Journal of Responsible Technology 25 (2026) 100152

Contents lists available at ScienceDirect

Journal of Responsible Technology

journal homepage: www.sciencedirect.com/journal/journal-of-responsible-technology

Review Article

An exploratory thematic review of the emerging field of RRI education

Ferdinand Griesdoorn *, Maarten Kroesen, Pieter Vermaas

Delft University of Technology, Delft, Zuid Holland, the Netherlands

ARTICLE INFO

Keywords:
Responsible research and innovation (RRI)
Thematic review
Competencies
Successful practices
Obstacles

ABSTRACT

This exploratory thematic review examines the emerging landscape of Responsible Research and Innovation (RRI) education. It reviews 17 peer-reviewed studies published over the past two decades, using the PRISMA methodology. These studies were categorized into four themes to identify recurring successes and obstacles. The review highlights several successful practices, including the contextualization of RRI, promotion of reflexivity, participatory methods, interdisciplinary collaboration, and instances of institutional integration. Simultaneously, it uncovers persistent obstacles such as conceptual ambiguity, institutional resistance, scalability limitations, and the difficulty of translating abstract RRI principles into measurable competencies. The relations between some of these obstacles suggests a vicious reinforcing cycle that hinders progress in RRI education. We argue that resolving conceptual and definitional ambiguities could foster a more coherent and sustainable RRI education.

1. Introduction

In an era of fast-paced technological advancement, science and innovation are deeply intertwined with ethical, social, and environmental concerns. As global challenges reshape the research and innovation landscape, ensuring that scientific progress and innovation align with societal values and expectations has become a priority in Europe (EU, 2022). Although there is no agreed-upon definition for Responsible Research and Innovation, nor a clear methodology to educate on RRI, the European Union has positioned Responsible Research and Innovation (RRI) as a key approach to fostering an inclusive society that steers research and innovation through lessons from the past, actions in the present, and responsibility for the future (Von Schomberg, 2013; Owen & Pansera, 2019).

If one agrees that research and innovation should align with societal values and expectations, and that RRI is an important approach in achieving this alignment, then teaching RRI is not just an academic subject of importance but a societal necessity: it equips the next generation of students, scientists, innovators, and policymakers to navigate complex ethical and societal dimensions of research and innovation.

Despite the increasing recognition of RRI in European policy documents and scholarly discussions, its integration into education is underexplored and unevenly applied across disciplines. Many efforts have been made to teach RRI, however the current academic insight lacks a thematic exploration of these efforts, their challenges, opportunities, and solutions. As education on research and innovation often focuses on technical skills and knowledge, this focus leaves a gap with regards to the broader societal impacts and ethical responsibilities. This gap highlights the need for studies on how RRI should be educated effectively and embedded into education and training programs at all levels, for all disciplines. Understanding how to foster RRI education is crucial for the education of students, scientists, innovators, and policymakers.

By reviewing education in RRI, we provide insights that facilitate the development of educational RRI approaches. We hold that, exploring how to advance education on RRI is not only an essential step toward a research and innovation ecosystem that is innovative and forward-looking, but also one that is deeply reflective, inclusive, and accountable. Our exploration shows successes with RRI education and uncovers several interrelated challenges with RRI and its education. In the conclusions section we give some suggestions how these challenges could be resolved.

To navigate the landscape of RRI education, this review sets out to document existing practices, and to uncover deeper patterns. At its core, this review is guided by the following central question:

How is RRI taught, implemented, and evaluated in educational settings, and what thematic patterns of success and challenge emerge across recent scholarly studies?

To answer the central question, we examine a series of interrelated inquiries that structured our analysis:

* Corresponding author.
E-mail address: ferdinand.griesdoorn@minbuza.nl (F. Griesdoorn).

https://doi.org/10.1016/j.jrt.2026.100152

Available online 13 January 2026

- What educational strategies and methodologies have been used to teach RRI, and how can we categorize them?
- Which elements of RRI education are consistently reported as successful across studies?
- What are the most persistent obstacles or limitations faced in implementing RRI education?
- How do these successes and challenges interact, and what systemic patterns can be identified?
- What conceptual or philosophical gaps in RRI education hinder its coherence and long-term viability?

After the introduction of this paper, the method section describes the three-step process we used for reviewing literature on RRI education, including search strategies, and the selection criteria we used to identify studies. The results and analysis section categorizes the studies into four themes. Per theme the findings are discussed first and then the results and analysis section concludes with a synthesis. The conclusions section summarizes key findings, emphasizes the need for a more coherent, practical, and science based educational approach, and proposes future directions. The paper concludes with a discussion of its limitations, highlighting the dependency on the completeness of the reviewed studies and the opportunities for future research.

2. Methodology

To discover themes, challenges, opportunities, and solutions in RRI education we used PRISMA, the "Preferred Reporting Items for Systematic Reviews and Meta-Analyses" method.

Search entries definition and literature discovery

A comprehensive search was performed using the following defined entries:

- Responsible Research and Innovation
- RRI Education
- RRI Approaches
- RRI Training
- Understanding RRI
- Implementing RRI
- Science Education
- Science Communication
- Teaching Science
- Successful Education Strategies

These search terms were applied across multiple sources, including:

- Google Scholar
- Web of Science
- Research Gate

Academic references in papers have been used as well, and we explored additional literature through snowballing. This step resulted in the identification of 52 publications.

Screening & eligibility

To ensure relevance and quality the identified publications were screened using four criteria:

- Publications should report on the education of RRI.
- Publications should be peer-reviewed studies.
- Publications should be of the last two decades.
- Publications needed sufficient detail to support the review, such as information on the studied subject, targeted audience, goals, and results.

With these criteria, 17 publications out of the initial 52 publications, were included in the review Fig. 1.

3. Results and analysis

We use an inductive thematic synthesis within our exploratory thematic review, clustering studies by their primary epistemological function (describing phenomena, assessing impact, testing methodological innovations, or applying and extending frameworks). An inductive approach is an approach where the recognition of categories and the assignment of studies emerge from the data itself (Gough, Oliver & Thomas, 2012).

Given the diversity of studies reviewed, many could have reasonably been placed in more than one category. For example, a methodological study might also provide evaluative insights, or a descriptive account may introduce an applied framework. To maintain clarity and comparability across the dataset, each study was assigned to the category that best reflected its primary research emphasis. This approach follows common practices in reviews, where overlapping contributions are acknowledged but classification is guided by the dominant objective of the study.

Both European and UK frameworks were included in the search strategy. However, the resulting corpus was dominated by studies situated within the European Union. The UK's Responsible Innovation (RI) framework, although conceptually aligned, has produced fewer published studies explicitly addressing educational practice. As a result, UK-based studies were underrepresented in the final sample, reflecting the current distribution of published research rather than a deliberate exclusion.

This review used the following categorization:

- Descriptive Studies, these studies document or map practices, principles, or contexts of RRI in educational settings. Their aim is to provide foundational understanding rather than test interventions or frameworks. Providing foundational understanding, these studies focus on describing RRI practices and concepts in educational contexts.
- Review Studies, these studies assess the outcomes, effectiveness, or impact of RRI-related education or training. They often combine qualitative and quantitative approaches to capture multidimensional effects. informing on the effectiveness, these studies assessed the effectiveness of RRI-related interventions, training, or methodologies.
- Intervention Studies, these studies propose, develop, or test new methods, approaches, or interventions for embedding RRI principles in education. Their focus lies in innovation and experimentation. Suggesting novel approaches, these studies introduced or tested new methods for implementing RRI in education.
- Conceptual Studies, these studies extend RRI concepts into practical applications, institutional frameworks, or cross-disciplinary contexts. They aim to bridge theory and practice, often proposing tailored frameworks or strategies for implementation. Bridging theory with real-world applications, these studies explored the practical application of RRI and proposed expansions to existing frameworks.

This categorization captures the breadth of approaches while ensuring analytical clarity. It also makes the review more comparable to similar reviews in education, health sciences, and social sciences Fig. 2.

We use this categorization for the following reasons:

- To enhance the comprehensiveness of this review as categories collectively encompass the breadth of approaches and perspectives in RRI education, ensuring a holistic understanding.
- To gain practical utility, which enables targeted analysis of what is successful, what are potential obstacles, and potential solutions for specific aspects of RRI education.

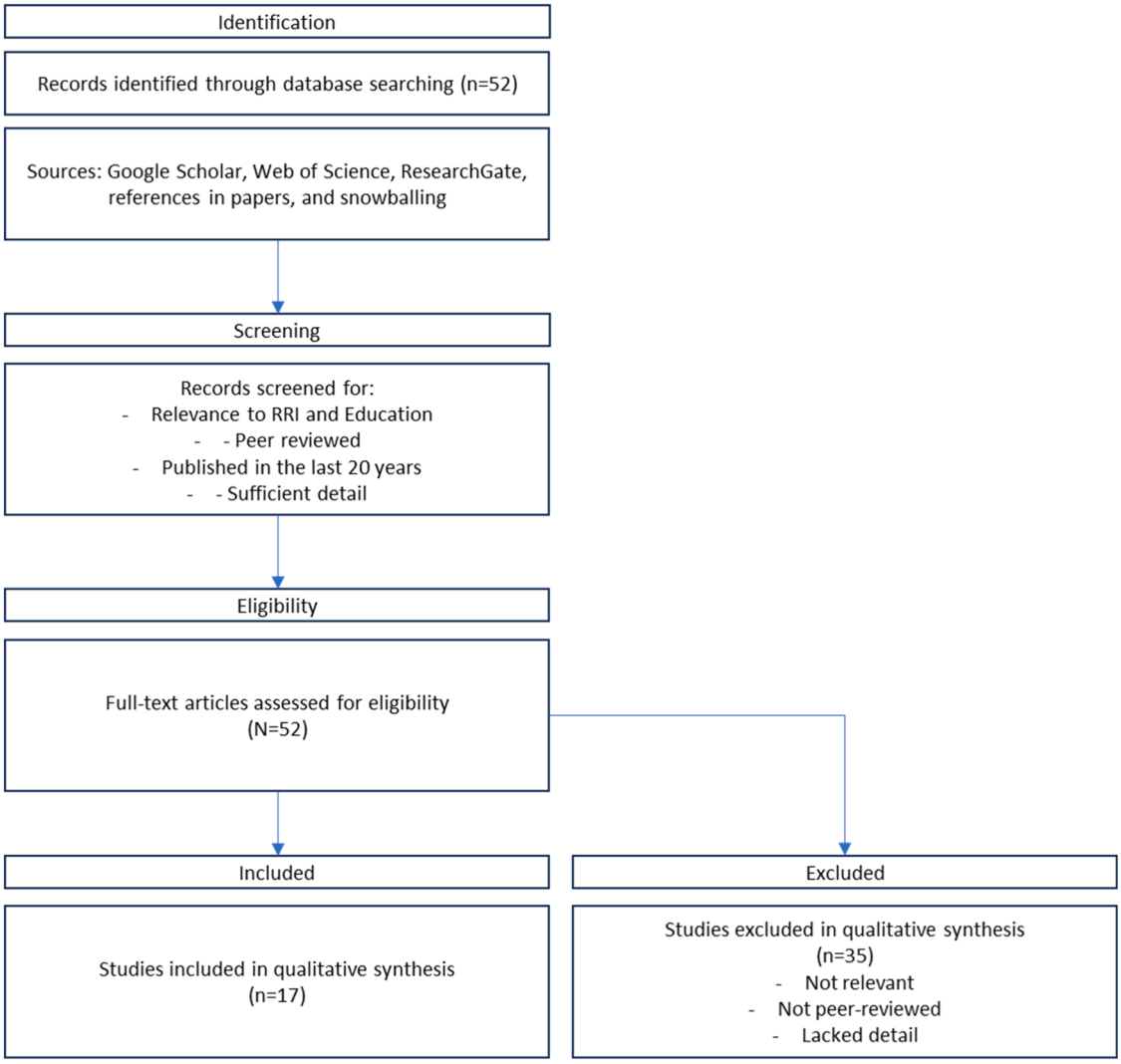

Fig. 1. PRISMA based visual chart showing the process of literature selection.

- To compare studies within and across categories to identify patterns, gaps, and trends.

3.1. *Descriptive studies*

In the review we identified three studies that primarily fall into the Descriptive Studies category. These studies focus on documenting and analysing existing practices, principles, and contexts of RRI, providing foundational understanding rather than testing or evaluating interventions.

The success of these studies, particularly in fostering critical thinking and deepening awareness of the societal implications of science and technology, highlights key facilitators for RRI education. These facilitators include contextualizing RRI education, bridging the gap between theory and practice, and encouraging active participation in real-world socio-scientific issues. However, significant obstacles were identified, such as limited institutional recognition, resource constraints, and the inability to assess or evaluate long-term impacts. Additionally, the studies emphasized the need for a more systematic and coherent approach to RRI education.

For example, on the island of Malta, Levikov, Quacinella and Ducac (2020) studied endeavours aimed to embed a culture of RRI best practices within a local university ecosystem. The research project used a two-tier preliminary stage of formative research. The first tier was to understand RRI in the local context by doing interviews and appealing to academics, staff, and students at the University of Malta (UM) for the participation in questionnaires. The interviewees were chosen from different faculties and centres to ensure that opinions from a variety of disciplines and sectors were addressed. The second tier was to create a culture of RRI best practice and establish key figures that can support this endeavour. By engaging academics, students, and staff through interviews and questionnaires, the project created actionable frameworks like the Committee for Engaged Research, demonstrating how localized initiatives can foster broader cultural shifts. However, while fostering a culture of RRI best practices demonstrated the power of localized initiatives, challenges such as limited institutional recognition and resource constraints highlighted the systemic barriers to scaling such efforts.

While the study in Malta highlighted localized efforts to foster RRI within a university ecosystem, Elster, Barendziaks, and Birkholz (2020) demonstrated how science education can embed RRI through participatory and reflective practices. Their research at the University of Bremen's Faculty of Biology and Chemistry, under the STARBIOS2 project, aimed to integrate RRI structurally into academia. The Citizen-SIP Model, a participatory assessment tool used in the study, effectively enhanced students' understanding of RRI by engaging them in real-world socio-scientific issues, fostering critical thinking, and raising

Papers included in the review

Authors	Category	Pub.
Doris Elster, Tanja Barendziak, Julia Birkholz	Descriptive	2020
Nika Levikov, Daniela Quacinella, Edward Duca	Descriptive	2020
Mirjam Schuijff, Anne M. Dijkstra	Descriptive	2019
Stahl et al.	Review	2023
Vestergaard Bidstrup et al.	Review	2024
Emanuele Bardone, Mirjam Burget, Margus Pedaste	Review	2023
Maria Heras, Isabel Ruiz-Mallén	Review	2017
Stahl et al.	Intervention	2021
Lukovics et al.	Intervention	2019
Hesjedal et al.	Intervention	2020
Malagrida et al.	Intervention	2022
García-Melón et al.	Conceptual	2022
Sandu et al.	Conceptual	2021
Marschalek et al.	Conceptual	2017
Portillo et al.	Conceptual	2022
Mejlgaard et al.	Conceptual	2018
Gry Oftedal	Conceptual	2014

Fig. 2. Overview of literature included and their categorization .

awareness of science's societal impact. Successful integration into university education included modules on risk literacy in nanotechnology and ethical dilemmas like biodiversity loss and climate change. While evaluations confirmed meaningful learning outcomes, such as inclusiveness, anticipation, and responsiveness, scaling the model to larger, diverse student groups posed logistical challenges. Additionally, its resource-intensive implementation and lack of long-term impact assessment of the training on students' professional practices may limit widespread adoption.

The challenge in establishing long-term metrics for impact and engagement, seems to be reflected in the limited long-term evaluations in a study done by Schuijff and Dijkstra (2019). This study offers a foundation by systematically analysing 79 RRI practices across various contexts, anchored in the core dimensions of stakeholder engagement, anticipation, reflexivity, and responsiveness. It highlights the importance of aligning research with societal needs and values while identifying success factors such as collaboration, adaptability, and institutional support. By providing practical recommendations such as enhancing training, fostering co-creation, and developing governance frameworks, the study bridges theoretical ideals and practical applications. However, next to these successful applications this review reveals significant drawbacks, such as inconsistent methodologies, limited evaluation of long-term impacts, and insufficient integration of RRI principles into institutional structures. The study emphasizes the need for a more systematic and coherent approach to embedding RRI into research practices, offering a foundation for future advancements.

While these studies reveal foundational challenges and contextual insights into RRI education, the Review Studies build on this by critically assessing the effectiveness and measurable impact of these educational efforts.

3.2. Review studies

Four studies were classified as Review Studies. These works assess the effectiveness, outcomes, and broader impact of RRI education or implementation, highlighting both promising practices and persistent challenges.

In these studies, the participatory methods show to enhance reflexivity and societal alignment by involving students and citizens in shaping outcomes. Furthermore, structured frameworks that link RRI to educational philosophies and assess multidimensional learning have shown promise in improving RRI education, with holistic assessments benefiting from the combination of qualitative and quantitative methods. However, evaluating reflexivity, ethical awareness, and societal impact remains challenging. Furthermore, these studies report underrepresentation of ethical and gender issues, resource limitations hindering participatory and exploratory methods, and a tendency to prioritize short-term outcomes at the expense of long-term sustainability, reducing the depth of RRI alignment. While the Descriptive Studies highlight the need for a more systematic and coherent approach, the Review Studies point to the need for a more comprehensive understanding of RRI. Finally, these studies show that traditional metrics fall short in addressing challenges like assessing long-term societal impact, engagement, and ethical awareness.

An argument that could benefit institutional recognition and potentially bridge the gap between theory and practice was made by Bardone, Burget and Pedaste (2023). In relation to science education, their study shows that the connections between RRI and existing philosophies (Nature of Science, Inquiry-Based Learning, Socio-Scientific Issues, Citizenship Education) have been made, yet there is a need for a more comprehensive and coherent understanding of RRI. This is consistent with a broader pattern across the reviewed literature: multiple studies report conceptual ambiguity (for example, whether RRI is empirical, normative, or procedural) and difficulties translating this into teachable and assessable learning outcomes (Stahl et al., 2023; Hesjedal et al., 2020), including challenges in distinguishing scientific reasoning from ethical reflection (Malagrida et al., 2022). The RRI Map Study suggests overcoming fragmentation and gaining a coherent understanding of RRI in science education, by using an RRI mapping and link it to existing educational philosophies. In this way the study contributes to making RRI relatable in education. The mapping has three ways corresponding to time and engaging RRI. This gains three practical meanings that the term responsibility “as care” may acquire, which are: (1) responsibility for sense-making (past), (2) responsibility for action-taking (present), and (3) responsibility for exploring (future). Where the connections between RRI and philosophies in science

education makes RRI more relatable, the same would be so if RRI would connect to other scientific philosophies in institutes, as this would enhance the recognition of RRI in those institutes.

Bridging the gap between theory and practice and enhancing the recognition of RRI in institutes should advance RRI education, and this may increase the need for the assessment of education on RRI. Stahl et al. (2023) focus on training and evaluating RI and RRI, acknowledging their conceptual distinctions, while remaining under development. While training on RRI principles is widely supported, there is less consensus on its content, delivery, and assessment. The study raises key questions, such as who holds responsibility in RRI, whether it is cognitive or empirical, and how training can address internal tensions. It highlights essential requisites for effective training and assessment, including conceptual clarity, a balance of qualitative and quantitative methods, and transparent evaluation criteria with constructive feedback. However, it argues that no universal assessment method suits all institutions and disciplines. The study also underscores the challenges of measuring complex competencies like reflexivity and ethical awareness, as well as assessing societal impact and long-term ethical engagement using traditional metrics.

Ethical awareness can be assessed using an analytical framework developed by Heras and Ruiz-Mallén (2017), which focused on young participants in science education while aligning with pedagogical practices. The framework includes indicators for cognitive aspects (knowledge acquisition, conceptual understanding), experiential aspects (emotions, attitudes, engagement), transversal competences (reflective thinking, collaboration, communication), and RRI dimensions (inclusiveness, ethical awareness, gender equality). A review of 72 studies revealed a reliance on quantitative methods (for example, closed-ended questionnaires), limiting insights into experiential and emotional learning. While mixed method approaches better captured RRI's multidimensional nature, qualitative methods were underutilized. The study emphasized linking assessment indicators to RRI process requirements and advocated for holistic assessments integrating cognitive, emotional, and ethical learning dimensions. Challenges included time and resource constraints, curricular pressures favouring standardized assessments, and difficulties in measuring abstract concepts like ethics and reflexivity. Cognitive learning was the most frequently assessed, while ethical and gender issues, as well as emotional engagement, were often overlooked. The study stressed the need for participatory methods, involving students and educators in developing assessment tools, and argued that RRI science education assessments must integrate metacognitive skills, emotional aspects, and procedural dimensions for meaningful impact evaluation.

Tweed (2009) explains that the metacognitive function in the context of education refers to a learner's ability to monitor, control, and reflect on their own thinking and learning processes. It involves awareness and regulation of one's cognitive activities and is crucial for effective learning and problem-solving. In science education the metacognitive function is regarded as a core learning objective. Mastering this objective enables students developing a capacity to identify what they understand and misunderstand about scientific concepts and what they should do to learn what they misunderstand. These seem to be good arguments for using and enhancing the metacognitive function, to self-teach RRI and develop RRI specific competences. This takes the learning effect from the short term into a longer-term development.

While in the Descriptive Studies, the importance of contextualizing RRI education was underlined as important for the endeavour to bridge the gap between theory and practice, as a part of the Review Studies, we find a study that showed opportunities to understand the context of RRI and RRI education better. Vestergaard Bidstrup et al. (2024) did this by integrating citizen feedback into policymaking. Furthermore, it enhanced the opportunities to gain public trust and accountability. The specific objectives were to determine whether citizens' input could contribute to aligning territorial policies with RRI principles and making them more responsible, inclusive, and reflective of societal needs.

The study involved two territories:

- Western Macedonia, Greece: Facing challenges like economic vulnerability and the transition from lignite-based energy to clean energy.
- Sofia Municipality, Bulgaria: Addressing urban development issues such as air pollution, climate change, and sustainable mobility.

The panels included 23 citizens in Western Macedonia and 27 in Sofia Municipality, selected to represent diverse demographic groups and ensure broad societal representation.

The citizens provided valuable insights into improving territorial policies, aligning them with societal needs and values. They introduced additional elements of RRI, such as ethics, responsiveness, and inclusiveness, which were often missing in initial policy proposals. In Western Macedonia, citizens emphasized the importance of immediate societal benefits, like job creation and community support, over broader infrastructural projects. In Sofia, citizens prioritized sustainability and digital transformation, focusing on practical actions that could address urban challenges. The success of this study showed improved policy recommendations reflected citizens' feedback, particularly in aligning actions with RRI principles such as inclusivity and anticipation. However, the study has some challenges as well: The panels were limited to a one-day event, potentially constraining the depth of deliberation and the scope of feedback, citizens prioritized short-term financial and social benefits, sometimes neglecting longer-term or broader sustainability goals, and some RRI principles, like gender equality, received minimal attention, reflecting a narrow focus on economic and practical outcomes.

The gaps found in the Review Studies, coupled with the complexity of assessing emotional aspects, procedural dimensions, and nuanced competencies, create substantial challenges that maybe addressed by innovative methodological advancements.

3.3. *Intervention studies*

In this review we grouped four studies as Intervention Studies. These contributions propose and test innovative approaches for embedding RRI principles into research, education, and policymaking.

These studies emphasize the importance of fostering reflection on assumptions, values, and societal dimensions to promote deeper understanding and transformative learning. They also highlight how interdisciplinary interactions enhance both the understanding and application of RRI in research practices. Several of these studies demonstrated that linking science education with societal challenges through structured frameworks helps map RRI competencies into educational settings, aligning learning with real-world applications. Participatory methods were shown to enhance collaboration and inclusivity, effectively integrating education with societal values. Additionally, embedding RRI into governance and compliance mechanisms through institutional frameworks offers the potential for long-term sustainability beyond project funding. However, persistent debates over RRI's definition and scope hinder consistent integration, while resource-intensive methodologies face scalability issues in broader educational settings. The prioritization of immediate outcomes over long-term impact remains a significant challenge, affecting both Review Studies as well as Intervention Studies. Many educational resources fail to comprehensively address RRI competencies, particularly in areas like reflexivity and ethical analysis. The expertise of facilitators and the resource demands for implementing RRI principles continue to influence RRI education. Challenges in evaluating the long-term societal and cultural impact of RRI education, first identified in Descriptive Studies, persist in Intervention Studies. Additionally, participants often struggle to articulate and apply RRI concepts, reinforcing the complexity of translating RRI into competence and supporting the argument that no straightforward approach exists.

For example, a methodological advancement addressing the

complexity of nuanced competencies maps STEAM education competencies into a framework of eight RRI-linked categories. Malagrida et al. (2022) developed this framework by engaging teachers in focus groups to assess RRI competencies using existing educational tools like virtual experiments, card games, and community-based participatory research (CBPR). Teachers emphasized RRI's role in bridging science education and societal challenges, aligning learning with real-world applications. However, challenges emerged: Some resources failed to comprehensively address RRI competencies, such as distinguishing scientific from ethical issues or anticipating societal impacts, students often struggled to critically analyse scientific evidence, and not all tools were equally effective (CBPR fostered reflexivity better than virtual experiments but faced logistical barriers), and teachers needed further training to facilitate discussions and engage stakeholders.

Before this study, Lukovics et al. (2019) emphasized the importance of early RRI integration into Research Development Innovation (RDI) to strengthen its assimilation. Their study adapted the Socio-Technical Integration Research (STIR) method, encouraging interdisciplinary dialogue and reflection on societal impacts during research. Pre- and post-study questionnaires showed increased awareness, reflexivity, and broader intellectual engagement. Challenges included initial student resistance to integrating non-scientific factors, scalability issues due to STIR's intensive nature, a lack of long-term impact assessments, and facilitator dependency, which affected replicability. Despite these challenges, both studies highlight the need for flexible, interdisciplinary, and participatory methods in RRI education.

To address the challenge of prioritizing short-term outcomes over long-term RRI sustainability, Stahl et al. (2021) explored transitioning from project based RRI to a more embedded model, "Responsibility by Design" (RbD). This approach stems from eight years of RRI implementation in the Human Brain Project (HBP) through its Ethics and Society subproject, which integrated the RRI AREA framework (Stilgoe, Owen & Macnaghten, 2013). The study assessed RRI's long-term impact within HBP and its potential for future European research infrastructures, particularly EBRAINS. Successes included public engagement shaping project decisions, development of ethical guidelines, and the embedding of RRI principles into interdisciplinary collaboration, governance, and compliance. The study demonstrated how RRI can extend beyond research projects via RbD. Challenges included ongoing debates over RRI's definition and scope, disciplinary differences complicating alignment, legal and compliance demands overshadowing broader RRI goals, difficulties ensuring RRI continuation post-funding, and conceptual and practical challenges in assessing RRI's long-term cultural impact.

This review suggests that the ongoing debate over RRI's definition and scope, emphasizes the need for a more comprehensive and coherent understanding. Hesjedal et al. (2020) addressed this by advocating a shift from procedural compliance to transformative understanding, introducing double-loop learning, which challenges underlying assumptions rather than reinforcing existing procedures. This reflective approach aimed to make RRI education more adaptive and self-critical. Their study analysed a Ph.D. course for early-career researchers in biotechnology fields, examining how it encouraged participants to reflect on their practices, values, and theoretical assumptions. The course successfully created "breakthroughs", fostering deeper awareness of science's societal dimensions and the role of social actors in research. Some participants continued reflecting on these concepts long after the course. However, while participants felt their perspectives had shifted, many struggled to articulate their learning outcomes or apply them in daily work. Some found it difficult to grasp RRI and double-loop learning, reinforcing the argument that there is no straightforward way to translate RRI into competence.

Building on the experimental approaches explored in the Intervention Studies, the Conceptual Studies shift focus toward embedding RRI into institutional structures and expanding its applicability across diverse educational and professional contexts.

3.4. Conceptual studies

Six studies were classified as Conceptual Studies. These works seek to expand or adapt RRI frameworks to new contexts, offering practical strategies for broader and more sustainable applications.

Many Conceptual Studies successfully incorporated RRI principles into educational methodologies, such as problem-based learning and co-creation, while fostering collaboration across disciplines, including the integration of philosophy of science with nanomedicine and diverse stakeholder training. Others developed tailored indicators for public engagement and RRI implementation, aligning them with specific regional or disciplinary needs. These efforts promoted critical reflection and reflexivity, enabling participants to assess the ethical and societal dimensions of their work. Several studies also demonstrated success in involving stakeholders and the public in co-creation processes, ensuring that research outcomes align with societal needs. However, RRI remains resource-intensive, time consuming, and dependent on high level expertise. Scaling participatory and reflective methodologies to broader contexts presents challenges, and the lack of a clear, standardized definition of RRI complicates consistent implementation. Additionally, some studies reported resistance from researchers and participants to integrating non-traditional aspects such as societal and ethical considerations. Institutional reward systems continue to prioritize traditional research outputs over RRI practices, limiting their adoption and long-term sustainability. Measuring societal impact and long-term effectiveness remains difficult, especially for qualitative and abstract outcomes, further complicating the broader application of RRI principles.

For example, Mejlgaard et al. (2019) advanced RRI education by drawing on Aristotelian ethics to instil practical wisdom (phronesis), helping students navigate ethical dilemmas with confidence. Their study, based on interviews and policy reviews, explored fostering a culture of RRI in higher education. It argued that phronesis captures the core objectives of RRI teaching, emphasizing practical ethics and decision-making in complex, uncertain contexts. The study identified key RRI-related competencies, including critical reflection, real-life case integration, and external actor engagement. It highlighted the effectiveness of problem-based learning (PBL), inquiry-based learning (IBL), and Neo-Socratic dialogue in engaging students with societal implications of science. Beyond producing skilled professionals, RRI education should cultivate responsible citizens capable of addressing ethical and societal challenges. However, barriers to adopting a phronetic perspective include resource intensity, academic reward systems prioritizing traditional research outputs, and heavy reliance on educators' personal experience. These barriers limit standardization and scalability. Additionally, conveying abstract ethical concepts to disengaged students remains challenging, though the approach offers valuable opportunities to develop practical wisdom for real-world dilemmas.

Regardless of the impressive results with the phronetic perspective, collectively, the barriers make the practical implementation of the phronetic perspective in RRI teaching a complex and contested endeavour. Recognizing the limitations of existing studies with regards to the gap between societal and scientific domains, several studies propose expansions into new domains to narrow the gap, for example in the field of art and science. Sandu et al. (2021) integrated co-creation and creativity into RRI education, using real-life case studies to connect societal and scientific domains. Their study explored the role of art and science collaborations in schools, emphasizing the active participation of learners, teachers, scientists, and societal actors to align research with societal values. The study aimed to integrate art and science into curricula, enhance creativity, communication, and civic participation, and align education with societal needs to foster inclusivity and sustainability. It presented case studies such as digital storytelling for heritage education, community-based co-creation, and open schooling initiatives. Qualitative assessments, including open-ended questionnaires and interviews, demonstrated that students developed greater creativity, critical thinking, and teamwork skills. Through digital

storytelling projects, they engaged with cultural heritage and societal issues, while collaborative learning methods such as focus groups and peer evaluations encouraged active participation. Despite its successes, the study also identified challenges. Students faced difficulties in accessing appropriate digital tools and technical expertise, and many were initially reluctant to adopt creative and participatory methods, requiring additional facilitation. The resource-intensive nature of the methodologies limited their scalability across different educational settings. Additionally, teachers required further training to achieve balance between providing structure and fostering students' creative autonomy. With that this study shows the gap between the theoretical concept of RRI and its practical application. It highlights the need for tailored training for different stakeholder groups to ensure effective implementation.

Marschalek et al. (2017) developed interactive reflection training programs aimed at diverse stakeholder groups, including policymakers, educators, researchers, civil society, and industry representatives. The training fostered a shared understanding of RRI principles and enabled participants to reflect on and implement these concepts in their own work through tailored methodologies. Post-training feedback indicated that participants felt more equipped to address RRI in their contexts, with many describing an "energizing effect" that motivated them to integrate RRI principles. The interactive approach and variety of methodologies were particularly appreciated, leading to increased familiarity with RRI, a clearer understanding of its societal and ethical relevance, and improved ability to contextualize it within their professional environments. Beyond knowledge acquisition, the training helped participants develop actionable strategies for implementing RRI, supported by tools such as the RRI Toolkit and self-reflection exercises. It also fostered dialogue across sectors, encouraging mutual learning and collaboration. However, the study identified several challenges. Differences in pre-existing knowledge and willingness to engage with RRI required customized training approaches, while the broad and abstract nature of RRI made it difficult for some participants to fully grasp and implement its principles. Recruiting and ensuring active participation from a diverse range of stakeholders was resource-intensive and complex. Although the interactive, hands-on methodology proved effective, scaling it to larger groups or different regions posed significant challenges due to the high resource investment required. Additionally, the study lacked mechanisms to assess the long-term impact of the training on participants' professional practices.

Portillo et al. (2022) sought to support researchers in applying RRI principles through tools and activities, such as the Moral-IT cards and ORBIT RRI self-assessment and co-create guidelines for RRI implementation tailored to digital and autonomous systems research (including the ethical use of personal data). They explored current RRI practices amongst researchers using two research programs as case studies, to identify barriers and facilitators for RRI implementation. They used interviews, in depth case studies, a preliminary development of a program approach to RRI, the development of a system that supported the creation, use and analysis of card-based tools and activities to support RRI and its policy impact. The Interviews and workshop discussions revealed that participants appreciated the value of RRI-focused sessions, the tools like Moral-IT cards were promoting meaningful discussions and anticipation of potential research impacts, that researchers developed actionable strategies to integrate RRI into their workflows and aligned their projects with ethical and societal considerations. However, there were some obstacles as well, like the short duration of projects that made it difficult to fully implement RRI principles, especially activities requiring significant reflection and collaboration. Prior most participants had limited exposure to RRI, resulting in uneven understanding and application of its principles. Tools like the ORBIT RRI self-assessment were time-intensive and not always tailored to specific project contexts, limiting their perceived effectiveness. Researchers lacked sufficient institutional support, such as recognition for RRI-related efforts or dedicated roles to facilitate these practices. Evaluating the impact of RRI practices was challenging, particularly for abstract concepts like reflection and societal alignment.

The difficulty of implementing RRI principles as a result of the short duration of projects, might be reduced if the Philosophy of Science would be a central feature of RRI, as advocated by Oftedal (2014). The study argues for integrating the philosophy of science into the RRI framework, emphasizing its role in examining foundational scientific assumptions, particularly in nanomedicine. Using nanomedicine as a case study, it explored how philosophical perspectives can address reductionist models and their ethical and societal implications. The study's objectives included identifying key assumptions in nanomedical research, demonstrating their ethical risks, and advocating for a broader philosophical perspective within RRI to address societal concerns. An in-depth analysis of nanomedical practices revealed dominant reductionist approaches, which focus on molecular-level causes while neglecting systemic or higher-level factors in disease causation. The study highlighted the ethical risks of reductive modelling, such as oversimplifying disease causation and potentially alienating patients from their health responsibilities. It argued for integrating systems-oriented and higher-level perspectives into nanomedical research to foster a more holistic understanding of health and disease. By aligning its findings with RRI dimensions like reflexivity and inclusivity, the study demonstrated how philosophy of science provides conceptual tools to challenge assumptions. However, integrating these perspectives into practical RRI frameworks proved challenging due to their abstract nature and perceived lack of direct applicability. Resistance from the dominant reductionist paradigm in nanomedicine and limited interaction between philosophers of science and natural scientists further hindered the adoption of broader perspectives. Additionally, the study raised concerns that the emphasis on nanomedicine might divert resources from exploring preventative or environmental approaches to health.

García-Melón et al. (2022) expanded the assessment of RRI institutionalization in Technological Centres (TCs) through a maturity framework, offering insights for future implementation. The study aimed to adapt European-level RRI public engagement indicators to the Spanish scientific and innovation context by identifying relevant aspects of public engagement, developing tailored indicators, prioritizing them based on stakeholder feedback, and validating a participatory methodology for adaptation. Using the Analytic Hierarchy Process (AHP) method, the study provided a quantitative ranking of indicators, incorporating participatory sessions and decision-making tools to ensure stakeholder ownership. Iterative feedback confirmed the relevance and applicability of the indicators, which aligned with existing gaps in Spanish science and innovation policies. A total of 20 indicators were developed, reflecting Spain's unique priorities in policy, co-creation, and competence-building. Through a multi-criteria decision-making approach, seven key indicators were prioritized for immediate implementation. The participatory process also revealed gaps in the European indicators, underscoring the need for context-specific adaptations. While stakeholders recognized the effectiveness and participatory nature of the methodology, the study also identified challenges. Disparate stakeholder views led to inconsistencies in policy-related indicator prioritization, and differences in aggregation levels (micro vs. macro) complicated comparisons. Although effective in Spain, the methodology required adaptation for broader and international contexts. Additionally, quantitative indicators alone were insufficient to capture qualitative aspects of public engagement, such as its depth or societal impact.

4. Synthesizing successes and obstacles in RRI education

Having analysed the studies thematically, we now synthesize the findings to identify overarching themes of successful implementation and recurring obstacles in RRI education. We start with the definitions of these themes, give some examples, and give some deliberations with regards to the finds to finish with a deeper understanding of the relations

between the themes.

4.1. Successfully implemented themes

There were five themes that were reported to be successfully implemented. We list them in the upcoming sections.

4.1.1. The contextualization of RRI education

The Contextualization of RRI Education refers to the mapping of RRI competencies into specific educational or regional settings, creating frameworks that align learning with real-world applications and societal challenges. It is reported as the most successfully implemented (14 out of 17 studies). One of the examples that shows an immediate impact on society is the tailored RRI-based policymaking with citizen panels to match territorial challenges in Greece and Bulgaria studied by Vestergaard Bidstrup et al. (2024). However some of the reported implementations seem to stay within the academic or educational field itself, like the embedding of RRI within a local university, using a two-tier model and created frameworks like the Committee for Engaged Research studied by Levikov et al. (2020) or the mapping of Science, Technology, Engineering, Arts and Mathematics (STEAM) education competencies to RRI categories via participatory workshops with teachers, as studied by Malagrida et al. (2022), suggesting that adjacent fields of study, that already work on engagement or an educational approach that integrates STEAM to encourage inquiry, creativity, and critical thinking could be more susceptive towards ethical and critical thinking themed subjects like RRI.

4.1.2. Reflection and reflexivity

With Reflection and Reflexivity, we refer to the promotion of critical thinking and reflexivity which enables participants to evaluate the ethical and societal dimensions of their work. This includes fostering awareness of assumptions, values, and the broader impact of science and technology. This is reported as the second most successful (12 out of 17 studies). Reflexivity was often successfully encouraged through participatory, interdisciplinary, or philosophical methods. But it remains unclear what depth of reflection is necessary to effectively contribute to enable responsible research or innovation having real world effects, except the Improved regional policy recommendations in Greece and Bulgaria Vestergaard Bidstrup et al. (2024), most of the studies emphasized reflexivity in theory but did not trace clear outcomes or behavioural changes. Actual behavioural or institutional change because of reflexivity seems harder to measure or demonstrate. The studies repeatedly cite difficulty in scaling reflective methods and measuring long-term impact (especially in soft or emotional learning dimensions). Mixed-method approaches were suggested as better suited to capture the depth of reflexivity, however these were underused.

4.1.3. Participatory methods

This refers to the act of involving stakeholders, students, and citizens in co-creation processes that enhances inclusivity and aligns research and education with societal needs. It is reported as the third most successful (11 out of 17 studies). The Participatory methods emerge as a key mechanism for enhancing both reflexive learning and practical alignment with societal needs. The Participatory approaches consistently helped participants question assumptions, consider diverse values, and connect abstract concepts to real-world issues. By involving the citizens, stakeholders, students, and professionals, participatory methods ensured that research and education respond to societal needs and the participatory approaches had a more observable connection to behaviour change than abstract instruction alone. However, the Participatory Methods often require facilitation, time, and expertise, making them hard to scale. Some of the methods (like the citizen panels) were one-day events, limiting the depth of deliberation. The success of Participatory Methods depends heavily on participant willingness and the competences of the facilitator.

4.1.4. Interdisciplinary collaboration

This refers to bridging disciplines, such as the integration of philosophy with science and stakeholder engagement, enriches perspectives and embeds RRI principles in education and practice, is reported as the fourth most successful (9 out of 17 studies). Interdisciplinary collaboration is deemed a core enabler of both reflexivity and real-world relevance in RRI education, as highlighted across several studies in the review. When different disciplines interact, especially those combining social sciences, ethics, philosophy, and Science, Technology, Engineering, and Mathematics (STEM), it is reported that they broaden the scope of what's considered relevant and reflective. For example, in the study done by Hesjedal et al. (2020) participants in a biotech PhD course reflected more deeply on societal roles in science when exposed to interdisciplinary views. The interdisciplinary collaboration seems to support Institutional Integration. This is substantiated by Malagrida et al. (2022), they showed that STEAM education, with its natural interdisciplinarity, was a conducive platform for embedding RRI principles via teacher collaboration and Portillo et al. (2022) involved engineering, ethics, and digital research groups to co-create tools like the ORBIT RRI toolkit. This seems to support institutional integration; however, it has challenges as well. Challenges like epistemic differences which are visible as clashes of paradigms, a lack of a common language which creates difficulty in achieving mutual understanding across disciplines, Institutional silos which have reward systems that discourage collaboration beyond traditional academic metrics and there is a perceived irrelevance as ethics and societal concerns are often viewed as "soft" or peripheral, this is especially so in the hard sciences.

4.1.5. Institutional integration

This refers to the embedding of RRI principles into governance and compliance mechanisms to offer sustainability for RRI practices beyond project funding. Only 6 out of the 17 studies show to have success with regards to Institutional Integration. Insight into the low success rate of the institutional integration is given by Mejlgaard et al. (2019), they found that there is an emphasis on traditional metrics of excellence, as current reward systems and the dominant understanding of research excellence does not accommodate the transition towards a greater focus on responsibility in research and innovation. Measures of merit, performance, and success tend to favour traditional components of academic work such as publishing in high impact journals and patenting the results of research and innovation activities. They also discovered that there seems to be a marginalization of RRI teaching, which is often considered peripheral, and under constant threat of funding cuts as new university administration gives it a low priority. And finally, there are tensions in engineering disciplines, ethical considerations are seen as detached from core engineering practices and applied ethics research is disregarded as lower-status and less important.

4.2. The obstacles

There were six themes that were reported as an obstacle. We list them in the upcoming sections.

4.2.1. Resource intensity and scalability

With Resource intensity and Scalability, we refer to the high demand for time, facilitation, and expertise in RRI education methods, particularly those involving participatory and reflective elements, which limits their scalability. This is reported as an obstacle in at least 10 out of the 17 studies reviewed, it is one of the most consistently cited challenge. For example, Malagrida et al. (2022) demonstrated that while participatory workshops linking STEAM education to RRI principles were effective, they required substantial teacher training and coordination. Similarly, Lukovics et al. (2019) highlighted that the STIR method, while impactful, was too intensive for large-scale adoption. Most studies report methods that are seen as vital to meaningful RRI education; however, their labour-intensive nature makes them difficult to

implement broadly. This challenge is compounded by short project timelines and the absence of sustained institutional support, often leaving successful pilot projects unscalable beyond their original context.

4.2.5. Institutional and cultural barriers

With Institutional and Cultural Barriers, we refer to limited institutional recognition of RRI practices and the academic reward systems favouring traditional outputs that hinder broader adoption and sustainability. This has been found in 13 studies, particularly those examining university structures or STEM focused disciplines. For example, Mejlgaard et al. (2019) showed that current academic reward structures emphasize publications and patents, sidelining RRI teaching and implementation. Portillo et al. (2022) reported that without institutional incentives or roles, RRI efforts often faded after project completion. Institutional barriers reduce access to the time and resources needed for RRI, reinforce disciplinary silos, and hinder attempts to embed responsibility as a core value.

4.2.2. Definitional and conceptual challenges

With Definitional and Conceptual Challenges, we refer to the persistent ambiguity regarding what RRI entails and how it differs from the nature of science, the philosophy of science and related frameworks which undermines consistent integration and understanding. In at least 12 out of the 17 studies it was pointed out that definitional uncertainty is a barrier to both education and institutional adoption. For example, Stahl et al. (2023) reported confusion among trainers and institutions regarding whether RRI is empirical or normative, and how to assess its impact. Hesjedal et al. (2020) noted that participants struggled to articulate what they had learned from RRI training, highlighting conceptual vagueness. When foundational concepts lack clarity, educational programs struggle to define learning objectives, and institutions hesitate to embed RRI into curricula or policy. This ambiguity hinders the development of competencies and perpetuates resistance, particularly in disciplines where abstract ethical constructs are viewed as peripheral. These definitional and conceptual challenges are a primer for other obstacles uncovered and deliberated on in Section 4.3.

4.2.3. Short-Term focus vs. long-term impact

With Short-Term Focus vs. Long-Term Impact, we refer to the emphasis on immediate outcomes in educational interventions, leaving the long term societal, institutional, or behavioural effects of RRI training unassessed or unsupported. This is a recurrent theme in 10 studies, often linked to evaluation limitations or funding models. For example, Vestergaard Bidstrup et al. (2024) illustrated how short-term citizen panels provided valuable input yet lacked follow-up mechanisms to assess long-term policy influence. Stahl et al. (2021) discussed the difficulty of embedding RRI beyond project timelines. Without longitudinal strategies, it is impossible to determine whether RRI education truly cultivates lasting responsibility. This short-term focus perpetuates shallow engagement, undermining RRI's transformative potential with regards to societal goals, especially when societal goals need long term engagement.

4.2.4. Challenges in competency development

With Challenges in Competency Development, we refer to the struggle of Participants to articulate and apply RRI concepts in practice, emphasizing the difficulty of translating RRI into concrete competencies. This is cited in 5 studies, often in conjunction with measurement difficulties and conceptual ambiguity. For example, Hesjedal et al. (2020) noted that while participants experienced shifts in perspective, they had difficulty applying RRI concepts in real-world contexts. Malagrida et al. (2022) found that students lacked the ability to distinguish scientific facts from ethical implications, limiting competence development. When students cannot practice RRI, its education fails to generate meaningful change. The lack of applied uniform learning tools and clear competency models exacerbates this challenge, particularly in disciplines that lack tradition in ethics or social reflection.

4.2.6. Measurement and evaluation challenges

With Measurement and Evaluation Challenges we refer to the struggle with existing metrics to capture the qualitative, long-term, and often abstract outcomes of RRI education, such as ethical awareness, reflexivity, or social responsiveness. Identified in 9 studies, especially in connection to the limits of traditional evaluation tools. For example, Heras and Ruiz-Mallén (2017) highlighted that standard tools like closed questionnaires overlook emotional and ethical learning dimensions. Stahl et al. (2023) noted the lack of universal criteria for assessing RRI competencies, particularly in interdisciplinary settings. This inability to measure what matters most in RRI education weakens arguments for its legitimacy and funding.

4.3. Interpreting the tensions between successes and obstacles

Several factors identified as successes in the reviewed studies also appear, under different conditions, as obstacles. This tension reflects the context-dependent nature of progress in RRI education. For instance, institutional integration was often cited as a positive outcome, signalling the inclusion of RRI principles in curricula, strategy documents, and project design. Yet the same institutionalization may become a barrier when it leads to bureaucratic formalization, tokenistic compliance, or rigid assessment frameworks that undermine reflexivity. Similarly, interdisciplinary collaboration was praised as a driver of innovation but also reported as a source of tension due to differing epistemic traditions and communication styles. These examples illustrate that RRI education advances through ambivalent dynamics, where enabling structures may also constrain transformative learning. Recognizing these dualities provides a more realistic understanding of why progress remains uneven and helps explain how the field becomes caught in a self-reinforcing cycle of partial success and persistent limitation. These ambivalent dynamics between progress and limitation form the foundation of what can be described as a vicious reinforcing cycle in RRI education.

4.4. The building blocks of a vicious reinforcing cycle

The previous sections highlighted both promising practices and persistent obstacles in RRI implementation. A deeper analysis reveals that many of these challenges share a common root: Conceptual and philosophical ambiguity regarding what RRI is, what kind of knowledge it represents, and how it relates to scientific inquiry itself. Notably, definitional and conceptual uncertainty is reported as a barrier in at least 12 of the 17 included studies, and it frequently co-occurs with difficulties in assessment and competency development. Without a clear epistemological and ethical foundation, educational initiatives remain fragmented, producing short-lived successes that rarely translate into institutional or behavioural change.

This review identified three dominant interpretative strands in the literature, each of which frames RRI differently:

One prominent perspective positions RRI as a strategy where stakeholders become mutually responsive to one another, actively anticipating the outcomes of research and innovation to address the "grand challenges" of our time, for which they share responsibility (Von Schomberg, 2013).

This perspective presupposes that grand challenges are knowable and that stakeholders can meaningfully anticipate outcomes. From a science philosophy standpoint, particularly one informed by fallibilism and complex systems theory, this is problematic. Scientific and technological trajectories are often unpredictable, nonlinear, and emergent. It risks instrumentalizing science, reducing it to a tool for solving predefined societal problems, thereby downplaying the epistemic autonomy and exploratory nature of scientific inquiry. It also implies a degree of consensus about what constitutes a "grand challenge", ignoring the

pluralism and contestation inherent in both science and society.

Alternatively, RRI is viewed as a policy-driven framework introduced in the early 2010s by the European Commission (Owen & Pansera, 2019). This perspective emphasizes fostering inclusive and sustainable research and innovation, particularly through co-creation and co-production with society. Within this perspective RRI seeks to align research and innovation with societal values, needs, and expectations, focusing on addressing societal grand challenges.

By positioning RRI as a policy tool to align science with societal values and needs, this perspective risks normative paternalism, assuming that values and needs can be clearly identified and agreed upon, which contradicts philosophical insights into the value-influenced and contestability of science. It encourages conformity over critique, potentially muting scientific dissent or radical innovation that challenges current societal values, and it reflects a linear model of governance, assuming values can be translated into research agendas without loss or distortion, a notion critiqued in STS (Science and Technology Studies) and critical realism.

A third perspective suggests that stakeholder inclusion, participation, and partnership foster harmony and consensus by reducing divergent judgments and value frameworks to a shared understanding (Blok, 2019). In this context, dialogues on innovation are seen as opportunities for ethical engagement, shifting from merely understanding ethical considerations to embodying ethical behaviour in practice. This evolution emphasizes the ethical responsibilities of stakeholders in shaping research and innovation outcomes collaboratively.

This perspective implies that stakeholder participation will lead to ethical harmony and shared understanding, which can be counterproductive due to the reduction of ethical pluralism. It treats ethical disagreement as something to be managed away, rather than a productive force in shaping responsible innovation. Furthermore, it depends on deliberative idealism, assuming dialogue will naturally lead to shared ethics, overlooking power dynamics, strategic behaviour, and incommensurable worldviews. Therefore, this perspective risks creating a veneer of legitimacy without grappling with deeper epistemological and ontological tensions in how knowledge and responsibility are constructed.

While each perspective contributes valuable insights, they also reflect distinct philosophical assumptions about predictability, normativity and consensus that are rarely examined explicitly. As a result, educational efforts often oscillate between procedural compliance and moral aspiration, without a coherent understanding of RRI's epistemic status.

This lack of philosophical grounding creates a vicious reinforcing cycle. Ambiguous definitions lead to inconsistent curricula and assessment criteria, which in turn undermine institutional recognition and resource allocation. Weak institutionalization then prevents long-term investment in conceptual development, perpetuating the very ambiguity that impedes progress. The cycle thus connects definitional vagueness, methodological inconsistency, and institutional inertia into a self-sustaining loop.

Several studies in this review illustrate different stages of this cycle. Stahl et al. (2023) and Hesjedal et al. (2020) reported persistent uncertainty over whether RRI is empirical or normative, while Malagrida et al. (2022) showed that teachers struggled to separate ethical reflection from scientific reasoning. Oftedal (2014) and Mejlgaard et al. (2019) demonstrated that introducing philosophical reflection, whether through analysis of scientific assumptions or through Aristotelian practical wisdom (phronesis), may break this loop by providing conceptual clarity and moral coherence.

These examples suggest that engaging explicitly with the philosophy of science offers a pathway out of the reinforcing cycle. Importantly, we do not propose philosophy of science as additional curricular content, but as a conceptual support that helps address repeatedly reported problems in the literature, like the conceptual vagueness, the difficulty distinguishing ethical reflection from scientific reasoning, and the assessment of reflexivity and related competencies.

This may be done by clarifying how responsibility, reflexivity, and scientific reasoning interrelate, that clarification enables more consistent pedagogy and institutional embedding of RRI principles. The interrelations between these subjects are summarised schematically in Fig. 3, which visualises how conceptual ambiguity, inconsistent practice, and weak institutionalisation form a self-perpetuating cycle that constrains the development of coherent RRI education.

This figure illustrates how unresolved conceptual and philosophical ambiguities surrounding Responsible Research and Innovation (RRI) education create a self-reinforcing cycle. Ambiguous definitions lead to inconsistent educational practices, which reduce institutional support and result in superficial implementation. Limited reflection and short-term evaluation prevent deeper theoretical development, thereby perpetuating the original ambiguity.

The persistence of this reinforcing cycle indicates that the main barriers to advancing RRI education are not merely practical but conceptual. The reviewed literature collectively suggests that without addressing the underlying philosophical questions reform efforts risk remaining superficial or short-lived. Breaking this cycle therefore requires a deeper theoretical foundation that can clarify RRI's epistemic and ethical dimensions while supporting coherent educational design. This insight leads directly to our concluding discussion on the role that philosophy of science could play in strengthening the conceptual and institutional foundations of RRI education.

5. Conclusions and future research

5.1. conclusions

This review examined the emerging field of RRI education and revealed that RRI is conceptually fragmented and methodologically diverse. Across the 17 studies analysed, we found a recurring theme: The lack of definitional and philosophical clarity about what RRI entails, how it relates to existing educational paradigms, and what kinds of knowledge and competencies it should cultivate.

Several studies reported uncertainty about whether RRI should be approached as an empirical, normative, or procedural framework (Stahl et al., 2023; Hesjedal et al., 2020), while others highlighted educators' difficulty in distinguishing scientific from ethical reasoning (Malagrida et al., 2022). This persistent conceptual ambiguity underpins many of the practical challenges, such as the inconsistent assessment of reflexivity, limited institutional uptake, and the difficulty of developing measurable competencies.

While only a subset of the included studies explicitly propose that philosophy of science could play a role in strengthening the conceptual and institutional foundations of RRI education, the underlying problem it addresses, the conceptual ambiguity and its downstream effects on pedagogy, assessment, and institutional embedding, recurs across the reviewed literature. Furthermore, within this fragmented field, a small number of studies demonstrated that introducing explicit philosophical reasoning strengthens both conceptual coherence and educational depth. Oftedal (2014) showed that engaging the philosophy of science can reveal hidden epistemic assumptions in research practices, broadening RRI's ethical and conceptual foundations. Similarly, Mejlgaard et al. (2019) illustrated that teaching RRI through a phronetic, or practical-wisdom-based, approach helps learners navigate moral complexity and link theoretical reflection with professional decision-making. Bardone, Burget and Pedaste (2023) likewise argued that mapping RRI within established educational philosophies such as the Nature of Science framework increases its intelligibility and legitimacy. These findings collectively suggest that the recurring obstacles in RRI education, conceptual vagueness, weak institutional integration, and inconsistent competence development, stem from an insufficient philosophical grounding.

Therefore, we suggest that systematically integrating philosophy of

Fig. 3. The vicious reinforcing RRI cycle of conceptual ambiguity.

Table 1
Themes that are reported as successfully implemented in the studies.

Succesfull implementations per study, 1 = reported, 0 = not reported

Studies that reported as succesfully implemented:	Nika Levikov, Daniela Quacinella, Edward Duca	Doris Elster, Tanja Barendziak, Julia Birkholz	Mirjam Schuijff, Anne M. Dijkstra	Emanuele Bardone, Mirjam Burget, Margus Pedaste	Stahl et al. (2023)	Maria Heras, Isabel Ruiz-Mallén	Vestergaard Bidstrup et al.	Malagrida et al.	Stahl et al. (2021)	Hesjedal et al.	Mejlgaard et al.	Sandu et al.	Marschalek et al.	Portillo et al.	Gry Oftedal	García-Melón et al.	Lukovics et al.	Total horizontal axis
Reflection and reflexivity	0	1	1	1	0	0	1	1	1	1	1	0	1	1	1	0	1	12
Interdisciplinary Collaboration	1	0	1	0	0	0	1	1	1	0	0	1	1	0	1	1	0	9
Participatory Methods	0	0	1	0	0	1	1	1	1	1	0	1	1	1	0	1	1	11
Contextualization of RRI Education	1	1	0	1	1	1	1	0	1	1	1	1	1	0	1	1	1	14
Institutional Integration	1	1	1	0	1	0	0	0	1	0	0	0	0	0	0	0	1	6
Total vertical axis	3	3	4	2	2	2	4	3	5	3	2	3	4	2	3	3	4	

science into RRI education seems a promising direction that could address this core deficiency. This integration is not intended as a standalone philosophy curriculum, but as a pedagogical support embedded in existing RRI teaching practices to make core distinctions and learning outcomes teachable and assessable.

Philosophical reflection offers the analytical tools needed to clarify RRI's epistemological status, articulate its normative dimensions, and reconcile tensions between scientific autonomy and societal accountability. Doing so would not replace existing pedagogical models but rather provide the conceptual support that connects them, enabling more coherent curricula, more reliable assessment, and more sustainable institutional support.

5.2. Limitations and scope conditions

This review is an exploratory thematic synthesis of an emerging field, based on 17 peer-reviewed studies published over the last two decades. As such, it should be read as identifying recurring patterns rather than providing definitive estimates of prevalence or effectiveness.

First, our findings depend on the reporting quality of the included studies: relevant successes or obstacles may have occurred but remained unreported, which can affect how themes appear in Tables 1 and 2. Second, the search strategy (databases, keywords, language, and emphasis on peer-reviewed outlets) may have excluded relevant work, particularly non-English publications and grey literature where educational innovation is often documented. Third, the corpus is dominated by European contexts, which may limit global generalisability given RRI's strong linkage to European policy frameworks. Finally, the inductive categorisation and thematic clustering involve interpretive judgement; while used here to enhance comparability and practical utility, alternative categorisations could foreground different relationships between studies.

5.3. Future research directions

Future research on RRI education would benefit from broadening evidence coverage by extending searches to additional databases, non-English outlets, and grey literature, thereby capturing practice-based innovations and reducing publication bias.

In parallel, the field would gain from improved reporting and comparability through the development and adoption of clearer reporting conventions for RRI education studies, including transparent descriptions of context, learner populations, intervention components, duration, and assessment approaches, which would enable stronger cross-study synthesis and more robust meta-evaluation.

Given RRI's close association with European policy frameworks, further work should also diversify geographical contexts by conducting studies outside Europe and, where possible, explicitly comparative designs to examine which obstacles and success factors persist across institutional and cultural settings.

To address concerns about short-termism, longitudinal designs and follow-up measurements are needed to determine whether RRI education yields sustained behavioural, institutional, or policy effects beyond project timelines. Moreover, research should incorporate practitioner and learner perspectives by complementing literature-based synthesis with interviews, participatory inquiry, or co-analysis with educators, students, and practitioners to identify which challenges are most consequential in practice and which pedagogical supports are feasible under real-world constraints.

All these advances should be coupled with empirical testing of pedagogical supports, such as mapping RRI to established educational paradigms, competency-based rubrics, tool-supported reflection, and, where appropriate, philosophical or epistemic supports to establish which approaches most effectively reduce conceptual ambiguity and improve assessment across disciplines.

Finally, Future research should examine how Philosophy of Science

Table 2
Obstacles for RRI education found in the studies.

Obstacles found per study, 1 = reported, 0 = not reported																		
Obstacles found during the study:	**Nika Levikov, Daniela Quacinella, Edward Duca**	**Doris Elster, Tanja Barendziak, Julia Birkholz**	**Mirjam Schuijff, Anne M. Dijkstra**	**Emanuele Bardone, Mirjam Burget, Margus Pedaste**	**Stahl et al. (2023)**	**Maria Heras, Isabel Ruiz-Mallén**	**Vestergaard Bidstrup et al.**	**Malagrida et al.**	**Stahl et al. (2021)**	**Hesjedal et al.**	**Mejlgaard et al.**	**Sandu et al.**	**Marschalek et al.**	**Portillo et al.**	**Gry Oftedal**	**García-Melón et al.**	**Lukovics et al.**	Total horizontal axis
Resource Intensity and Scalability	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	15
Definitional and Conceptual Challenges	0	0	1	1	0	0	1	1	0	1	1	1	1	1	1	1	1	12
Short-Term Focus vs. Long-Term Impact	0	1	1	0	1	1	1	1	1	1	0	0	1	0	0	0	1	10
Challenges in Competency Development	0	1	0	0	1	1	0	0	0	0	0	0	0	1	0	0	1	5
Institutional and Cultural Barriers	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	0	1	13
Measurement and Evaluation Challenges	0	1	1	0	1	1	1	0	1	1	0	0	1	0	0	0	1	9
Total vertical axis	2	4	5	2	4	5	5	4	4	3	3	3	5	4	3	2		

based instruction can operationalize reflexivity, anticipation, and inclusiveness as measurable competencies across disciplines.

CRediT authorship contribution statement

Ferdinand Griesdoorn: Writing – review & editing, Writing – original draft, Visualization, Resources, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Maarten Kroesen:** Validation, Supervision. **Pieter Vermaas:** Supervision.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Bardone, E., Burget, M., & Pedaste, M. (2023). The RRI map: Making sense of responsible research and innovation in science education. *Journal of Responsible Innovation, 10* (1), Article 2198183. https://doi.org/10.1080/23299460.2023.2198183

Blok, V. (2019). From participation to partnership: Towards a responsible dialogue on responsible research and Innovation (RRI). *Journal of Responsible Innovation, 6*(1), 1–6. https://doi.org/10.4337/9781784718862.00024

Elster, D., Barendziak, T., & Birkholz, J. (2020). Science education as a trigger for RRI structural change (Note#5) in STARBIOS2 (Structural Transformation to Attain Responsible BIOSciences) final version of the guidelines and a model on RRI implementation in biosciences structures, 87–92, https://ec.europa.eu/research/participants/documents/downloadPublic?documentIds080166e5c7f639e4&appIdPPGMS (October 2025).

EU. (2022). Horizon Europe. *The Next EU Research & Innovation Investment Programme*, 2021–2027. https://ec.europa.eu/info/sites/default/files/research_and_innovation/strategy_on_research_and_innovation/presentations/horizon_europe_en_investing_to_shape_our_future.pdfandhttps://researchandinnovation.ec.europa.eu/funding/fundingopportunities/funding-programmesandopencalls/horizon-europe_en (November 2022).

García-Melón, M., Gómez-Navarro, T., Gonzalez-Urango, H., & Corona-Sobrino, C. (2022). Adapting RRI public engagement indicators to the Spanish scientific and innovation context: A participatory methodology based on AHP and content analysis., Central. *European Journal of Operations Research, 30*, 1483–1512. https://doi.org/10.1007/s10100-022-00796-3

Gough, D., Oliver, S., & Thomas, J. (2012). *An Introduction to Systematic Reviews. in Systematic Reviews by Sage Publications. an Introduction to Systematic Reviews | Sage Publications Ltd*.

Heras, M., & Ruiz-Mallén, I. (2017). RRI indicators in science education. *Science and Society Studies, 6*(1), 45–63. https://doi.org/10.1080/09500693.2017.1392643

Hesjedal, M. B., Åm, H., Sørensen, K. H., & Strand, R. (2020). Transforming scientists' Understanding of science–Society relations. Stimulating double-loop learning when teaching RRI. *Science and Engineering Ethics, 26*, 1633–1653. https://doi.org/10.1007/s11948-020-00208-2

Levikov, N., Quacinella, D., & Ducac, E. (2020). embedding RRI in a higher education institution: Lessons learned from Malta in the. *International Journal of Higher Education Management (IJHEM)*. https://doi.org/10.24052/ijhem/v07n01/art-2. Vol. 7 Number 1fatcat:K2ojumwjyjf7poo5w2s6q72wc4.

Lukovics, M., Udvari, B., Nádas, N., & Fisher, E. (2019). Raising awareness of researchers-in-the-making toward responsible research and innovation. *Journal of the Knowledge Economy, 10*, 1558–1577. https://doi.org/10.1007/s13132-019-00624-1

Malagrida, R., Klaassen, P., Ruiz-Mallén, I., & Broerse, J. E. W. (2022). Towards competencies and methods to support Responsible Research and Innovation within STEAM secondary education – the case of Spain. *Research in Science & Technological Education*. https://doi.org/10.1080/02635143.2022.2123790

Marschalek, I., Schrammel, M., Unterfrauner, E., & Hofer, M. (2017). Interactive reflection trainings on RRI for multiple stakeholder groups. *Journal of Responsible Innovation, 4*(2), 295–311. https://doi.org/10.1080/23299460.2017.1326262

Mejlgaard, N., Christensen, M. V., Strand, R., Buljan, I., Carrió, M., Cayetano i Giralt, M., Griessler, E., Lang, A., Marušić, A., Revuelta, G., Rodríguez, G., Saladié, N., & Wuketich, M. (2019). *Teaching Responsible Research and Innovation: A Phronetic Perspective in Science English Ethics, 25*, 597–615. https://doi.org/10.1007/s11948-018-0029-1

Oftedal, G. (2014). Philosophy of science and RRI: Insights from Nanomedicine. *Ethics and Philosophy in Science, 8*(3), 193–211. https://doi.org/10.1186/s40504-014-0005-8

Owen, R., & Pansera, M. (2019). Responsible Innovation and Responsible Research and Innovation. *Oxford Research Encyclopedia of Business and Management*. https://doi.org/10.4337/9781784715946.00010

Portillo, V., Craigon, P., Dowthwaite, L., Greenhalgh, C., & Perez-Vallejos, E. (2022). Supporting responsible research and innovation within a university-based digital research programme: Reflections from the "hoRRIzon. *project Journal of Responsible Technology, 12*. https://doi.org/10.1016/j.jrt.2022.100045

Sandu, P., Tudisca, V., & Valente, A. (2021). Co-creating in schools through art and science, lessons learned in community engagement within the responsible research and innovation framework. *Published as a Part of Springer Briefs in Research and Innovation*. https://link.springer.com/book/10.1007/9783030726904.

Schuijff, M., & Dijkstra, A. M (2020). Practices of responsible Research and Innovation: A review in. *Science and Engineering Ethics, 26*(2020), 533–574. https://doi.org/10.1007/s11948-019-00167-3. Nature B.V.

Stahl, B.C., Aicardi, C., Brooks, L., Craigon, P.J., Cunden, M., Burton, S.D., De Heaver, M., De Saille, S., Dolby, S., Dowthwaite, L., Eke, D., Hughes, S., Keene, P., Kuh, V., Portillo, V., Shanley, D., Smallman, M., Smith, M., Stilgoe, J., Ulnicane, I., Webb, H., (2023) assessing responsible innovation training in the journal of responsible technology, Volume 16, 2023, https://doi.org/10.1016/j.jrt.2023.100063.

Stahl, B. C., Akintoye, S., Bitsch, L., Bringedal, B., Eke, D., Farisco, M., Grasenick, K., Guerrero, M., Knight, W., Leach, T., Nyholm, S., Ogoh, G., Rosemann, A., Salles, A., Trattnig, J., & Ulnicane, I. (2021). From responsible research and innovation to responsibility by design. *Journal of Responsible Innovation, 8*(2), 175–198. https://doi.org/10.1080/23299460.2021.1955613

Stilgoe, J., Owen, R., & Macnaghten, P. (2013). Developing a Framework for responsible Innovation,. *In Research Policy, 42*(9), 1568–1580. https://doi.org/10.1016/j.respol.2013.05.008

Tweed, A. (2009). *Designing Effective Science Instruction* (p. 177). published by the national science teachers association (NSTA-press).

Vestergaard Bidstrup, M., Gabova, S., Kilintzis, P., Samara, E., & Kouskoura, A. (2024). The RRI Citizen Review Panel: A public engagement method for supporting responsible territorial policymaking. *Journal of Innovation and Entrepreneurship, 13* (2). https://doi.org/10.1186/s13731-023-00360-6

Von Schomberg, R. (2013). Responsible innovation: Managing the Responsible emergence of science and innovation in society. In R. Owen, J. Bessant, & M. Heintz (Eds.), *Responsible Research and Innovation: Managing the Responsible Emergence of Science and Innovation in Society* (pp. 51–74). John Wiley & Sons. https://doi.org/10.1002/9781118551424.ch3. Ltd.