249–252. https://doi.org/10.1063/1.3266728

Rouinfar, A., Agra, E., Larson, A. M., Rebello, N. S., & Loschky, L. C. (2014). Linking attentional processes and conceptual problem solving: Visual cues facilitate the automaticity of extracting relevant information from diagrams. *Frontiers in Psychology*, *5*. Retrieved February 20, 2024, from https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2014.01094

Rubbia, C. (1985). Experimental observation of the intermediate vector bosons W-, W+, and Z0 [Publisher: American Physical Society]. *Reviews of Modern Physics*, *57*(3), 699–722. https://doi.org/10.1103/RevModPhys.57.699

Schwartz, R. S., Lederman, N. G., & Crawford, B. A. (2004). Developing views of nature of science in an authentic context: An explicit approach to bridging the gap between nature of science and scientific inquiry. *Science Education*, *88*(4), 610–645. https://doi.org/10.1002/sce.10128

Schweber, S. S. (1994). *QED and the Men Who Made It: Dyson, Feynman, Schwinger, and Tomonaga* (Vol. 104). Princeton University Press. https://doi.org/10.2307/j.ctv10crg18

Schwinger, J. (1948). Quantum Electrodynamics. I. A Covariant Formulation [Publisher: American Physical Society]. *Physical Review*, *74*(10), 1439–1461. https://doi.org/10.1103/PhysRev.74.1439

Schwinger, J. (1949a). Quantum Electrodynamics. II. Vacuum Polarization and Self-Energy [Publisher: American Physical Society]. *Physical Review*, *75*(4), 651–679. https://doi.org/10.1103/PhysRev.75.651

Schwinger, J. (1949b). Quantum Electrodynamics. III. The Electromagnetic Properties of the Electron—Radiative Corrections to Scattering [Publisher: American Physical Society]. *Physical Review*, *76*(6), 790–817. https://doi.org/10.1103/PhysRev.76.790

Seskir, Z. C., Migdał, P., Weidner, C., Anupam, A., Case, N., Davis, N., Decaroli, C., Ercan, Foti, C., Gora, P., Jankiewicz, K., Cour, B. R. L., Malo, J. Y., Maniscalco, S., Naeemi, A., Nita, L., Parvin, N., Scafirimuto, F., Sherson, J. F., … Chiofalo, M. (2022). Quantum games and interactive tools for quantum technologies outreach and education [Publisher: SPIE]. *Optical Engineering*, *61*(8), 081809. https://doi.org/10.1117/1.OE.61.8.081809

Sexl, R., Kühnelt, H., Stadler, H., Jakesch, P., & Sattlberger, E. (2019). *Sexl Physik 8*. öbv Österreichischer Bundesverlag Schulbuch GmbH & Co. KG. Retrieved April 25, 2022, from https://www.oebv.at/produkte/sexl-physik-8-schulbuch-0

Shannon, C. E. (1948). A mathematical theory of communication [Conference Name: The Bell System Technical Journal]. *The Bell System Technical Journal*, *27*(3), 379–423. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shic, F., Chawarska, K., Bradshaw, J., & Scassellati, B. Autism, eye-tracking, entropy. en. In: In *2008 7th IEEE International Conference on Development and Learning*. Monterey, CA: IEEE, 2008, August, 73–78. ISBN: 978-1-4244-2661-4. https://doi.org/10.1109/DEVLRN.2008.4640808.

Shifman, M. (1995, October). ITEP Lectures in Particle Physics [arXiv:hep-ph/9510397]. https://doi.org/10.48550/arXiv.hep-ph/9510397

Stöltzner, M. (2018). Feynman Diagrams: Modeling between Physics and Mathematics. *Perspectives on Science*, *26*(4), 482–500. https://doi.org/10.1162/posc_a_00284

Stückelberg, E. C. C. (1942). La mécanique du point matériel en théorie de la relativité et en théorie des quanta [ISBN: 0018-0238]. *Helvetica physica acta*, *15*, 23–37.

Susac, A., Planinic, M., Bubic, A., Ivanjek, L., & Palmovic, M. (2020). Student recognition of interference and diffraction patterns: An eye-tracking study [Publisher: American Physical Society]. *Physical Review Physics Education Research*, *16*(2), 020133. https://doi.org/10.1103/PhysRevPhysEducRes.16.020133

Susac, A., Planinic, M., Bubic, A., Jelicic, K., Ivanjek, L., Matejak Cvenic, K., & Palmovic, M. (2021). Effect of students' investigative experiments on students' recognition of interference and diffraction patterns: An eye-tracking study [Publisher: American Physical Society]. *Physical Review Physics Education Research*, *17*(1), 010110. https://doi.org/10.1103/PhysRevPhysEducRes.17.010110

Sutton, C., & Zerwas, P. (2004). The W and Z at LEP. *CERN Courier*. Retrieved February 5, 2024, from https://cerncourier.com/a/the-w-and-z-at-lep/

Sweller, J. (1993). Some cognitive processes and their consequences for the organisation and presentation of information. *Australian Journal of Psychology*, *45*(1), 1–8. https://doi.org/10.1080/00049539308259112

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction*, *4*(4), 295–312. https://doi.org/10.1016/0959-4752(94)90003-5

Sweller, J. (2010). Element Interactivity and Intrinsic, Extraneous, and Germane Cognitive Load. *Educational Psychology Review*, *22*(2), 123–138. https://doi.org/10.1007/s10648-010-9128-5

Swinbank, E. (1992). Particle physics: A new course for schools and colleges [Publisher: IOP Publishing]. *Physics Education*, *27*(2), 87–91. https://doi.org/10.1088/0031-9120/27/2/006

Tanedo, F. (2012, March). Dissecting the Penguin. Retrieved March 1, 2024, from http://www.quantumdiaries.org/2012/03/19/dissecting-the-penguin/

Tobii. (2023, June). Tobii Customer Portal. Retrieved February 22, 2024, from https://connect.tobii.com/s/article/Gaze-Filter-functions-and-effects?language=en_US

Tobii AB. (2023). Tobii Pro Lab [Place: Danderyd, Stockholm Version Number: 1.217]. https://www.tobii.com/

Tobii Pro AB. (2019). Tobii Pro Fusion eye tracker [Place: Danderyd, Stockholm Type: Appartus Published: Apparatus]. https://www.tobii.com/

Tomonaga, S. (1943). On a Relativistic Reformulation of Quantum Field Theory. *Bull. IPCR (Riken-iho)*, *22*, 545–557.

Tomonaga, S. (1946). On a Relativistically Invariant Formulation of the Quantum Theory of Wave Fields. *Progress of Theoretical Physics*, *1*(2), 27–42. https://doi.org/10.1143/PTP.1.27

Tuzón, P., & Solbes, J. (2016). Particle Physics in High School: A Diagnose Study [Publisher: Public Library of Science]. *PLOS ONE*, *11*(6), e0156526. https://doi.org/10.1371/journal.pone.0156526

van Gog, T., Paas, F., van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the Problem-Solving Process: Cued Retrospective Reporting Versus Concurrent and Retrospective Reporting [Place: US Publisher: American Psychological Association]. *Journal of Experimental Psychology: Applied*, *11*(4), 237–244. https://doi.org/10.1037/1076-898X.11.4.237

van Kampen, N. G., & Meijer, P. H. E. (2000). *Views of a Physicist: Selected Papers of N.G. Van Kampen* [Google-Books-ID: xSWx0216qHUC]. World Scientific.

Veritasium. (2013, June). Can We Really Touch Anything? Retrieved February 29, 2024, from https://www.youtube.com/watch?v=bKldI-XGHIw

Wagner, W. Quarks und Feynman-Diagramme: Basiswissen im Schulunterricht? (O. Passon, T. Zügge, & J. Grebe-Ellis, Eds.). de. In: *Kohärenz im Unterricht der Elementarteilchenphysik: Tagungsband des Symposiums zur Didaktik der Teilchenphysik, Wuppertal 2018* (O. Passon, T. Zügge, & J. Grebe-Ellis, Eds.). Ed. by Passon, O., Zügge, T., & Grebe-Ellis, J. Berlin, Heidelberg: Springer, 2020, pp. 3–11. ISBN: 978-3-662-61607-9. https://doi.org/10.1007/978-3-662-61607-9_1.

Weingard, R. Virtual particles and the interpretation of quantum field theory (H. R. Brown & R. Harré, Eds.). In: *Philosophical foundations of quantum field theory* (H. R. Brown & R. Harré, Eds.). Ed. by Brown, H. R., & Harré, R. New York: Oxford University Press, 1988, pp. 43–58. ISBN: 0-19-824975-6.

Welch, B. L. (1947). The Generalization of 'Student's' Problem when Several Different Population Variances are Involved [Publisher: [Oxford University Press, Biometrika Trust]]. *Biometrika*, *34*(1/2), 28–35. https://doi.org/10.2307/2332510

Wertheimer, M. Laws of organization in perceptual forms. In: In *A source book of Gestalt Psychology*. Harcourt Brace, 1923, pp. 71–88.

Wiener, G. (2017). *Elementary particle physics in early physics education* (Doctoral dissertation). Universität Wien. Wien. https://utheses.univie.ac.at/detail/41103#

Wiener, G. J., Schmeling, S. M., & Hopf, M. (2015). Can grade-6 students understand quarks? Probing acceptance of the subatomic structure of matter with 12-year-olds [Publisher: Bastas]. *European Journal of Science and Mathematics Education*, *3*(4), 313–322. https://doi.org/10.30935/scimath/9440

Wiener, G. J., Schmeling, S. M., & Hopf, M. (2017a). An Alternative Proposal for the Graphical Representation of Anticolor Charge [Publisher: American Association of Physics Teachers]. *The Physics Teacher*, *55*(8), 472–474. https://doi.org/10.1119/1.5008340

Wiener, G. J., Schmeling, S. M., & Hopf, M. (2017b). Introducing 12 year-olds to elementary particles [Publisher: IOP Publishing]. *Physics Education*, *52*(4), 044001. https://doi.org/10.1088/1361-6552/aa6cfe

Wiener, G. J., Woithe, J., Brown, A., & Jende, K. (2016). Introducing the LHC in the classroom: An overview of education resources available [Publisher: IOP Publishing]. *Physics Education*, *51*(3), 035001. https://doi.org/10.1088/0031-9120/51/3/035001

Wilhelm, T., & Hopf, M. Design-Forschung (D. Krüger, I. Parchmann, & H. Schecker, Eds.). de. In: *Methoden in der naturwissenschaftsdidaktischen Forschung* (D. Krüger, I. Parchmann, & H. Schecker, Eds.). Ed. by Krüger, D., Parchmann, I., & Schecker, H. Berlin, Heidelberg: Springer, 2014, pp. 31–42. ISBN: 978-3-642-37827-0. https://doi.org/10.1007/978-3-642-37827-0_3.

Wittrock, M. C. (1989). Generative Processes of Comprehension [Publisher: Routledge _eprint: https://doi.org/10.1207/s15326985ep2404_2]. *Educational Psychologist*, *24*(4), 345–376. https://doi.org/10.1207/s15326985ep2404_2

Woithe, J. (2020, August). *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) [Number: CERN-THESIS-2020-089 Publication Title: CERN Document Server]. Kaiserslautern U. Retrieved December 9, 2021, from https://cds.cern.ch/record/2727453

Woithe, J., Boselli, M., Chatzidaki, P., Dahlkemper, M. N., Duggan, R., Durey, G., Herff, N., Horvat, A. K., Molaro, D., Scheerer, G. W., Schmeling, S., Thill, P. G., Wiener, J., & Zoechling, S. (2022). Higgs in a Box: Investigating the Nature of a Scientific Discovery [Publisher: World Scientific Publishing Co.]. *The Physics Educator*, *04*(04), 2250019. https://doi.org/10.1142/S2661339522500196

Woithe, J., Schmidt, R., & Naumann, F. (2019). Track inspection: How to spot subatomic particles. *Science in School*, (46). https://www.scienceinschool.org/article/2019/track-inspection-how-spot-subatomic-particles/

Woithe, J., Wiener, G. J., & van der Veken, F. F. (2017). Let's have a coffee with the Standard Model of particle physics! [Publisher: IOP Publishing]. *Physics Education*, *52*(3), 034001. https://doi.org/10.1088/1361-6552/aa5b25

Wolchover, N. (2020, November). What Is a Particle? Retrieved February 28, 2024, from https://www.quantamagazine.org/what-is-a-particle-20201112/

Wolchover, N. (2021, April). 'Last Hope' Experiment Finds Evidence for Unknown Particles. Retrieved August 8, 2022, from https://www.quantamagazine.org/last-hope-experiment-finds-evidence-for-unknown-particles-20210407/

Wüthrich, A. Introduction: Origin, Use and Interpretation of Feynman Diagrams (A. Wüthrich, Ed.). en. In: *The Genesis of Feynman Diagrams* (A. Wüthrich, Ed.). Ed. by Wüthrich, A. Archimedes. Dordrecht: Springer Netherlands, 2011, pp. 1–21. ISBN: 978-90-481-9228-1. https://doi.org/10.1007/978-90-481-9228-1_1.

Wüthrich, A. (2012). Interpreting Feynman Diagrams as Visual Models [Number: 1]. *Spontaneous Generations: A Journal for the History and Philosophy of Science*, *6*(1), 172–181. https://doi.org/10.4245/sponge.v6i1.16131

Yukawa, H. (1935). On the Interaction of Elementary Particles. *Proceedings of the Physical-Mathematical Society of Japan*, *17*, 48–57. https://www.jstage.jst.go.jp/article/ppmsj1919/17/0/17_0_48/_pdf/-char/en

Zaman Ahmed, S., Jensen, J. H. M., Weidner, C. A., Sørensen, J. J., Mudrich, M., & Sherson, J. F. (2021). Quantum composer: A programmable quantum visualization and simulation tool for education and research. *American Journal of Physics*, *89*(3), 307–316. https://doi.org/10.1119/10.0003396

't Hooft, G., & Veltman, M. Diagrammar (D. Speiser, F. Halzen, & J. Weyers, Eds.). en. In: *Particle Interactions at Very High Energies: Part B* (D. Speiser, F. Halzen, & J. Weyers, Eds.). Ed. by Speiser, D., Halzen, F., & Weyers, J. NATO Advanced Study Institutes Series. Boston, MA: Springer US, 1974, pp. 177–322. ISBN: 978-1-4684-2826-1. https://doi.org/10.1007/978-1-4684-2826-1_5.