

De- and Reconstructing Physics Education

A Female Gaze

Contents

1. Introduction	2
2. Interdisciplinarity	2
2.1. Definitions.....	3
2.2. Disciplinary perspectives in Analysis, Visions and Voices.....	5
2.3. The lens behind the project.....	12
3. Critical Reflections.....	15
3.1. Analysis.....	15
3.2. Voices	19
3.3. Visions.....	22
4. Conclusion	26
Bibliography	28

1. Introduction

In such moments, she'd think of [...] the surprise of her peers upon seeing her on the first day of the third semester. "You're still here?", one had commented with amazement so genuine she almost didn't register it as condescending.

(“Eva”, p. 6)

This is an excerpt from a short story that was produced specifically for this project, making experienceable the impact of an interconnected web of bias and obstacles that women in physics face to this day. Alongside analysis, other narratives and foresight, it constitutes the project [“Physics Education: A Female Gaze”](#) and investigates women’s underrepresentation in physics and the role of education in it. The core of the project is presented to the public on a website, making this report a complement to the contents presented there. The website is both a means and an end, it is a coherent system in and of itself that serves as a bridge between research, public discourse and education practice. It is its own contribution to the co-creation of physics education futures that it invites and empowers others to participate in as well.

Before continuing with this report, I would encourage readers to engage with the contents of the website. In the following sections, I will refer to specific parts of the site and complement the public contents with further academic framing. This includes theoretical considerations on interdisciplinarity, clarifying key terms and concepts in their application to my work, as well as a discussion of various disciplinary perspectives as present in the project. Further elaborating on the contributions and challenges of the work, I will draw on evaluative frameworks from the literature to reflect critically on the three parts of the project, concluding with its strengths, limitations and indications for future research.

2. Interdisciplinarity

The project in its set up, research questions, goals and audiences carries a complexity with inherent conceptual challenges and dilemmas. To give but one example at this point, a question as simple as how to cite sources in this project has led me to adapt rather than simply use an established citation style. The goal is to maintain the academic integrity of this project while making the contents on the website accessible to non-academic audiences as well. I followed the guidelines by the American Psychological Association, a style that is generally used in education

research (*About APA Style*, 2023), except for the use of footnotes: On the website, I chose to cite academic sources in footnotes, rather than through in-text citations, for a smoother reading experience that would appeal to non-academic readers as well – which is non-conforming with APA guidelines (*Concise Guide to APA Style*, 2020). This report, on the other hand, is directed entirely at academic readers and follows all APA style guidelines, including in-text citations. This choice may well be criticized as inconsistent. I would, however, consider this slight inconsistency between the website and this report acceptable as a deliberate, self-reflective choice bearing the different requirements for the website and the report.

Other conceptual challenges throughout the project are addressed by integrating theories, concepts and methods from various disciplines, making interdisciplinarity a key characteristic of the website. As such, the following sections are dedicated to defining core concepts and exploring why, where and how different disciplines contribute to and connect within the project.

2.1. Definitions

For an understanding of different disciplinary perspectives in this project, first a definition of a discipline itself is needed. Menken and Kestra (2016) make an attempt to do so and

“define a discipline as a field of science with a particular object of research and a corresponding body of accumulated specialist knowledge. This knowledge is effectively organized by and expressed through theories, concepts, and assumptions inclusive of its discipline-specific terminologies and technical language. Furthermore, a discipline has its own specific research methods and has an institutional manifestation (usually at universities).” (p. 27)

This definition is a useful starting point in that it highlights how disciplines express their knowledge in different theories and concepts and use certain methods to investigate their object of research. In this project, I drew precisely on such theories, concepts and methods from different disciplines to solve conceptual challenges. Menken and Kestra (2016) go on to classify academic disciplines into natural sciences, social sciences and humanities, and provide an overview of major disciplines in this classification. I will use their terminology for some of the disciplines I drew on in this project, such as educational studies, psychology and sociology from the social sciences; however, other disciplines from the humanities that I integrated in my work, like literature and film studies, are not contained in their high-level overview.

With this delineation of separate disciplines, a definition of interdisciplinarity is now in order. The National Academy of Sciences (2005) offers one of the most widely used definitions, stating that interdisciplinary research integrates “information, data techniques, tools, perspectives, concepts and/or theories from two or more disciplines [...], with the objective to advance fundamental understanding or to solve problems whose solutions are beyond the scope of a single discipline or area of research practice”. The first part of this definition describes adequately my work in integrating aspects of various disciplines. In addition to ‘advancing understanding’ and ‘solving a problem’, the objective for this interdisciplinarity in my project is also the “construction of a more comprehensive perspective” (Newell & Klein, 1997, p. 394).

Both definitions, by the National Academy of Sciences (2005) and the one provided by Newell and Klein (1997), include the word “integrate” to define interdisciplinarity. This becomes an important detail in distinguishing between related concepts like multi- and transdisciplinarity. The first is characterized precisely by the absence of integration, that is, it involves more than one discipline, but does not integrate the results (Menken & Keestra, 2016, p. 32). Transdisciplinarity, on the other hand, takes interdisciplinarity one step further and includes collaboration with stakeholders outside the academic world, integrating disciplinary and non-academic knowledge (Menken & Keestra, 2016, p. 32). At first glance, this project then seems transdisciplinary rather than interdisciplinary; however, even though the website is designed specifically for this broadening of academic *and* non-academic audiences, the work itself was not carried out as a collaboration. Menken and Keestra’s (2016) definition emphasizes the aspect of a joint process where decisions are made together with academic and non-academic stakeholders, which was not the case here. That being said, the website may well serve as a basis for *future* transdisciplinary research, connecting academic and non-academic readers.

Having thus established and clarified the interdisciplinarity of this project, there remains a fundamental question: Why choose an interdisciplinary approach for this work? It arguably adds further complexity in an attempt to bring together disciplines as varied as educational studies, psychology, sociology, literature, film studies and futures studies.

Brinkmann (2020) identifies four drivers of interdisciplinary research, and indeed two of them serve to justify the need for interdisciplinarity in this project, namely the inherent complexity of nature and society, and the need to solve societal problems (pp. 38-39). As for the first driver, it is clear that the project is concerned with a topic deeply rooted in society as a complex system. Inequity in physics education cannot be understood in disciplinary isolation of, say, educational

studies investigating curricular contents and contexts; disadvantages experienced in physics by different groups, especially women, are part of this inherently complex web of interactions within society. Secondly, the need to solve societal problems further drives interdisciplinarity in this project. Menken and Keestra (2016) assert that “there is an expectation that scientific research will, in some way, contribute to solving important societal issues” (p. 38). The project attempts to adhere to this expectation and contribute to dismantling barriers for women and other minorities in physics, thriving for more equity. Such societal problems, however, are often highly complex, in some cases even wicked, and it is important to recognize that any intervention might have unexpected consequences. This bi-directional relationship between ‘fixes’ and problems is another reason why addressing those complex interactions requires an interdisciplinary approach.

2.2. Disciplinary perspectives in Analysis, Visions and Voices

With this understanding of disciplines and interdisciplinarity, I have integrated methods, theories and concepts from various disciplines, merging and adapting them to adequately address the research questions posed by the three parts of the project.

The first part, [Analysis](#), rests on a systematic literature review, a method that stems originally from the field of medical research (Booth et al., 2016) but has since evolved and is now practiced in many different disciplines, including educational studies (Newman & Gough, 2020). I considered the systematic literature review a useful method for the objective of this first part, namely to provide an overview of the research on the topic¹. This “secondary level analysis [...] that brings together the findings of primary research” (Newman & Gough, 2020, p. 5) is precisely what I set out to do: Bringing such primary research to a diverse audience through synthesis.

However, the systematic review uses “explicit, accountable, rigorous research methods” (Gough et al., 2017, p. 4) and is usually characterized by elements such as precise selection and exclusion criteria, a clear and documented search strategy and an assessment of the quality of the studies (Newman & Gough, 2020). All of these are elements that make perfect sense in the context of

¹ At this point, this is a rather vague objective. The more precise research questions are formulated later following the general process of a systematic literature review.

meta-analyses in medical research, and indeed for some topics in educational studies as well, but are too restrictive for my endeavor. This is why I would not call the Analysis a 'systematic review', but rather a review with methodological contributions from and some adaptations of a systematic review.

The formulation of a concise research question stands at the very beginning of this process (Newman & Gough, 2020). In purely systematic reviews, this can be a semi-closed question that invites precise answers, for example *What is the effect of flipped learning versus traditional learning on student achievement?*, categorizing findings from the review into pre-determined answers, for example as 'More effective', 'less effective' or 'no difference' (Lo, 2020). In contrast, my research questions are deliberately open, inviting a range of different answers that are not predetermined: *What do statistics tell us about female participation in physics? What role do learners, teachers, curricula and cultural narratives play?* These questions also invite answers that go beyond the strict scope of educational research, for example asking about cultural narratives touches upon sociological questions – but as section 3.1. will examine in more detail, such broadness is deliberate and necessary for this project to achieve greater interpretative depth.

Suggestions for search methods from the literature on systematic reviews such as citation checking, contact with experts and bibliographic database search (Newman & Gough, 2020; Booth et al., 2016) have been valuable for my search strategies as well. However, for the purpose of making the information from primary research accessible to various audiences, it did not seem necessary to document all search results systematically. Similarly, I did not define a precise set of inclusion and exclusion criteria, such as a time frame for publication or a minimum number of study participants. The scope I defined in the [Project](#) section of the website invites both quantitative and qualitative research, as well as various levels of evidence, from conference reports to meta-analyses. This scope is also at times consciously expanded in the review, for example including information about intersectionality and other minorities in physics despite the focus on gender-related issues.

The systematic review offers a toolkit of methods for presenting findings, some of which have informed my output. The overall visual presentation, for example, draws on the idea of a concept map, "a picture of the territory under study, and [representing] the concepts in that area and the relationships between them" (Rowley and Slack, 2004, p. 36). However, I adapted this method to show more than just concepts. On a higher level, visually presented with darker background,

white text and larger font, the concepts are depicted in relationship to each other, for example mindsets, identity, self-efficacy and sense of belonging as concepts with relation to 'Learners' – a relationship that is visually expressed through color-coding (s. Figure 1). On a more detailed level below the concepts, my application of this method also includes, for example, the results of studies. For this level of information, I made use of a different method from systematic reviews: A narrative synthesis, which “adopts a textual approach to the process of synthesis to ‘tell the story’ of the findings from the included studies” (Popay et al., 2006, p. 5). This story-telling aspect is featured in my presentation of the results below the high level of concepts, for example narrating the chronology of conflicting findings on interactive teaching and its effects for women in physics, or the dialogue style of question-and-answer boxes (s. Figure 2). However, this approach “relies primarily on the use of words and text to summarise [sic] and explain the findings” (Popay et al., 2006, p. 5), whereas my synthesis includes visual storytelling in addition to this purely text-based approach, as also evidenced in Figure 2.

Further expanding this narrative synthesis approach into an entirely different discipline led me to produce a data story from statistical data. Upon inquiry, I had received the raw datasets on women in physics from the American Physical Society and used Microsoft Excel to produce meaningful charts to visualize the data. However, a simple graph is not sufficient to make the data *meaningful*, so I devised a variety of different visualizations from the same dataset with varying parameters, and connected the charts in a guiding narrative to demonstrate how the same data can tell different stories (s. Figure 3).

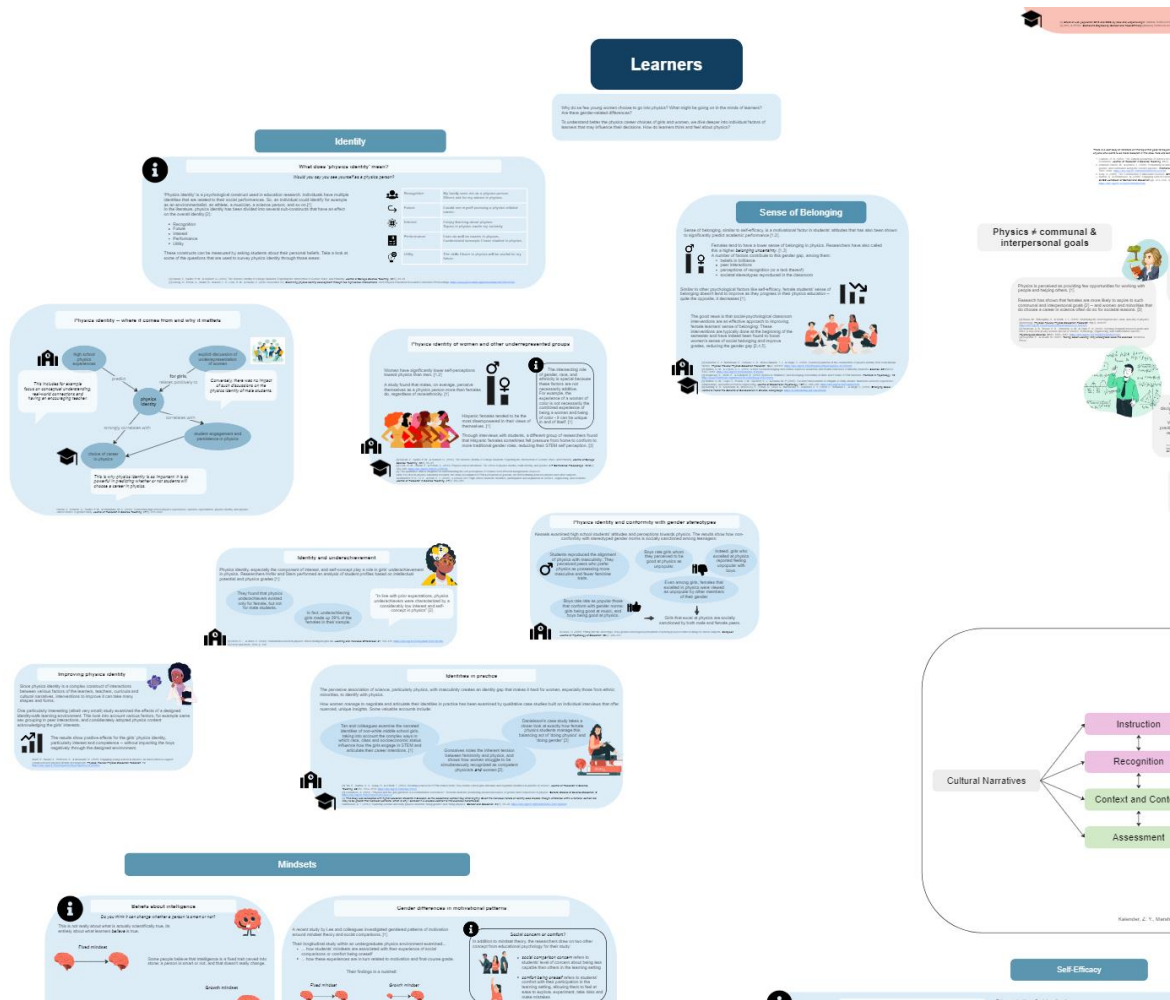


Figure 1: Example of the application of a concept map, with darker boxes, white text and larger font representing concepts and relations as color coded.

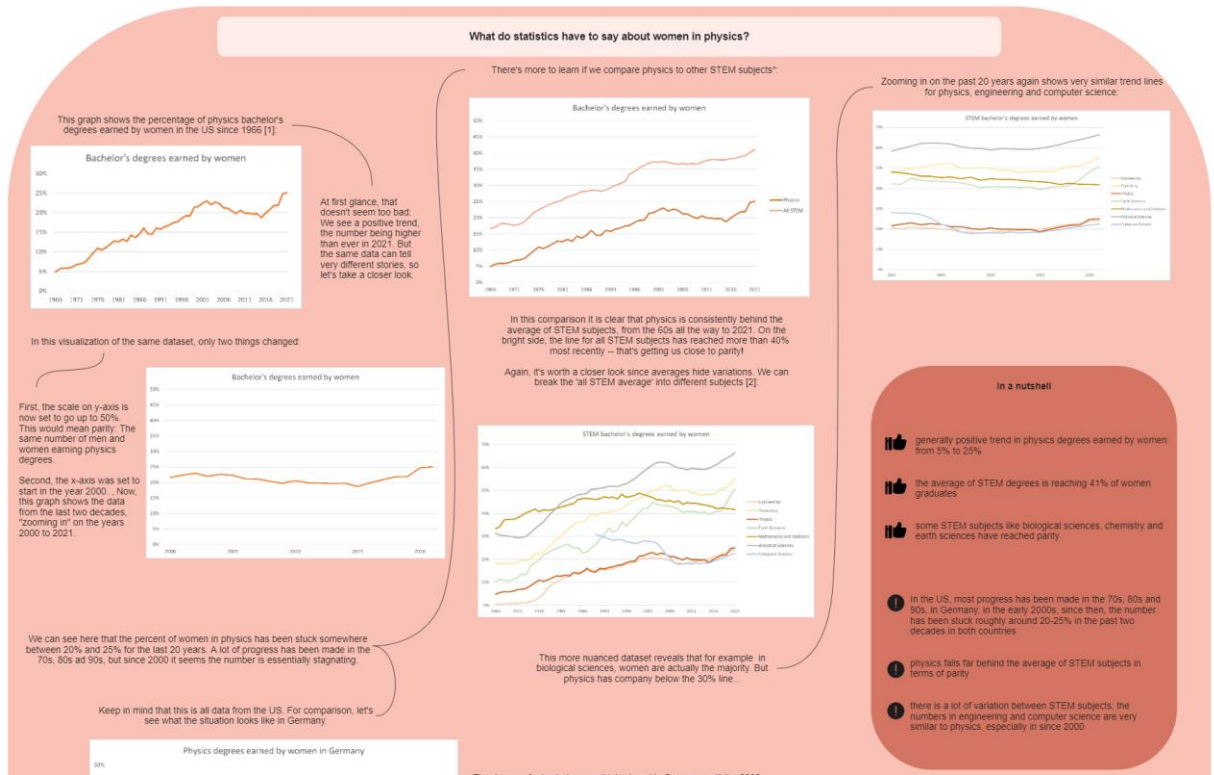


Figure 3: A data story from the APS's statistics on women in physics.

The second part, [Voices](#), asks how the factors examined academically in the Analysis manifest in practice. It seeks to make the abstract findings of studies accessible and personal. For this objective, part two starts out with research interviews, a method that is not specific to one discipline but rather “a key method in the human and social sciences and in many other corners of the scientific landscape, from education to the health sciences” (Brinkmann, 2020, p. 424). Such qualitative research interviews are characterized by three main features, according to Gillham (2005): First, the questions and topics are ‘open’ in the sense that the interviewee is determining their own answers; Second, the relationship between researcher and interviewee allows certain ‘adjustment’ as it is interactive and responsive; Third, the interviewer follows a structure and purpose (pp. 3-4). These three characteristics are precisely why research interviews emerged as a fitting method to illustrate on an individual, personal level what, for example, low self-efficacy beliefs signify in context.

For the five interviews I conducted with an activist, a physicist, two students and a teacher, I followed Gillham's (2005) guidance for question and topic development in the preparatory phase. As a first step, I brainstormed potential questions, proceeded to group similar ones into wider topics, and finally organized those in a logical 'narrative' order. While there are some similarities between the prepared questions, each of the interviewees brought a slightly different perspective to the topic and required an individualized approach in preparing for the interview. For example, the teacher interview covered topics such as classroom interactions, student behavior, assessment and curricula – the purple and some green parts in the mind map – whereas the students opened up about their own self-efficacy beliefs, sense of belonging and physics identity – the blue parts – and yet again with a different perspective, the activist and physicist interviews brought more personal insights about representation, stereotypes and cultural narratives – the red and gray part of the mind map.

Two of the interviews are accessible directly as slightly edited recordings on the website. Those were conducted with what I called 'experts' – the activist and the physicist. The framing of the interviews as podcast episodes specifically for a public audience is a deviation from the most traditional sense of research interviews, where data confidentiality and protection of identity information is usually valued (Gillham, 2005, Chapter 2). For this project, the public accessibility of the interviews forms an important part of the objective, which is why informed consent was obtained from the expert participants to publish their edited interviews on the website.

The case of the other three interviews, namely the two students and the teacher, is very different. As vulnerable groups – two young adults and a teacher sharing stories about minors in her own teaching practice – the contents of their interviews could not be published directly. Instead, I crafted fictionalized narratives in the form of two short stories from the contents of the interviews. Clough (2002) exemplifies this method in his deliberately non-prescriptive book on narratives and fictions in education research, building upon prior work in sociology by Richardson (1994) and in education by St. Pierre (1997). The theoretical framing for this method as well as a detailed analysis of my application is given for a broader audience on the page [Narratives and Fictions in Education Research](#).

The third part, [Visions](#), makes extensive use of methods and theories from the field of futures studies. A disciplinary homeland for futures studies is challenging to determine for a fundamentally interdisciplinary field that has even been contested as a separate field or discipline at all (Marien, 2002). Broadly, Bell (2003) views futures studies as part of the social sciences,

arguing that decision-making and action are social processes. Kristóf (2023) specifies this classification further and describes futures studies as rooted in sociology and policy sciences. For understanding why and how I chose specific methods and theories from this field, Son's (2015) sketch of (western) futures studies in three phases constitutes a helpful basis: The first period of mid-20th century futurists was mainly concerned with technological advancements, aiming to forecast visions of the future; since the 1970s, the second phase focused more and more on global business interests, whereas the third period of neoliberal views begins in the 90s and is marked by an intensified fragmentation of futures.

The work I aim to do in this project in general and in Visions specifically displays the same characteristics as the third period of futures studies. The core question *Where do we go from here?* is not concerned with technological advancements, forecasts or business trends – it invites greatly diverging, fragmented answers, with the precise goal of producing various images of futures from and with a diverse audience, through a multitude of methods. Indeed, the previously described fragmentation of futures studies in the third period also refers to “the lack of disciplinary consensus” (Son, 2015, p. 128), which evidently is pronounced in this project as well. For this objective, I draw on Voros (2003) widely cited and accepted generic foresight process framework (see for example Nemeth et al., 2018; Riddell et al., 2020; Fiel'ardh, 2024; Directorate-General for Research and Innovation (European Commission) & Soeiro de Carvalho, 2023) and adapt it slightly for the purposes of the project. His theory and terminology of the ‘cone of futures’ frames the activity I invite my audience to participate in, incorporating elements of diverse futuring methods such as the Delphi method, brainstorming, wildcards and scenarios. The precise application and adaptation of those is explained in greater detail for the public audience on the page [Futures and Foresight](#). Notably, Voros' generic foresight process spans across all three parts of the project, connecting Analysis and Voices to Visions as precursors for the prospection activity. This connection is expressed on the above referenced theory page, as well as the [Recap](#) page preceding the prospection in Visions.

2.3. The lens behind the project

In the face of such complexity in theories and methods drawn upon for the various parts of the project, I introduce yet another theory from a different discipline that connects the three parts as

a lens, a perspective from which to examine the topic – and again, I adapt and develop it further for an application outside its discipline of origin.

In film studies, the theory of the female gaze emerged as a response to Laura Mulvey's (1975) feminist critique of the male gaze in mainstream Hollywood films. As explained in more detail on the corresponding page of the website, [The Female Gaze](#), the concept encompasses more than the literal gaze and its connection to sexual objectification; it can be understood as a lens, a certain perspective that impacts the levels of camera, plot and audience. Re-stating the concept in broader terms that make it applicable to other fields, I formulated a basic question at the heart of investigations around the (fe)male gaze: *What stories are told, and how?* This formulation breaks the concept into two constituents that have different answers depending on the perspective taken – *what* and *how*. Where the male gaze tells a particular kind of story in a particular way, as analyzed and exemplified on the website, the female gaze, in my own words, tells different stories, and tells stories differently – a formulation I chose in analogy to St. Pierre (s. footnote 11 on the website page). This 'other way of looking' (*Female Gaze*, n.d.), which defines itself mainly in opposition to a certain norm, governs the content (*What*) and method (*How*) throughout the three parts.

Concerning the first constituent, *What*, the content of the website seeks to apply the female gaze to tell different stories. 'Different' in this context refers to a distinctive 'otherness', a challenge of prevalent gendered narratives about physics. Such common stories about physics are exposed in a dedicated section in the [Analysis](#) part, for example stereotypes and cultural narratives that position science as a male domain, hold that physics requires innate brilliance for success, and associate such genius generally with white males. *Finding stories* that gaze upon physics education from a female point of view, the mind map presents research that challenges such stereotypes and paints a complex picture of the web of interactions that contribute to the underrepresentation of women in physics. The podcast episodes, then, *tell stories* of female experiences in physics that thematize problematic narratives and offer alternatives through individual perspectives. Similarly, the prospection exercise is an invitation to readers to *create new stories* for physics education futures. It becomes clear that, throughout the three parts on the website, the 'different stories' aim to produce explicit counternarratives.

The second constituent, *How*, takes the female gaze as an analogy on a methodological level. In film studies, examples of movies that exhibit a female gaze appeal to different audiences than the male-centric Hollywood blockbuster. To give but one example, *Forbidden Love* is analyzed by

Dirse (2013) as a movie made by lesbian directors for lesbian audiences. This consideration for the audience is also what governed my methodological choices in this project, although the simple by-for-relationship does not hold for application here – I may be a female researcher, but the project is not just for female researchers. In fact, quite the contrary: It is the explicit goal to reach a public, diverse audience, male and female, academic and non-academic. As such, the ambition to 'tell stories differently' in this case refers to the balancing act of adhering to academic standards in quality of research while designing it in an accessible fashion, employing academic methods but adapting every single one for the specific needs of this project with a diverse audience. In this sense, the subversion of the female gaze is pronounced in my work not in appealing to an altogether *different* audience, but rather a much larger audience than conventional methods would allow. Importantly, and in differentiation to the female gaze in film, the gender-perspective thus becomes secondary, if not irrelevant, in this analogy on a methodological level; the methods chosen are not selected in order to appeal to females instead of males, but (hopefully) both, although perhaps producing different reactions.

Such an application and adaptation of a theory from film studies to a vastly different discipline is undoubtedly ambitious and not without conflict². Nonetheless, I believe it lends itself to such a disciplinary and theoretical stretch, as French (2021) reminds us that “the female gaze is not homogeneous, singular or monolithic, and it will necessarily take many forms” (pp. 53-54). For this project, it takes the form of a unifying lens backgrounding three methodologically diverse parts.

As a summary, the connections of methods and theories drawn upon in the three parts of the project are visualized in Figure 4.

² For example, the abandonment of the gender perspective for methodological choices as touched upon above. On the website, the [corresponding page](#) also highlights recent criticism of the concept in its narrow focus on gender as a binary opposition.

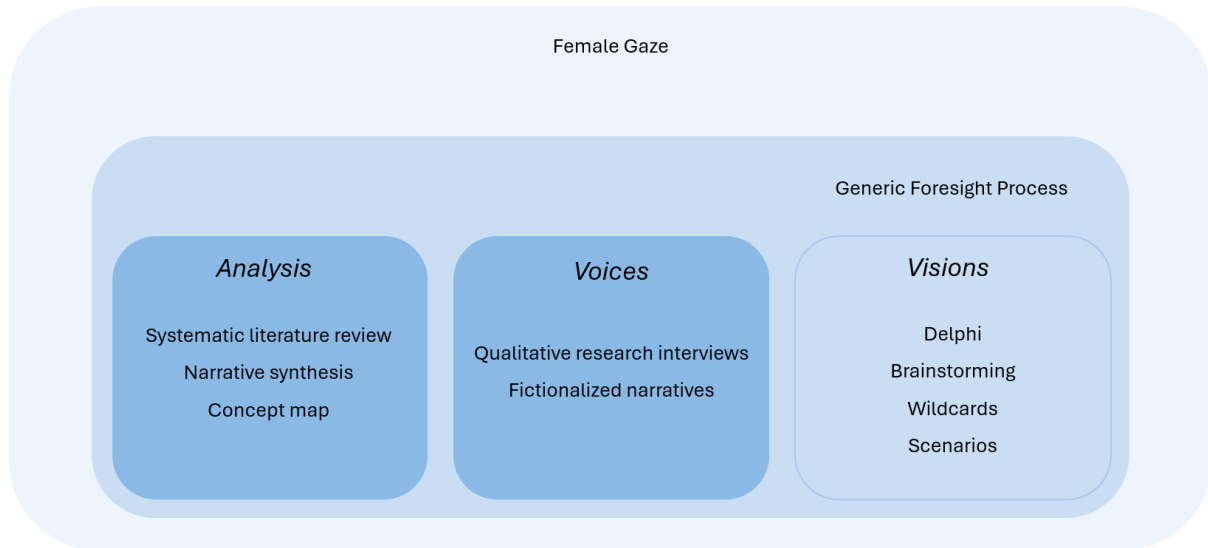


Figure 4: An overview of methods and theories and their connections throughout the project.

3. Critical Reflections

Having established the interdisciplinary approaches and connections of Analysis, Voices, and Visions, the focus now shifts to a more detailed examination of each component. This section will critically reflect on the strengths and weaknesses of these parts individually, drawing on evaluative frameworks from the literature to provide a nuanced understanding of their contributions, challenges and limitations.

3.1. Analysis

As discussed in section 2, the Analysis presented on the website forms part of the foresight in pulling together the stages of input, analysis and interpretation described in Voros' (2003) generic foresight process framework. For this stage, he suggests methods like Inayatullah's (1998) causal layered analysis and depth approaches like Slaughter's (2002) layering of futures-thinking, but then in a later publication offers a five-layer framework – incorporating both previously suggested approaches and more (Voros, 2006, p. 46). These five layers of interpretive depth are summarized in Table 1.

1) Event	Discrete events and occurrences
2) Trend	Patterns and trends
3) System	System drivers, social causes, policy analysis
4) Worldview	Mental models, discourses, perspectives, myths, metaphors, symbols, structures
5) Historical	Social change and related factors and forces, (macro-) historical factors and forces

Table 1: Layering schema of interpretative depth (Voros, 2006, p. 46).

While Voros uses this layering schema of interpretive depth as a basis to analyze and classify other prospective methods, for this report it serves as a framework to guide critical reflections on the work in the Analysis part of the website. Classifying the contents of the mind map along the proposed categories reveals that the analysis and interpretation conducted at this stage extends from the first to the fourth layer.

Discrete *events* are presented through direct quotes of women narrating occurrences, for example recalling experiences of microaggressions (s. Figure 5). Such perceived to be isolated instances are shown to follow *trends* and patterns, for example by discerning in detail the data on women's representation in physics and embedding the direct quotes in more research findings. The *system* of physics education is examined more closely to reveal on an even deeper level the systemic structure behind the events and trends, for example by shining light upon systemic bias in physics tests and the impacts of curricular contents and contexts on female interest and participation in physics. Social causes like teacher or peer behavior in the classroom, impacting the learners, are also examined closely at this level. The mind map extends to one more level of depth, placing the section about cultural narratives and stereotypes at the literal center of these events, trends and systems. This is where the underlying *worldviews* are exposed, the discourses, mental models and myths that directly or indirectly impact the more superficial levels. The last and deepest level of Voros' layering schema is not reached in this project. This "level of societal, historical and macrohistorical change" (Voros, 2006, p. 46) involves a large scope of such theoretical and philosophical complexity that I consider it too removed to be of practical use for the audience this website is intended for. It should be noted that this depth of analysis comes at the cost of

broadening the scope of the literature review, which consequently means that I cannot claim the comprehensiveness of a more narrowly defined rigorous systematic literature review.

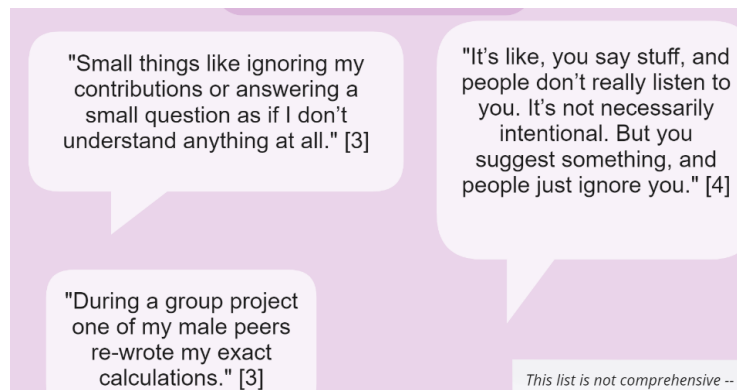


Figure 5: Example of discrete events presented in the analysis.

This application of the layering schema demonstrates how the mind map provides what I considered a sufficient level of interpretive depth to serve as a basis for the later prospective work³. Considering and critically reflecting on this level of depth is important for the outcomes the prospection might produce: Voros (2003) himself points out how his generic foresight process framework can lead to suboptimal results when analysis and interpretation are missing as steps of the process. He calls *reactive* such approaches that proceed from inputs to strategy without engaging in any analysis, interpretation, prospection or output formulation at all. *Shallow* applications engage in analysis but fail to probe beneath the surface to reach a level of depth in interpretation, and disregard explicit prospection before jumping to “rather suspect” outputs (p. 19). On the other hand, *narrow* approaches do include prospective activities but again without thoroughly interpreting the superficial results of an analysis. The mind map covering four levels of interpretive depth is therefore a crucial basis for Visions in part three.

To give a concrete example of such suboptimal action from the context of female participation in physics, consider the case of women guest speakers in K-12 science classes. University outreach

³ Voros doesn't define what would constitute 'sufficient' in depth, so this judgment is entirely my own.

programs and schools often organize presentations by female scientists in class to talk about their work and serve as role models. Despite such efforts, Hazari et al. (2013) criticize this type of role modeling as not having the assumed impact, since it lacks crucial components like relationships and self-realization for female learners. Such outputs seem like the result of a shallow approach, reacting to the concerning trend of low female representation and jumping to a 'solution' that is, at best, not impactful. In contrast, with the step of an in-depth interpretation of the situation, the same group of researchers identified problematic cultural narratives as underlying causes and demonstrated effectiveness for lessons that explicitly challenge such narratives (Potvin et al., 2015; Potvin et al., 2023).

Indeed, this project runs the risk of producing such shallow and/or narrow outputs as well. Since the level of engagement from readers is entirely open and voluntary depending on individual interest, time, expertise and other factors, it is possible that readers will gloss over the mind map and skip ahead to the prospective exercise – or leave the prospective exercise incomplete and jump to outputs. To mitigate this risk, the 'Welcome' section in the mind map suggests options for how to engage with the content, all of which include a recommendation for cultural narratives, the deepest level of analysis covered (s. Figure 6). However, it remains a limitation that this important step is presented merely as an invitation to readers and may fail to unfold its full potential.



Where to start?

Whichever section sparks your interest! There is no defined chronology you need to follow to make sense of the content -- think of it more like Lego set: You can assemble the building blocks in your own individual way.

However, if you're feeling overwhelmed, feel free to follow or start with one of the suggested Journeys:

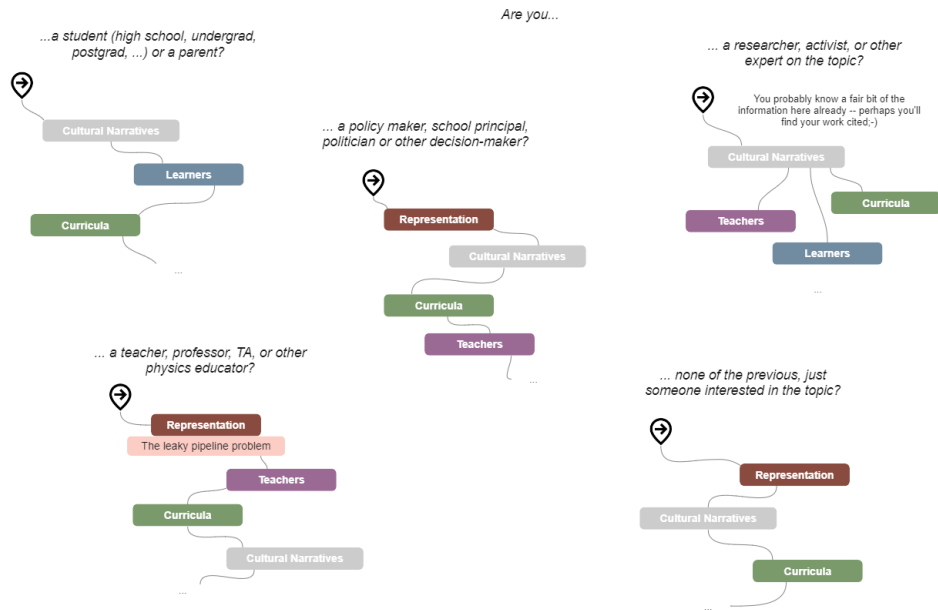


Figure 6: Welcome section in the mind map to guide readers.

3.2. Voices

The podcast episodes are comprised of research interviews that are published in two formats following different methodologies, as outlined in section 2.2. The expert interviews in Episode 2 and 3 have been slightly edited for a pleasant listening experience, for example by adding a musical intro and outro, as well as audio quality enhancements, background noise reduction and fluency improvements. The contents of the conversations, however, have not been significantly edited, resulting in natural conversations that occasionally go beyond strictly relevant topics of the project. For example, Ulrike as an activist for women in physics criticizes wider societal problems of gender equity such as the distribution of unpaid care work as a factor in the issue of female underrepresentation in physics. The decision to include the expert interviews in an almost unaltered format stays true to the participants' answers but comes at the cost of focus and relevance at times. Nonetheless, I consider their full accounts to provide unique personal insights that outweigh the perceived benefits of trimming their answers down to strictly relevant aspects.

Producing fictionalized narratives from the student/teacher interviews for Episode 1 comes with considerably more involvement and deliberate choices from me as a researcher. This artistic defamiliarization – ‘*Verfremdung*’, to use the German literary term – requires some critical reflections guided by criteria from the literature (Richardson & St. Pierre, 2005⁴). Since “mere novelty does not suffice” (p. 964), Richardson holds writings in qualitative research to high standards and elaborates on four evaluative criteria that will serve as a basis for my reflections: Substantive contribution, aesthetic merit, reflexivity and impact (Richardson & St. Pierre, 2005, p. 964).

To evaluate the *substantive contribution*, Richardson poses the question “Does the piece seem ‘true’ – a credible account of cultural, social, individual or communal sense of the real?” (p. 964). As highlighted on the website, the question of truth is indeed a complicated one. By drawing on identifiable sources of data I would argue that the stories fulfill this criterion of credibility. The [corresponding page](#) on the website discerns in detail what those sources are and how they are woven together, alongside some examples.

In contrast, the question of *aesthetic merit* has not been given any attention on the website. For Richardson, the text should invite interpretive responses from readers, it should be artistically shaped, satisfying and complex. The two stories unfold their aesthetic merit through different stylistic and narrative means. For example, the role of the plot itself is vastly different for both stories: Where *Eva* is carried by a clear line of narrative events progressing through rising and falling action – pondering over the problem, solving it, experiencing the adverse reactions by peers and instructors – *Maya* is characterized by a collection of classroom observations that are only loosely framed by the chronology of the exam. Consequently, they invite readers’ interpretative responses in different ways. In *Eva*, the internal focalization, to use Genette’s (1972) narratological framework, allows a narration of her inner thoughts and feelings, inviting readers to identify and empathize with her. This also serves the complexity of characters, avoiding a simplistic narrative that victimizes the female and demonizes the male characters. Instead, ambiguity is embraced through the subjectivity of *Eva*’s feelings: Her male peers are not constructed as bullies – nicknaming her “*Evie*” can be caring or condescending, or both –, and

⁴ The chapter is divided into three parts, the first and third are explicitly attributed to Richardson. The following considerations refer to the contents of her text specifically.

neither does the fellow student who made a joke “that she was too pretty to be smart, and had meant it as a compliment” (p. 8) have any malicious intents – but these comments do have an emotional impact on Eva, explored through the internal focalization. In contrast, *Maya* is told with an external focalization, that is, without insights into the characters’ thoughts or feelings. This purely observational perspective leaves space for the reader to fill in the gaps that are not narrated, for example what Sarah might be thinking as she receives her exam paper and immediately packs it away. One challenge with this perspective lies in capturing appropriate student voice and tone in a narration that is carried forward mostly through direct speech. For example, Max declaring upon receiving his exam “Maaan I’m so smart, I’ll be a doctor of something one day!” (p. 3) is a translation of a student quote recounted by the teacher in a German secondary school. Translating and phrasing such student statements in words that reflect – without the further contextualization of thoughts, feelings or indirect speech – their character in confidence, tone and other social nuances has been challenging and, upon critical reflection, perhaps not always successful.

The *reflexivity* of the text refers to the degree of self-awareness and self-exposure for the reader to make judgments about the point of view. This is why I placed myself as an embodied fictional narrator in both narratives, in *Eva* adding a diegetic layer to the story in the very end, while being present and explicit in *Maya* from the beginning. Despite this intradiegetic presence – existing in the narrative – I chose to remain heterodiegetic, stepping back behind the protagonists (Genette, 1972). Instead of appropriating the interviewee’s stories and telling them as my own, I as a narrator appear only as a side character with an unobtrusive narratorial presence. This follows Richardson’s assertion that “[q]ualitative writers do not have to try to play God, writing as disembodied omniscient narrators, claiming universal and atemporal general knowledge” (Richardson & St. Pierre, 2005, p. 961). Instead, the intra- and heterodiegetic narrator, coupled with internal and external (but not zero!) focalization, emphasizes the self-awareness that is fundamental for a text in qualitative research. While I consider the narrative situation in both stories sufficient to meet the bar of this criterion, an even more provocative text could make use of an unreliable narrator, adding a completely new layer of self-reflexivity and complexity.

Lastly, the *impact* evaluates if the piece affects emotionally or intellectually, generating new questions or moving to action. This is by far the most important criterion by my own standards, and simultaneously the one most difficult to determine for me as the author. Though challenging

to evaluate objectively, I would encourage readers to consider this dimension when assessing the work.

3.3. Visions

As part of the 2012 special issue of 'Futures', focusing on academic works around evaluating futures studies and foresight, Piirainen et al. (2012) introduced a systematic evaluation framework for futures research. This multidimensional framework not only combines various levels of evaluation criteria but also builds on and incorporates Voros' generic foresight process. Both aspects make it a particularly useful basis for critical reflection on the Visions part of the project, specifically, the prospection activity as well as outputs and strategy.

The systemic framework unfolds on three levels of evaluative analysis. First, the level of utility and delivery is concerned with practical questions around objective fulfilment. Second, the technical level deals with the technical execution, suitability of methods and quality of data. Lastly, the third level of ethics draws attention to motivations, underlying worldviews and power relations (Piirainen et al., 2012). For each level, evaluative questions targeting specific stages of the foresight process are proposed⁵. However, "while the principles of form are intended to be generally applicable, the implementation, that is, the principles of function, should be determined case-by-case" (Piirainen et al., 2012, p. 470). With this in mind, I will not apply their complete collection of questions as a checklist, but rather use the three-level framework and relevant questions from each level as prompts for critical reflection.

On the first level, utility and delivery are considered as important factors in creating the impact of the foresight. Feasibility of strategies is among the questions posed for evaluating the foresight on this level. In the activity proposed on the website, the outcomes of the prospection are open and inviting audience contributions, which makes it challenging to support, let alone evaluate, the feasibility of strategies for realizing the output. Nonetheless, the page [Outputs and Strategy](#) includes a collection of resources that can serve as a concrete starting point for possible strategies depending on the context, audience and output. Without assessing the foresight of other

⁵ They adapt Voros' framework slightly and expand it by a sixth phase of monitoring and updating (Piirainen et al., 2012, p. 467). However, this will not be considered for the evaluation of this project's foresight process.

contributions, I can reflect on the examples I have produced during the prospection, output and strategy. “Do these prospections answer the question? Are they challenging and inspiring? Are the actions and strategies based on the prospects?”, are questions proposed by Piirainen et al. (2012, p. 471). My prospection, from the perspective of a physics education researcher, is concerned with future images around societal perceptions and education research. As such, they answer only a small portion of the question *What does physics education in ten years look like?*, leaving other aspects open for the imagination of others, for example the perspectives of the fictional characters Maya and Eva from the short stories. My prospects aim to be challenging and inspiring by imagining probable and plausible, but also possible futures and problematizing some of the shortcomings of the project, for example the binary construction of gender as male or female. Taking up the insights from the previous interpretation step, my prospects develop futures around shifting cultural narratives and reflect on the role of research in this process, acknowledging that despite the vast research of the past decades, little has changed for women in physics in the last twenty years. From this realization, the resulting actions and strategy center around this idea of actively shaping narratives and bridging this gap between research, public discourse and educational practice. The website itself therefore is (or becomes?) the delivery.

However, it should be noted that the questions posed in the systematic evaluation framework are based on criteria for scenario planning (Piirainen et al., 2012, p. 467) and therefore are targeted to assess measurable outcomes like stakeholder satisfaction or relevance of the produced foresight for decision makers. As described in the [theory page](#) on the website, this is too narrow of an approach for the activity proposed here. The impact of the prospection exercise is mainly what Voros described as intangible output, “an expansion of perceptions and perceived options” (2003, p. 15). The purpose is to inspire readers to think about the future and feel empowered to contribute to shaping it, even though they may not perceive themselves to be in a position of power. As such, evaluation criteria like relevance for decision makers are not applicable. The question – which remains unanswerable at this point – is much rather “Is the activity engaging and inspiring enough to have an impact on the readers’ imagination?” (Piirainen et al., 2012, p. 471).

On the second, technical level, questions are concerned with validity. Some of them can be directly applied and answered for the examples of prospection I produced. “Do the prospects cover the future and challenge the status quo?” (Piirainen et al., 2012, p. 471) can be illustrated with some of the ideas presented from Maya’s perspective. Her prospects include scenarios of

technological advancement and reflect on various consequences of such developments for teachers and learners. For example, the 'probable' scenario of AI generated and tailored curricula for individual learners extrapolates current technologies into the near future, while also highlighting the problematic status quo that might be encoded into and reinforced by AI. The framework, again, is applicable only with limitations, as for this technical level the authors "draw extensively from literature on validation of computer-based simulation modeling" (Piirainen et al., 2012, p. 468). The focus of their evaluative questions is therefore on the realms of probable and plausible futures, rendering images of possible futures and wildcard-ideas less valuable. Given the goal of the activity explained above – the intangible output of expanded perceptions – I explicitly disagree with this focus. Therefore, I consider the question "Are the prospectations plausible, consistent with the analysis and coherent?" (Piirainen et al., 2012, p. 471) only with respect to consistency with analysis and coherence. As a matter of fact, the question prompts are devised specifically in derivation from the categories established in the analysis (s. Figure 7). By supporting participants with such concrete, deliberate prompts, I hope to facilitate prospects that are consistent with the analysis. For example, one of the ideas proposed from Maya's perspective is rooted in the insights on how contexts and contents of curricula impact female learners. Nonetheless, the examples on her slide are more than extrapolations but are presented as complex, connected webs with images of probable, plausible and possible futures that resist common pitfalls like simplistic technological fixes to such a social problem as well as fear-based reactions that constrain progressive thinking (Goode & Godhe, 2017). The evaluation of the foresight on a technical level is thus limited to consistency with the preceding analysis and coherence and evades further assessment of validity.

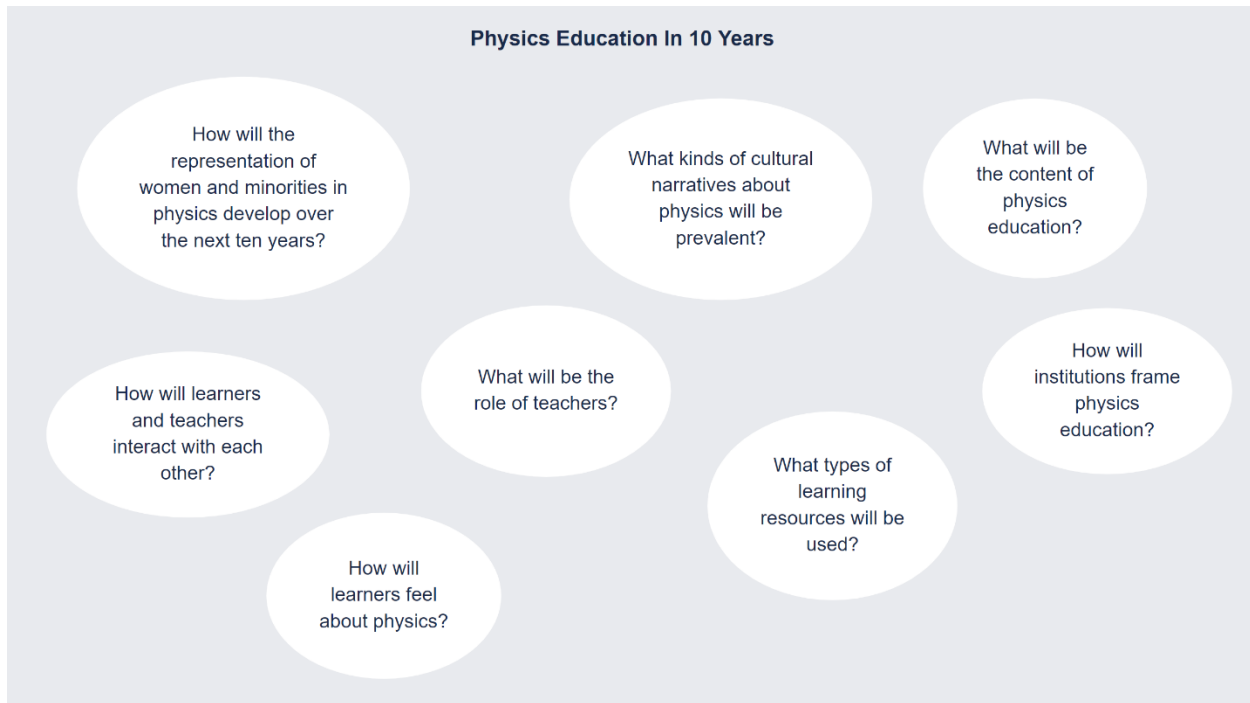


Figure 7: Question prompts to facilitate consistency with analysis. The questions cover the categories of the mind map, namely representation, learners, teachers, curricula and narratives.

Since foresight work in general, and this project is no exception, usually “provide[s] recommendations or opportunities for change, [...] the ethical dimension of the study and its results are of great importance” (Piirainen et al., 2012, p. 470). Such an ethical consideration asks “What worldview determines that which constitutes an improvement?” (Piirainen et al., 2012, p. 471). Voros’ term of ‘preferable futures’ is part of the foresight activity, and as also stated on the website, the answer depends on who is doing the preferring. For example, the probable/plausible future image of female physicists as more and more natural in the public perception of the subject is presented through Eva’s perspective and marked as ‘preferable’ from this worldview. Eva, the fictional student from the short story, makes this judgment from her specific positionality as a female PhD that struggled with gender-biased realities in her physics education. However, someone else may feel threatened and insecure by this prospect of women physicists perceived as normal in public discourse. In its set up, the activity makes no claim about generally preferable futures, but rather invites personal perspectives to be shared. To help place the proposed preferable futures in the context of ‘preferable to who?’, a box in the top left corner asks participants to specify (without compromising anonymity) basic information about their context.

In theory, a strength of the prospection activity in this project lies in its accessibility to diverse audiences, and therefore diverse worldviews. In contrast to expert panels and committees behind closed doors, the website aims to be open and inviting and offer a low-threshold, simple way to participate in the co-creation of futures. I add the caveat “in theory” to this ambitious ethical goal, considering that in practice, the broad invitation may reach only a small, privileged fraction of its possible audience. Who will hear about this offer, and who will have the resources and opportunity to engage? It is not unlikely that the true audience will remain confined to a bubble of western researchers with an interest in the field, and that the project ultimately fails to provide the diversity in perspectives it aims for.

In summary, the prospection, output and strategy in Visions are challenging to evaluate since criteria like decision-making utility, validity or plausibility are not easily assessed. With a focus on intangible, immeasurable impacts on participants’ imagination, the activity may be perceived as insufficiently rooted in basic academic criteria of reliability, validity and objectivity. Such limitations notwithstanding, Visions contributes to the project with a coherent, complex, inspiring and, in theory, largely ethical invitation to co-create futures of physics education.

4. Conclusion

This report has complemented the website with an extensive discussion of the various disciplinary perspectives that are merged, adapted and applied to produce new insights throughout the three parts of the project. It is through this interdisciplinarity that the website explores physics education in depth and breadth. It is, however, imperative to also acknowledge the limitations of the project that the critical reflections have shed light on. Particular challenges lie with potential impacts, which remain largely intangible, difficult to quantify or validate, and possibly unexpected. The core contribution of the project consists of inviting and empowering a diverse public audience to engage with the topic as applicable to their context, without being able to provide an easy fix for a complex societal problem. As a consequence, the nature and extent of the impact the website unfolds after its publication remains a topic for future research. Whether it succeeded in its ambitious goal to serve as a bridge between research, public discourse and educational practice, and what other consequences it entailed, awaits further investigation. Moving forward, the project’s success will ultimately be gauged by its ability to inspire and engage a broad audience. As researchers, educators, and the public continue to interact with and build upon the website’s

content, the hope is that it will catalyze meaningful change and drive further scholarly and practical advancements in the field.

Bibliography

- About APA Style*. (2023, March). <https://apastyle.apa.org>. <https://apastyle.apa.org/about-apa-style>
- Anthoni, E., Leemput, M. V., Schoffelen, J., & Hannes, K. (2020). Futures Studies. In P. Atkinson, S. Delamont, A. Cernat, J. W. Sakshaug, & R. Williams (Eds.), *Sage Research Methods Foundations*. SAGE Publications Ltd. <https://doi.org/10.4135/9781526421036949493>
- Bell, W. (2003). *Foundations of futures studies: History, purposes, and knowledge*. (Vol. 1). Transaction Publishers.
- Booth, A., Sutton, A., & Papaioannou, D. (2016). *Systematic Approaches to a Successful Literature Review* (Second edition). SAGE Publications Ltd.
- Brinkmann, S. (2020). Unstructured and Semistructured Interviewing. In P. Leavy (Ed.), *The Oxford Handbook of Qualitative Research* (p. 0). Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780190847388.013.22>
- Clough, P. (2002). *Narratives And Fictions In Educational Research* (1st edition). Open University Press.
- Concise guide to APA Style* (7th ed., pp. xxi, 326). (2020). American Psychological Association. <https://doi.org/10.1037/0000173-000>
- Directorate-General for Research and Innovation (European Commission), & Soeiro de Carvalho, P. (2023). *R&I foresight policy and practice: Futures studies, disruptions and the twin transition: fourth thematic report*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2777/952097>
- Dirse, Z. (2013). Gender in cinematography: Female gaze (eye) behind the camera. *Journal of Research in Gender Studies*, 3(1), 15–30.
- Female gaze*. (n.d.). Oxford Reference. Retrieved June 23, 2024, from <https://www.oxfordreference.com/display/10.1093/oi/authority.20110803095814800>
- Fiel'ardh, K. (2024). Futures Thinking in Middle School Science Textbooks: A Perspective from Japan. *Nordic Journal of Comparative and International Education (NJCIE)*, 8(2). <https://doi.org/10.7577/njcie.5647>
- French, L. (2021). The 'Female Gaze.' In L. French (Ed.), *The Female Gaze in Documentary Film: An International Perspective* (pp. 53–70). Springer International Publishing. https://doi.org/10.1007/978-3-030-68094-7_3
- Genette, G. (1972). *Figures III*. Éditions du Seuil.
- Gillham, B. (2005). *Research Interviewing: The Range of Techniques*. McGraw-Hill Education. <http://ebookcentral.proquest.com/lib/ed/detail.action?docID=287863>

- Goode, L., & Godhe, M. (2017). Beyond Capitalist Realism – Why We Need Critical Future Studies. *Culture Unbound*, 9(1), Article 1. <https://doi.org/10.3384/cu.2000.1525.1790615>
- Gough, D., Thomas, J., & Oliver, S. (2017). *An Introduction to Systematic Reviews*. 1–352.
- Hazari, Z., Potvin, G., Lock, R. M., Lung, F., Sonnert, G., & Sadler, P. M. (2013). Factors that affect the physical science career interest of female students: Testing five common hypotheses. *Physical Review Special Topics - Physics Education Research*, 9(2), 020115. <https://doi.org/10.1103/PhysRevSTPER.9.020115>
- Inayatullah, S. (1998). Causal layered analysis: Poststructuralism as method. *Futures*, 30(8), 815–829. [https://doi.org/10.1016/S0016-3287\(98\)00086-X](https://doi.org/10.1016/S0016-3287(98)00086-X)
- Kristóf, T. (2023). The Story of Futures Studies: An Interdisciplinary Field Rooted in Social Sciences. *Social Sciences* 12(3), 192. https://www.academia.edu/99083662/The_Story_of_Futures_Studies_An_Interdisciplinary_Field_Rooted_in_Social_Sciences
- Lo, C. K. (2020). Systematic Reviews on Flipped Learning in Various Education Contexts. In O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond, & K. Buntins (Eds.), *Systematic Reviews in Educational Research: Methodology, Perspectives and Application* (pp. 129–143). Springer Fachmedien. https://doi.org/10.1007/978-3-658-27602-7_8
- Marien, M. (2002). Futures studies in the 21st Century: A reality-based view. *Futures*, 34, 261–281. [https://doi.org/10.1016/S0016-3287\(01\)00043-X](https://doi.org/10.1016/S0016-3287(01)00043-X)
- Menken, S., & Keestra, M. (2016). *An Introduction to Interdisciplinary Research: Theory and Practice*. Amsterdam University Press. <http://ebookcentral.proquest.com/lib/ed/detail.action?docID=4460743>
- Mulvey, L. (1975). Visual Pleasure and Narrative Cinema. *Screen*, 16(3), 6–18. <https://doi.org/10.1093/screen/16.3.6>
- National Academy of Sciences. (2005). *Facilitating Interdisciplinary Research*. The National Academies Press. <https://doi.org/10.17226/11153>
- Nemeth, B., Dew, N., & Augier, M. (2018). Understanding some pitfalls in the strategic foresight processes: The case of the Hungarian Ministry of Defense. *Futures*, 101, 92–102. <https://doi.org/10.1016/j.futures.2018.06.014>
- Newell, W., & Klein, J. T. (1997). Advancing Interdisciplinary Studies. In J. G. Gaff & J. Ratcliff (Eds.), *Handbook of the Undergraduate Curriculum: Comprehensive Guide to Purposes, Structures, Practices, and Change* (pp. 393–415). Joseey-Bass.
- Newman, M., & Gough, D. (2020). Systematic Reviews in Educational Research: Methodology, Perspectives and Application. In O. Zawacki-Richter, M. Kerres, S. Bedenlier, M. Bond, & K. Buntins (Eds.), *Systematic Reviews in Educational Research: Methodology,*

Perspectives and Application (pp. 3–22). Springer Fachmedien.
https://doi.org/10.1007/978-3-658-27602-7_1

- Piirainen, K. A., Gonzalez, R. A., & Bragge, J. (2012). A systemic evaluation framework for futures research. *Futures*, 44(5), 464–474. <https://doi.org/10.1016/j.futures.2012.03.008>
- Popay, J., Roberts, H., Sowden, A., Petticrew, M., Arai, L., Rodgers, M., Britten, N., Roen, K., & Duffy, S. (2006, April 1). *Guidance on the Conduct of Narrative Synthesis in Systematic Reviews*. ESRC Methods Programme. <https://www.lancaster.ac.uk/media/lancaster-university/content-assets/documents/fhm/dhr/chir/NSsynthesisguidanceVersion1-April2006.pdf>
- Potvin, G., Hazari, Z., Khatri, R., Cheng, H., Head, T. B., Lock, R. M., Kornahrens, A. F., Woodle, K. S., Vieyra, R. E., Cunningham, B. A., Kramer, L., & Hodapp, T. (2023). Examining the effect of counternarratives about physics on women’s physics career intentions. *Physical Review Physics Education Research*, 19(1), 1–14. <https://doi.org/10.1103/PhysRevPhysEducRes.19.010126>
- Potvin, G., Hazari, Z., & Lock, R. M. (2015). Exposure to Underrepresentation Discussion: The Impacts on Women’s Attitudes and Identities. *2014 Physics Education Research Conference Proceedings*, 211–214. <https://doi.org/10.1119/perc.2014.pr.049>
- Richardson, L. (1994). Writing: A method of inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (1st ed., pp. 516–529). Sage Publications.
- Richardson, L., & St. Pierre, E. A. (2005). Writing: A method of inquiry. In N. K. Denzin & Y. S. Lincoln (Eds.), *The Sage handbook of qualitative research* (3rd ed., pp. 959–978). Sage Publications.
- Riddell, G. A., van Delden, H., Maier, H. R., & Zecchin, A. C. (2020). Tomorrow’s disasters – Embedding foresight principles into disaster risk assessment and treatment. *International Journal of Disaster Risk Reduction*, 45, 101437. <https://doi.org/10.1016/j.ijdrr.2019.101437>
- Rowley, J., & Slack, F. (2004). Conducting a literature review. *Management Research News*, 27(6), 31–39. <https://doi.org/10.1108/01409170410784185>
- Slaughter, R. A. (2002). Beyond the mundane: Reconciling breadth and depth in futures enquiry. *Futures*, 34(6), 493–507. [https://doi.org/10.1016/S0016-3287\(01\)00076-3](https://doi.org/10.1016/S0016-3287(01)00076-3)
- Son, H. (2015). The history of Western futures studies: An exploration of the intellectual traditions and three-phase periodization. *Futures*, 66, 120–137. <https://doi.org/10.1016/j.futures.2014.12.013>
- St. Pierre, E. A. (1997). Methodology in the Fold and the Irruption of Transgressive Data. *International Journal of Qualitative Studies in Education*, 10(2), 175–189.

Voros, J. (2003). A generic foresight process framework. *Foresight*, 5(3), 10–21.
<https://doi.org/10.1108/14636680310698379>

Voros, J. (2006). Introducing a classification framework for prospective methods. *Foresight*, 8(2), 43–56. <https://doi.org/10.1108/14636680610656174>