Horizon Europe Programme

HORIZON-WIDERA-2021-ERA-01-44: Societal trust in science, research and innovation

Deliverable 3.1 Conceptual, methodological and ethical framework

Disclaimer

Funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Executive Agency (REA). Neither the European Union nor the granting authority can be held responsible for them.

Funded by
the European Union

Project Acronym	IANUS
Project Title	Inspiring and Anchoring Trust in Science, Research and Innovation
Grant no	101058158
Type of action	Research and Innovation Action (RIA)
Start date	01/06/2022
Duration in months	36 months

Deliverable information

Deliverable leader	CEUT
Lead author(s)	**CEUT**
Contributors	**LSE, EUR, RU, KIT, ISI, RCLL**
Contractual delivery date	**30.09.2023**
Delivery date	**31.10.2023**
Dissemination level	

Version history

Version	Date	Author(s)	Summary of changes
0.1	05.10.2023	CEUT	Draft for presentation at the Vilnius meeting
0.2	31.10.2023	CEUT	Feedback from Vilnius meeting incorporated

Table of contents

List of figures

List of abbreviations used in the document

RE: Research Ethics
RI: Research Integrity
RRI: Responsible Research and Innovation

Executive Summary

Deliverable 3.1 provides conceptual clarification of trust and distrust and related concepts, identifies the key reasons of distrust in science and identifying effective mechanisms to build warranted trust in science, prepares the ground for recommendations for research policy.

This deliverable is based on several internal deliverables which are available for internal use and, when appropriate, will be revised and published as academic output. For this deliverable, work package 3 (WP3) conducted 5 tasks (T.3.1-T3.5). Below, you can find an executive summary organised per task.

Task 3.1 "Conceptual framework and basic principles of trust and distrust" surveys the philosophical and sociological literature on trust and develops a theoretical framework for coping with growing distrust in science (and more broadly in science, technology, and evidence-based innovation). However, our analysis revealed that distrust isn't detrimental, nor is trust always beneficial. What's essential is justified trust and justified distrust.

Trust and distrust are a continuum. The neutral point, lack of (dis)trust may be healthy scepticism. Based on the philosophical analysis of trust and distrust, we described the structure of the mental act of trust and the mechanism of building (dis)trust in science. Such map can be used to as a guide for a meta-analysis of the wealth of empirical studies in trust in science (to specify which aspects of trust are in focus in the different studies) and allow us for a more comprehensive understanding of the state of (dis)trust of science in the fast-changing society.

Our exploration emphasized the significance of emotions and how an individual's inclination to trust, which involves relinquishing control and exposing oneself to vulnerability, underpins all types of trust. By focusing on distrust as affective state, we put forth a new perspective: trust in science is not solely based on the credibility of the entity being trusted (whether a scientist or a scientific institution) and the education of the trustor, but also on the personal characteristics, past experiences, and values of the individual placing their trust.

While the cognitive deficit model suggested that distrust in science stemmed from a lack of knowledge, as noted by Wynne (1991) and Simis et al. (2016), it is evident now that enhancing trust isn't merely about better educating the public. The reasons for distrusting scientists and accepting or rejecting scientific findings are diverse. They are intertwined with psychological elements like readiness to lay oneself open, sociological aspects such as trust in other social institutions, and personal factors including personality traits, past experiences, moral values, religious beliefs, and political orientations. This sheds light on why, even when scientific institutions and scientists are credible, they can still face distrust.

However, even if trust in science depends on so many factors, it does not mean that we should not make efforts to make scientists, scientific institutions, and science as a system

more trustworthy. Initiatives to make science more trustworthy have taken the form of research ethics, research integrity and responsible research and innovation.

Task 3.2 "Contextualising the Framework" builds has three objectives. First, to characterize scientific research, technology, and innovation as a complex multi-level system and to hypothesize on the way in which different levels of the system interrelate. Second, to outline some of the current challenges confronting the research system. And thirdly, to report on contemporary actions designed to ameliorate the research system with new policy options and procedures that address these challenges in holistic approaches.

Task 3.3 "Trust and Truth" analysed the concept of trust as one of the conditions affecting trust in science which is of utmost importance in an era when phrases such as "post-truth" and the "crisis of truth" are omnipresent. Traditionally, philosophers tended to emphasize the distinction between truth (validated knowledge) and mere opinion and tried to define and consolidate concepts such as validated knowledge (e.g., via methods based on induction and deduction). More recently, philosophers, but also science studies experts in other fields (e.g., history of science and STS) have studied how knowledge claims are constituted in specific social contexts and how knowledge production may be vulnerable to political and economic interests rather than disinterested. At the same time, such scholars aimed to make knowledge more inclusive, e.g., more open to practical, experiential and indigenous forms of knowledge. Rather than exclusively focusing on expert knowledge, these authors aspired a democratization of knowledge. We acknowledge the positive value of these scholarly efforts. At the same time, in view of the challenges societies are facing, we need new ways to distinguish validated truth claims from mere speculation and warranted from unwarranted trust in truth claims. In other words, rather than turning back towards the traditional situation, concepts such as truth and expertise must be redefined in a dynamical knowledge landscape.

This is not an easy challenge. Research fields are facing many challenges, ranging from the replication crisis in the social sciences up to the increasing reliance on AI and algorithms which has given rise to the claim that anything can be "validated". RRI is not only a response to mistrust of science that is triggered by scandals related to research misconduct, but it may also involve questioning the quality and trustworthiness of large-scale, competitive scientific knowledge production on an industrial scale. How to ensure that "big science" is responsive to societal needs and concerns rather than turning into a self-serving enterprise, driven by researchers as knowledge entrepreneurs?

Many voices in contemporary societies are questioning whether scientific information is sufficiently valid, disinterested and objective. At the same time, in order to effectively address the global disruptions facing Europe and the rest of the world, evidence-based insights and potential solutions provided by science are indispensable. Therefore, the credibility, reliability and trustworthiness of science is an issue of crucial importance (Oreskes 2019). In order to restore a trust relationship between science and society, science and society must be brought closer together. And this entails developing a socially robust, inclusive and participatory concept of truth. This will be the guiding question for

our work in subtask 4.1.1 and 4.3 (co-creation event on trust and truth in virology), while using this experience to elaborate our conceptual views on trust and truth.

Task 3.4 "Ethical aspects of trust in science". Here, the relationship between ethics and trust in the scientific arena is dissected, the analysis earmarks ethics as the linchpin of trust and explores the ethical intricacies of decision making amid uncertainty. Firstly, we established ethics as the core value underpinning trust, where integrity and trustworthiness emerge as foundational elements within scientific ethos. Secondly, we emphasized trust's *telos* as essential for achieving significant societal objectives rooted in scientific endeavours. Thirdly, the study delved into the ethical quandaries associated with decision-making under uncertain conditions and the central role of trust therein. Utilizing a combination of philosophical literature and policy frameworks, this task endeavoured to holistically understand the ethical foundations underpinning trust in science and investigated guiding ethical paradigms through the perspectives of Research Ethics (RE), Research Integrity (RI), and Responsible Research and Innovation (RRI). These paradigms, sculpted by historical precedents, ethical discussions, organizational undertakings, and societal expectations, converge to foster and fortify trust in the world of science.

Task 3.5 "Governance and regulatory aspects". This task delves deep into trust governance at various levels – from individual researchers to systemic perspectives – and seeks strategies to enhance the trustworthiness of research across these tiers. In view of the complexity and multi-layered nature of the concept of trust in science, the various starting points regarding the governance of science shall be clarified with the help of a simple micro-, meso-, macro-level model, to which the various actors and aspects of trust in science are assigned. Besides mere legal and regulatory aspects, we also included scientific standards and virtues that actors of the science system want to promote. In this way, the heterogeneity of the various science systems and their different governance patterns will also be taken into account.

This task further fosters dialogue and interaction between WP3 and WP6, since a deeper analysis of science governance aspects is a fundamental prerequisite of developing recommendations for key recipients. As part of WP3, this task asks the following questions:

(1) What can science do to be trustworthy / to become more trustworthy?

(2) What legal and regulatory aspects regarding the governance of science need to be addressed for that?

Introduction

Recently, the relationship between scientists and public has been getting renewed attention, where concerns have been raised whether we have problem of trust in science (Cobern et al., 2022) and whether science scepticism is on the rise (Acar, 2023), which has also renewed interest in the question what causes science scepticism (Rutjens et al., 2021). A common sentiment among policymakers is the inherent virtue of public trust in science. However, such a perspective warrants scrutiny. A perceived deficit in trust should not be simply catalogued as another public shortcoming (Wynne, 2006). While distrust *per se* is not the issue, as fostering an attitude of scepticism and criticism is an intrinsic part of the scientific method, it is important to distinguish between warranted and unwarranted (dis)trust.

Current policy recommendations emphasize enhanced research ethics, integrative and participatory scientific processes, and improved science communication. Nevertheless, the efficacy of these solutions hinges on an accurate diagnosis of any underlying crisis of trust.

Studies conducted after the global COVID-19 pandemic have produced conflicting results on whether public trust in science has decreased (Pew Research Center, 2022), increased (Bromme et al., 2022; Fonseca et al., 2023) or become more polarised (Wellcome Global Monitor, 2020).

Currently, we lack a comprehensive analysis of how big trust deficit is and what causes (dis)trust in science. Public polls reveal which social groups are more likely to distrust science and identify various socio-economic and psychological factors that predict higher levels of distrust in science, but they do not manage to measure public trust in science as they do not have a shared concept of trust and distrust.

We believe that the explanation of increasing public distrust in science should start from the philosophical analysis of (dis)trust and trustworthiness of science and scientists. We will ask, what is involved in (dis)trust? What does it truly mean to trust or distrust? We further seek clarity on the entities we place our trust in, from abstract concept of science to tangible ones like institutions, individual researchers, or even specific scientific products. Trust relationships are different dependent on the object of trust. Utilizing a structured micro-, meso- and macro-level model, the intrinsic facets of trust in science must be outlined and associated with relevant actors. Also, it is important to define who is the 'public' in public (dis)trust. Is it the general public or some specific group (Resnik, 2011: 404).

1. Conceptual framework and basic principles of trust and distrust (T3.1)

Task 3.1: *Conceptual framework and basic principles of trust and distrust* (M1–M8) [CEUT, EUR, RU, KIT, ISI, LSE, RCL]

Trust permeates every facet of human existence. Given its ubiquity, a myriad of fields—including philosophy, sociology, psychology, organizational studies, science communication, STS, and political science—have undertaken its study. Sociologists, in their quest, often assess trust levels across diverse societies. Interestingly, while surveys consistently indicate a relatively steadfast and superior public trust in science compared to other institutions, public debates display a growing voice of skepticism. This skepticism is evident in controversies like those surrounding the Covid-19 vaccination, especially in our increasingly divided society. Such skepticism has applied pressure on scientific institutions to bolster the public's trust in their work.

Policymakers advocate for heightened emphasis on research ethics and integrity. They recommend that scientists adopt participatory methods to involve broader sections of society in the scientific process and stress the need for enhanced science communication. However, these strategies will only be effective if they address the root causes of the perceived trust deficit. Presently, a holistic understanding of the factors that influence trust and distrust in science remains elusive. While surveys shed light on demographic segments more inclined towards scepticism and highlight several socio-economic and psychological predictors, they fall short in explicating the intricate mechanics of trust and distrust.

We posit that to truly grasp the rising public scepticism towards science, one must initiate the inquiry from a philosophical examination of trust and distrust. Trust permeates every facet of human existence. Given its ubiquity, a myriad of fields—including philosophy, sociology, psychology, organizational studies, science communication, STS, and political science—have undertaken its study. Sociologists, in their quest, often assess trust levels across diverse societies.

Interestingly, while surveys consistently indicate a relatively steadfast and superior public trust in science compared to other institutions, public debates display a growing voice of skepticism. This skepticism is evident in controversies like those surrounding the Covid-19 vaccination, especially in our increasingly divided society. Such skepticism has applied pressure on scientific institutions to bolster the public's trust in their work.

Policymakers advocate for heightened emphasis on research ethics and integrity. They recommend that scientists adopt participatory methods to involve broader sections of society in the scientific process and stress the need for enhanced science communication. However, these strategies will only be effective if they address the root causes of the perceived trust deficit. Presently, a holistic understanding of the factors that influence trust and distrust in science remains elusive. While surveys shed light on demographic

segments more inclined towards scepticism and highlight several socio-economic and psychological predictors, they fall short in explicating the intricate mechanics of trust and distrust.

We posit that to truly grasp the rising public scepticism towards science, one must initiate the inquiry from a philosophical examination of trust and distrust.

1.1 What is trust?

Trust can be understood by exploring the fundamental nature of human social interactions (Luhmann, 1979; Simpson, 2012). At its core, our social existence is predicated on cooperation. Trust facilitates this cooperation by alleviating the need to constantly oversee or control others, making interactions more straightforward (Luhmann, 1979).

Choosing to trust involves handing over a certain degree of control and responsibility for future outcomes to another. This choice, highlighted by Luhmann in 1998, is accompanied by an inherent sense of risk, given that the outcome may not necessarily be favourable or in line with our desires. Thus, trust acts as a mechanism that aids us in navigating through uncertain scenarios. It involves a conscious decision to take a leap of faith without an assured safety net.

Why do we choose trust despite the uncertainties? Our very nature as social creatures dictates that we rely on others to achieve common objectives. To do so effectively, we must embrace a degree of vulnerability and risk, for without it, achieving our collective goals would be near impossible. The individual placing their trust (the trustor) becomes susceptible to the actions (or inactions) of the one being trusted (the trustee).

To mitigate this vulnerability, a trustor might employ measures like close monitoring or setting boundaries on the trustee's actions. Yet, as McLeod pointed out in 2015, there exists a delicate balance: beyond a certain point, increasing surveillance or constraints could inversely reduce the trust placed in the trustee.

Barber (1983) articulated that trust's foundation lies in shared values, the competency of those entrusted, and the assurance that the entrusted prioritize the collective good over self-interest.

***Value compatibility**:* This concerns expectations about the other's moral and value orientation. People assume that the political, legal and economic systems, in the rules they establish, the procedures they follow and the people involved, more or less reflect our shared basic values.

***Technical competence**:* The expectation that those with responsibility have the necessary skills to carry out their functions properly.

Public or fiduciary responsibility: The expectation that people will acknowledge and be motivated to serve other people's interests and not merely their own.

The philosophical discussion of trust usually starts from distinguishing between reliance and trust. In Annette Baier's seminal article "Trust and distrust" (1986), trust was defined as a kind of reliance, although it is not *mere* reliance. Trust is considered to involve reliance plus some extra factor which usually concerns why the trustor hopes that the trustee will fulfil her expectations. According to Baier, trust is reliance on another's **goodwill**, and this necessarily means being **vulnerable** (Baier, 1986: 234–235). She describes the difference between trust and reliance through a difference in our reactions. Breaches of trust make us feel **betrayed**, whereas if we relied on something to happen (e.g., when we trust that a car will start) and it doesn't, we simply feel **disappointed** or angry. Richard Holton argues that this difference in our reactions shows that trusting a person to do something "involves something like a participant stance towards the person whom one is trusting" (Holton, 1994: 4).
What constitutes the essence of trust? Three foundational components stand out:
1. **Risk and Vulnerability**: This involves willingly stepping into a position of potential harm or loss, marked by a relinquishing of control or power. It is characterized by consciously deciding against certain protective actions, thereby making oneself susceptible.
2. **Affinity and Belief**: Trust is often underpinned by a genuine fondness for, or faith in, the trustee. This may also involve a belief in the trustee's competence and their inherent goodwill.
3. **Anticipation of Commitment**: At its core, trust also encapsulates the expectation that the trustee will honour and act upon their given obligation or commitment.

Both trust and mistrust may be rational or irrational. Perhaps we have no reason to trust, but we trust anyway, irrationally. This kind of trust is unwarranted. Alternatively, there may be reasons to trust, but we refuse them and distrust. This kind of distrust is unwarranted. Irrational trust or irrational mistrust mean that we are not fully cognizant of risks; that we do not have sufficient knowledge of the trustee; that we do not fully take evidence or previous experience into account. Whereas irrational trust ignores the risks that something could betray trust, irrational mistrust overestimates the possibilities of betrayal.

1.2 Trust and trustworthiness

Trust is valuable when placed in trustworthy agents and activities, but damaging or costly when misplaced in untrustworthy agents and activities (O'Neill, 2018: 293). Trust without proper grounds can be as harmful as baseless scepticism. The aim should be to cultivate well-founded trust and justifiable distrust.

Trust is a disposition we hold towards individuals we expect to be trustworthy, with trustworthiness being an inherent quality. In an ideal world, those we trust would prove trustworthy, and individuals who are trustworthy would be granted trust. However, being

trustworthy is not always a prerequisite for justified trust. This is evident in "therapeutic trust" scenarios, when trust is placed in someone in the hope that they will live up to it (Holton, 1994; McGeer, 2008, Hertzberg, 2010). Conversely, even if someone demonstrates unyielding trustworthiness, they might not always be trusted, especially if there are valid reasons to be cautious (Scheman, 2020: 28).

Trustworthiness and trust are conceptually different. Trustworthiness is an evaluation by party A about whether party B merits their trust. Among the various determinants of trust, trustworthiness focuses on the attributes of the one being trusted. Mayer et al.'s influential 1995 model positions trustworthiness as stemming from perceived benevolence, ability, and integrity. Benevolence pertains to the genuine goodwill the trustee holds towards the trustor. In contrast, ability and integrity concern the trustee's competence to execute a given task and their adherence to accepted norms and established standards, respectively. Trustworthiness may be specific (to a certain relationship) or full in extent (one has a disposition to be trustworthy to everyone) (Potter, 2002: 8).

Various motivations underpin our decisions to trust others. Sometimes, trust is a result of deliberate judgment deeming others as trustworthy; other times, it is simply an innate confidence we feel towards them. In certain instances, the act of trusting occurs simultaneously with recognizing someone's trustworthiness. This happens in cases of "therapeutic trust" and "corrective trust." With corrective trust, despite prevailing doubts, one chooses to trust, rejecting what they consider unjust biases (Fricker, 2007; Frost-Arnold, 2014). In these scenarios, the act of trusting is not necessarily a reflection of perceived trustworthiness. Instead, by opting to trust, we position both ourselves and the other party within a mutual dynamic of trust and trustworthiness (Scheman, 2020: 28).

So, for example, disinformation campaigns designed to discredit scientific findings, such as research on the health implications of tobacco or human-induced climate change, don't diminish the genuine trustworthiness of that research.

Onora O'Neill (2002) posits that while regulatory systems can render entities like agencies and corporations trustworthy (ensuring they consistently perform duties, such as ensuring safety), these systems might not garner public trust if the regulatory bodies themselves are viewed with suspicion or if their attempts to establish trustworthiness prove ineffective.

1.3 What does (dis)trust involve?

Philosophical accounts of trust can be divided into doxastic and non-doxastic ones (Keren 2014). Doxastic accounts of trust hold that trusting a person involves holding a belief about the other person, either that this person is trustworthy or that she will do what the trustor expects. Non-doxastic accounts rely on the phenomenology of trust, and state that in trust the trustor holds some affective attitude towards the trusted person.

Many philosophical interpretations of trust either overlook **emotions** or incorporate them only to a limited extent. A prevailing sentiment is that trust is a rational construct, hinged on the belief in the trustee's trustworthiness. Yet, when exploring distrust, a potent emotional component emerges. It's compelling to argue that similar emotional substrates are intertwined with trust as well.

The notion that trust is interwoven with **belief** in the trustworthiness of the trustee (albeit sometimes misguided) might suggest its alignment with rationality. Paradoxically, trust often transcends or even contradicts empirical evidence. This departure from reason underscores the affective component of trust—a dimension that doesn't always adhere to logic. Analogously, one's fear of a harmless dog or mistrust of a seemingly loyal individual, despite evidence testifying to their benign nature, underscores the emotional crux of trust. Some argue for confining trust within rational parameters, yet such restrictive views fall short of capturing its holistic nature. Trust and mistrust, in our perspective, oscillate between **rationality and irrationality**, demanding an acknowledgment of the emotional undercurrents at play. This is why trust is laborious to cultivate, but precariously easy to shatter.

While numerous philosophical treatises on trust either sidestep or minimally address emotions, they often encapsulate trust through the prism of Strawsonian reactive negative attitudes (Strawson, 1962) - feelings of betrayal, resentment, or anger directed at the trust's focal point. In this context, if trust is anchored in normative expectations and these expectations are unmet—resulting in a betrayal of trust—then it is justified for the trustor to exhibit negative emotional reactions, such as resentment or anger, towards the trustee (Jones, 2004; Holton, 1994; Hawley, 2019). Faulkner (2007: 881) terms this emotion-tinted interdependence as "affective trust."

However, there's a burgeoning recognition of trust's intrinsic emotional aspects. We believe that beyond possible reactive responses, there's an inherent emotional component in every manifestation of trust. Trusting inherently involves a favourable disposition towards the trusted party, a kind of assurance that prompts us to be vulnerable and relinquish control. Karen Jones (1996; 2019) advocates for the presence of a positive affective disposition toward the trustee. In case of distrust, the anticipatory attitude is one of anxiety (Baier 1986: 234), fear (Jones 2019: 956), frustration, or even anger.

Moreover, Bernd Lahno (2001), drawing upon Ronald de Sousa's cognitive theory of emotions, underscores trust's emotive nature, emphasizing the causal interplay between trust and belief. Within this framework, emotions are not mere fleeting feelings but critical judgments shaping our world perception.

Bernd Lahno's insightful reflections offer clarity on this point. He notes: "In describing trust as an emotional attitude, what's intended is its intricate association with our perception of the world. Trust intrinsically impacts our cognitive patterns—how we view the world and associate our thoughts. Although there's an interplay between trust and belief, trust, in influencing our cognition, doesn't merely spring from rational deliberations." (Lahno, 2001: 177)

In the realm of trust, sometimes our judgments rest on sheer emotion, devoid of prior analysis. At other junctures, trust emerges from reasoned beliefs about the trustee. Our expectations, thus, can be sculpted by belief, emotion, or a blend of the two, painting a mosaic that is both intricate and intriguing.

Recognizing the affective underpinnings of trust is not a mere academic endeavour; it is an avenue to comprehend and foster genuine trust. Let us consider the mounting scepticism towards science, as illuminated by Furman (2020). Trust, as we discern, is not solely hinged on the perceived reliability of the trustee. The emotional realm of the trustor—shaped by context, past experiences, or overarching views of societal institutions—plays a significant role. Hence, appreciating the emotional texture of trust offers insights into the episodic and seemingly irrational distrust in science, and possibly suggests strategies for mitigation.

1.4 Trust and distrust

Philosophical explorations of trust are abundant, yet explicit forays into the realm of distrust remain scant. Katherine Hawley's assertion that to truly comprehend trust one must simultaneously delve into distrust highlights this oversight (Hawley, 2014). She notes the general tendency to erroneously equate distrust merely as the negation of trust.

Contrary to the idea of them being opposites, trust and distrust are not entirely exhaustive of each other. It is widely recognized that they stand as contraries rather than contradictories (Govier, 1992b, Jones, 1996, Hardin, 2001, Ullmann-Margalit, 2004). One can't trust and distrust someone in the same regard simultaneously.

Figure 1. Gradual nature of trust.

Both trust and distrust have a gradual nature. Mere lack of trust doesn't automatically translate to active distrust; it could mean a neutral stance of neither trust nor distrust (Ullmann-Margalit, 2004). There is a scale, on the one end of which is full trust and on the other, complete distrust. Full trust is confidence in someone, relying on the trustee without doubting them; confidence that the trustee will fulfil our expectations, without any need for control or precautionary mechanisms. At the other end of the scale from trust is distrust, where there is no trust at all, and there is a concomitant desire for control.

Absence of trust should not be equated with distrust. Rather, absence of trust describes the centre stage of this continuum between trust on the one hand and distrust on the other hand, where a trustor neither trusts nor distrusts a trusting object. Distrust/absence of trust can be described as an attitude of healthy scepticism.

Domenicucci and Holton (2017) posit that distrust is not merely an absence of trust but an active sentiment of viewing someone in negative light, often coloured by feelings of insecurity, cynicism, or even fear. This view distinguishes distrust from a neutral stance. D'Cruz (2020) further elaborates on the multifaceted nature of distrust, suggesting that its basis, be it rooted in scepticism of competence or quality of will, affects its scope and intensity.

Moreover, distrust, once made evident, holds significant potential for harm by signalling perceived incompetence, malevolence, or lack of integrity in the distrusted party (D'Cruz, 2020). There's also the alarming propensity for distrust to be generalized based on isolated instances, often leading to biases and stereotypes (D'Cruz, 2020). This susceptibility, coupled with its tendencies for self-perpetuation, necessitates reflection on our own distrustful attitudes. In summarizing, distrust is complex and multifaceted, and cannot simply be seen as the antithesis of trust.

Trust is valuable when placed in trustworthy agents and activities, but damaging or costly when misplaced in untrustworthy agents and activities (O'Neill 2018: 293). Unwarranted trust can be as dangerous as unwarranted distrust. Warranted trust means that it is rational: it is based on good evidence, it successfully targets a trustworthy person or institution. Warranted trust means that it is plausible in a particular situation. The conditions for warranted distrust, as outlined by Govier (1992), emphasize acts of deception, dishonesty, manipulation, and harm. What we need is warranted trust and warranted distrust. Warranted trust in scientific research can be strengthened by fostering participation in research as a co-creative and inclusive process, sensitive to societal values, concerns, and needs.

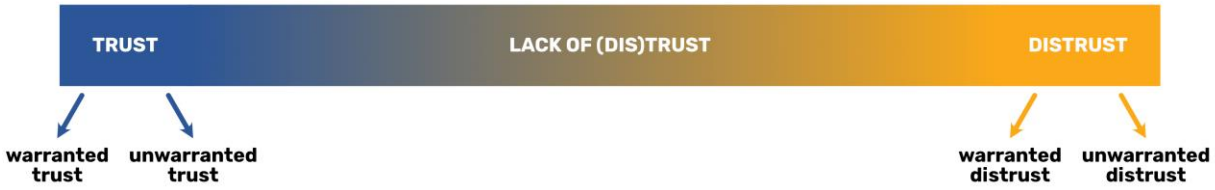

Figure 2. Warranted trust/distrust

1.5 What kind of relationship trust is?

Many philosophers seem to lean towards the doxastic view of trust based on the belief that trust functions as a three-place relationship (A trusts B to C; A trusts B to be C). Yet, there's substantial reasoning to consider that a two-part structure (A trusts B) or even a singular form (A's trust) may be more foundational.

Domenicucci and Holton (2017: 154) have recently posited that a dual relationship should serve as the foundation. Taking it a step further, Paul Faulkner asserts that to truly understand the inherent value of trust, we must begin with "X having a trusting attitude; meaning, in X trusting Y, or simply in X's act of trusting" (Faulkner, 2017: 118–119). We agree with Faulker (2017) that the one-place trust is the most basic form of trust that is involved also in two-place and three-place trust.

We should distinguish between three different forms of trust: 1) trust; 2) trusting some other person (A trusts B); 3) trusting somebody or something to do something (A trusts B to do C or A trusts B in the role C). The three forms of trust can be explained as follows:

1) One-place trust (A's trust) involves a feeling of confidence, readiness to give up control by refraining from taking precautionary measures;

2) Two-place trust (A trusts B) entails a feeling of confidence and a positive feeling towards the object of trust, as well as possibly a belief that the trustee is trustworthy, at least in some domains.

3) Three-place trust (A trusts B to do/to be C) involves, besides a feeling of confidence and positive attitude towards the trustee, also an expectation that C will be able and motivated to fulfil the expectations of A. It is an attitude of **hope** which involves both the **positive feeling** towards the object of trust and **a belief** that the desired outcome is possible or at least not impossible.

Three-place trust (A trusts B to be/to do C) may take the form of either a practical or a moral relationship. In a **practical relationship of trust**, A has a positive attitude towards B as well as relies on B to do something, arranging her life around the predictive expectation that B will do C. In a **moral relationship of trust**, A trusts B and relies on B to fulfil C as a **moral commitment** which she expects her to fulfil. This entails that A believes that B is capable and morally motivated to do C.

Katherine Hawley argues: "To trust someone to do something is to believe that she has a commitment to doing it, and to rely upon her to meet that commitment. To distrust someone to do something is to believe that she has a commitment to doing it, and yet not rely upon her to meet that commitment." (Hawley 2014:10). She explains that "commitments can be implicit or explicit, weighty or trivial, conferred by roles and external circumstances, default or acquired, welcome or unwelcome. In particular I will take it that mutual expectation and convention give rise to commitment unless we take steps to disown these" (p. 10-11). In order to be trustworthy, Hawley says, one must live up to one's commitments; correspondingly, attitudes of trust will involve "expectations of commitment-fulfilment." (Hawley, 2019: 23)

There is an important difference between practical and moral relationships of trust. When one relies on the watch to show the right time, the watch does not know anything about our expectations and cannot respond accordingly. In the same way Immanuel Kant did not know that his neighbours relied on his regular walks to notice that it is lunchtime. The

neighbours could not feel betrayed if Kant decides one day to go out at a different point of time. But if we were to ask someone to tell us the right time, we would be approaching her with normative expectations – to be honest, to be kind and to be accurate. Such expectations stem from our common morality. In a moral relationship one treats another person as a moral person (which in Kantian terms means to treat another person as an end and not as a means), expecting her to fulfil a commitment or an obligation that has been mutually agreed upon (either directly or because as members of a moral community we have committed ourselves to the same moral norms).

Besides common morality, normative expectations can also stem from the shared norms we relate to certain roles we fulfil, either in our personal relationships (being a mother, daughter or a wife is related to certain obligations) or in our professional life (as an academic working in a university we have roles as a researcher, teacher, supervisor, or leader, all of which are associated with certain commitments). The commitments of professionals are often articulated in professional codes of ethics, which inform the public about what can be expected from a representative of a given profession. Representatives of all professions must fulfil roles they have adopted and perform according to standards, as well as in view of certain values.

By trusting the bus driver, the cook, the teacher or a physician, we know what kind of values and standards the representatives of these professions are committed to follow. Also, the scientists must follow certain values and standards. Being trustworthy means to adhere to the commitments related to one's role or profession.

1.6 Trust in science

We can distinguish between **trust/distrust within science (between scientists)** as well as **public trust in science**. One should distinguish between **epistemic and social function of science**.

Trust within science refers to the reliance researchers place on their peers, especially concerning their claims and statements. Today's scientific advancements largely emerge from collaborative efforts, primarily because collecting and analyzing empirical data often surpasses the resources and time any single scientist can invest (Hardwig, 1991). Thus, researchers depend on the accuracy, integrity, and self-awareness of their colleagues in evaluating evidence. Knowledge, being a communal asset, means that as researchers venture into new realms of discovery, they build upon the work of others. Even though there is a practice of reproducing or replicating experimental and observational results, not all findings undergo this scrutiny. This is partly because exhaustive replication is resource-intensive, and the scientific community tends to value innovative findings over repeated ones (Kitcher, 1992). Furthermore, trust plays a pivotal role in the scientific communities responsible for upholding research quality via training, peer evaluations, and constructive criticism (Longino, 1990).

However, the inherent nature of science entails institutionalized skepticism, suggesting that trust within the scientific community is not an automatic guarantee, but something to be continually earned and validated.

Kristina Rolin has delved into exploring the foundations of rational epistemic trust within the scientific domain. She posits that "trust becomes epistemic when it furnishes epistemic justification for one's beliefs. Such epistemic trust is deemed rational when it is grounded in appropriate and sufficient evidence." (Rolin, 2020: 354)

Public trust in science can mean several things: trust in what scientists say (epistemic trust) and what they do – trust in a scientific method (reproducibility and replication of their work), in research findings, in individual scientists, in research institutions, in products of science and innovation, in science as a system.
Thus, we should distinguish between

- **public trust in science** (or specific disciplines, research fields, or techno-scientific areas);
- **public trust in scientists** (competence, goodwill/benevolence, integrity) or that their claims have truth-value (epistemic);
- **public trust in products of science and innovation**;
- **public trust in science as a system** (consisting of different stakeholders).

Sociologists are eager to measure the level of trust in different societies. What these surveys show is, that trust in science (either in scientists or institutions) is not independent of trust in **other social institutions**. Surveys show that public trust in different social institutions usually fluctuates together over time and on individual level trust in one institution (or social institutions in general) is a good indicator of trust in another. Also, in different countries public trust in science depends on different aspects (incl. science communication) but not only on it.

1.6.1 What are the roots of distrust in science?

The academic landscape about trust in science is far from being unanimous. The longstanding cognitive deficit model held the belief that people's scepticism towards science stemmed from a lack of scientific knowledge (Wynne, 1991; Simis et al., 2016). In this model, bolstering trust is achieved through better education and information dissemination. However, this perspective faced criticism in the last decades (see e.g. Jasanof, 2011: 362), as empirical studies revealed that simply amplifying information might not necessarily mend trust. Notably, Rutjens et al. (2018: 384) illuminated that acceptance and rejection of scientific data can vary based on the subject under examination and pinpointed religion, political views, morality, and science comprehension as key indicators of trust. Rowland et al (2022: 1399) highlighted the role of communication, emphasizing that the credibility of scientific information is dependent on its source, its prevalence in public discourse, and the nation's institutional trust levels.

To bolster trust in science, two predominant strategies emerge. The first emphasizes fostering a culture of research integrity and transparency in scientific research (Bouter,

2023; Ciubotariu & Bosch, 2022). This includes promoting open science initiatives (Haven, et al., 2022; Song et al., 2022, Rosman et al., 2022). The second underscores the importance of enhancing scientific communication (e.g., Intemann, 2023) and increasing participatory aspect of research (e.g., IKEA phenomenon mentioned by Acar, 2023) to fight growing science scepticism and spreading misinformation in the new digital environment. However, trust is multifaceted, implying that even the most ethical and transparent scientific practices do not guarantee trust. It is also essential to recognize instances where distrust is valid.

1.7 Take home messages

While many philosophers have primarily centred their attention on trust, our initial exploration of distrust provided a broader understanding of trust. This exploration emphasized the significance of emotions and how an individual's inclination to trust, which involves relinquishing control and exposing oneself to vulnerability, underpins all types of trust. By focusing on distrust as affective state, we put forth a new perspective: trust in science is not solely based on the credibility of the entity being trusted (whether a scientist or a scientific institution) and the education of the trustor, but also on the personal characteristics, past experiences, and values of the individual placing their trust.

While the cognitive deficit model suggested that distrust in science stemmed from a lack of knowledge, as noted by Wynne (1991) and Simis et al. (2016), it is evident now that enhancing trust isn't merely about better educating the public. The reasons for distrusting scientists and accepting or rejecting scientific findings are diverse. They are intertwined with psychological elements like readiness to lay oneself open, sociological aspects such as trust in other social institutions, and personal factors including personality traits, past experiences, moral values, religious beliefs, and political orientations. This sheds light on why, even when scientific institutions and scientists are credible, they can still face distrust.

However, even if trust in science depends on so many factors, it does not mean that we should not make efforts to make scientists, scientific institutions, and science as a system more trustworthy. Initiatives to make science more trustworthy have been described here as research ethics, research integrity and responsible research and innovation.

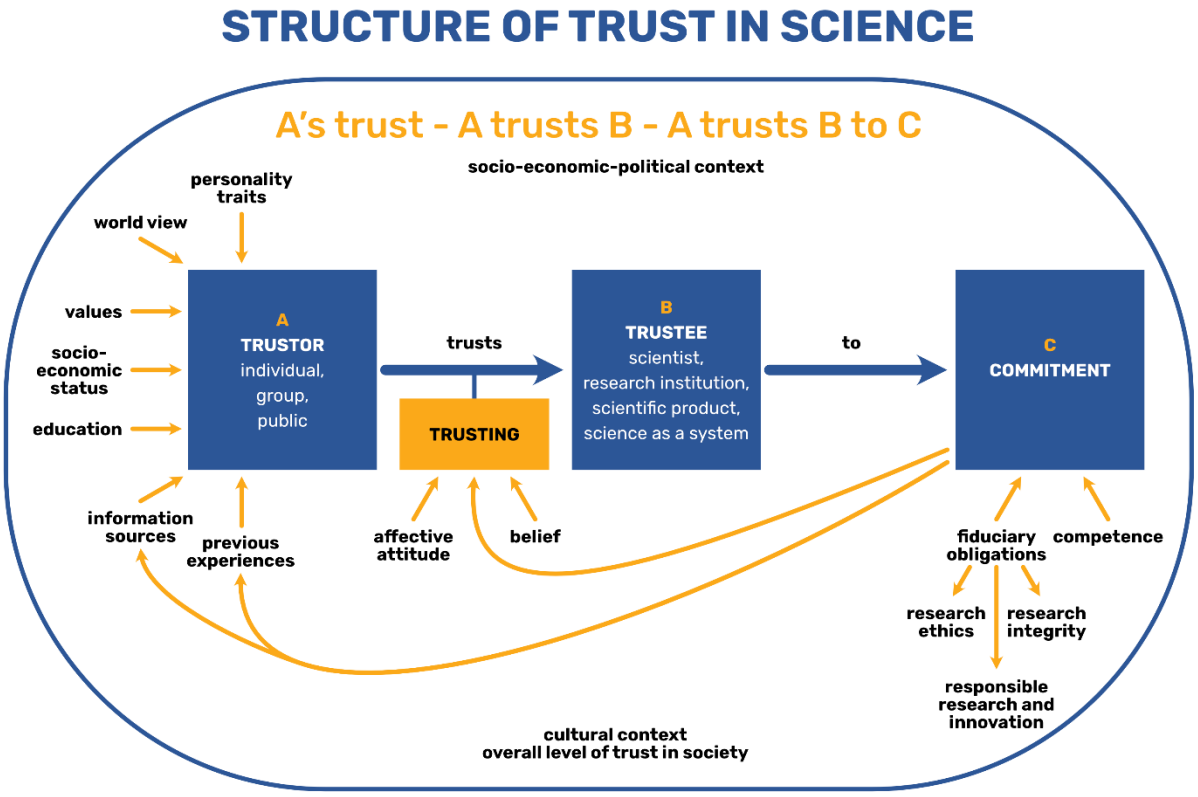

Figure 3. *Structure of trust in science*

2. Contextualizing the Framework (T3.2)

Task 3.2: *Contextualising the framework* (M4–M12) [LSE, CEUT, ISI, EUR]

2.1 A systems approach

A key conclusion of the five tasks in WP2 is that the behaviour of scientists and organizational characteristics are identified as a part of the trust equation, in other words trust and distrust arises in the context of a multi-level system. The way people act is to a great extent shaped by broader cultural frameworks of societal or institutional norms and values. Following Coleman the interaction between three levels is proposed. Type I (macro-micro) is a contextual mechanism that explains how macro-level conditions (e.g., culture or economic forces) affect the behaviour of individuals. Type II (micro) is an individual mechanism that evokes beliefs, attitudes, desires and opportunities as antecedents of action. Type III (micro-macro) describes how isolated individual actions generate collective outcomes –intended or not. A strong implication is that, if the aim is to change attitudes, behaviours and norms, it is more effective and efficient to target the normative environment than to focus on the individual level.

System resilience

Resilience is a property or a characteristic of a system. A resilient system is an adaptive system that readily assimilates incremental changes in the external environment, but crucially has established transparent procedures to address major external challenges. An example of a resilient system is that of medical ethics which has seen dramatic developments over the last fifty years due to the rapid pace of innovation in medical technologies. In the wider field of technological innovation European society has yet to establish widely accepted and resilient procedures to weigh up the pros and cons and to make recommendations to the research community and other key stakeholders in innovation management.

WP2 identifies a number of actions for the actors in the research eco-system that can contribute to system resilience.

- Increasing transparency and accessibility of scientific data
- Measures to safeguard transparency and academic independence
- Introducing strategies to prevent questionable research practices rather than adopting defensive reputation management
- Discourage the making of extravagant claims in research proposals and from preliminary findings
- Introduce institutional policies on conflicts of interest
- Recognize the risks of being used by policy makers to justify their decisions
- Treat science communication as both dialogical and participatory
- Avoid treating trust as an exercise in public relations
- View a goal of science communication as that of enabling participation
- Be open to learning lessons from public concerns and distrust

- Acknowledge that while science accepts scepticism, it should not be seen as a problem in the public
- Be aware of the risk of incentives leading to perverse outcomes (Goodhart's rule)
- Establish rules for ethically credible partnerships with commercial companies

Public opinion on science, scientists and innovation

The Wellcome Global Monitor 2020 finds that overall, science and the veracity of science are trusted by more than 8 in 10 Europeans. College and University scientists are viewed as having a greater commitment to the public good – fiduciary responsibility - than scientists working in companies. In parallel, transparency of funding is perceived to be less evident in company than in college/university scientists, raising questions of conflicts of interest.

2.2 Contemporary challenges: The Neo-liberal University

Neo-liberalism is defined as "a complex, often incoherent, unstable and even contradictory set of practices that are organized around a certain imagination of the 'market' that emerges with an economic and political dynamic to reform higher education". Universities and research centers adopt strategies for competitive advantage". Performance of universities, departments and individual researchers is audited with a variety of metrics that feed league tables. Research quality and impact are converted into funding levels. Productivity and entrepreneurial flair are rewarded, job insecurity is rife and those who fail are discarded. An unintended consequence is that some disciplines in the medical, life and social sciences face a replication crisis, the withdrawal of academic articles is not infrequent and, by researchers own admission, questionable research practices are not uncommon. A further outcome of this increasingly competitive system is that researchers tend not to trust their colleagues. A problem with the metrification of research outputs is captured in Goodhart's law. This states that those who are aware of a system of rewards and punishments will change their behaviour to achieve their desired results. As such, metrics of scientific evaluation will be abused by such activities as salami slicing of research and citation circles.

2.3 Questions about Research Integrity

The International Research Integrity Survey (IRIS), funded by the European Commission sought the opinions and reported behaviours of 2,300 active researchers in the US and 45,000 in Europe and asked them to assess their institutions' support for research integrity. The findings of the survey point to deficits in research environments, research activities and organisational practices. At the same time, there are encouraging signs of concerns about research integrity, and a willingness to adopt research and organisational practices that support RI. (Allum, N. et al., 2023).

With the experience of thirty years of RI policy, 80% of US researcher are confident about the integrity of their work; the same is the case for 60% in Europe, falling to 40% for Europe's early career researchers. Confidence in their organisations' support for RI is acknowledged by fewer than 50% in the US and less than a 30% in Europe.

Admissions of QRPs ranging from superficial peer reviewing, selective reporting and researching without ethical approval, are not uncommon and are more frequent in Europe than the US. In both cases, the rate of admission is twice as high for European researchers. In some instances, this may be indicative of sloppier practice in Europe, but it may also reflect variability in the formal requirements for ethical approval and declarations of interest in different national and institutional contexts.

Overall, half of researchers say their organization does not have high standards. There are, however, positive indications regarding the development of individual competences and organizational policies and practices. Researchers value the intrinsic benefits of RI, in terms of delivering greater research quality and trustworthiness far more than the extrinsic benefits of promotion and salary. Most researchers recognize that responsibility for RI is shared between them and their organization. In Europe a majority say they are willing to engage in RI training, consider that RI policies are well intentioned and should improve research quality. A non-trivial percentage of European and American researchers would welcome additional support across the nine RI topics.

2.4 Right wing populism

Rooduijn of the Populist research groups reports that about one-third of Europeans now vote for far-right and far-left parties with wide support for anti-establishment policies, challenging the mainstream. He reports that "the underlying trend is the numbers keep rising. Mainstream parties are losing votes, anti-establishment parties are gaining. Right and left wing populism say they offer a democratic corrective, privileging the ordinary person against elites, vested interests and an entrenched establishment". Issues such a climate change and global warming, and the covid-19 pandemic saw the gap between the politics and science closing leading to two possible outcomes – scientists categorized as part of the hated expertocracy, or distrust in the political elites generalizing to scientists.

The reach and volume of those who oppose science has been accentuated by social media platforms with a number of opinion formers with millions of followers articulating political conspiracy theories and anti-science rhetoric.

2.5 Scientific uncertainty

There is a striking similarity across a number of 'breakthrough' innovations in the life sciences. Consider recombinant DNA (genetic engineering), xenotransplantation (pig hearts for human transplants), cloning of animals (dolly the sheep) and genome editing (CRISPA cas9). All of these strategic developments have led to calls from sections of the scientific community for a moratorium while potential risks, both scientific and to the 'ordre publique' are investigated.

Calls for moratoria for novel scientific development start with rDNA in 1974 and the 'Berg letter' to Science and Nature inviting scientists around the world to join the signatories in a voluntary moratorium until potential hazards to laboratory staff and the public had been assessed. In 1999 a moratorium on xenotransplantation was proposed over the risks of porcine retroviruses. Four years later with the birth of 'dolly the sheep', and popular

images of cloning dreadful people, human cloning is banned around the world. Finally, in 2015, Berg, Charpentier and colleagues call for an international moratorium on all clinical uses of human germline editing to make genetically modified children.

Berg wrote in 2007 in a reflection the moratorium on rDNA "Scientists around the world hotly debated the wisdom of our call for caution, and the press had a field day conjuring up fantastical 'what if' scenarios. Meanwhile, the public seemed comforted by the fact that the freeze had been proposed by the very people who had helped to develop the technology". Berg's 'conjuring up of fantastical scenarios' is one of the key elements stimulating symbolic coping as theorised by Wagner and Kronberger. Symbolic coping is the process by which the novel innovation, comes to be understood in the public domain. Transgenics found resonance with the myths of Faust and Frankenstein.

Will history repeat itself with genome editing? Senior scientists called for a moratorium on heritable genome editing; little has been achieved thus far. With developments such as embryonic stem cell modes and concerns about germline interventions leading to eugenics and/or human enhancement, the stage is set for further symbolic coping that is unlikely to be led by scientists.

2.6 Measures to support trust in science and trustworthy science

The International Research Integrity Survey as part of the EC funded study SOPs4RI showed that there are concerns about the current state of health of integrity in research conduct and no magic bullets, no quick fixes. After decades promoting ethics and integrity in research conduct, there are integrity deficits in both researchers' behaviours and in how they judge their organization's practices. This leads to a cornerstone in the framing of RI; it is the responsibility of multiple actors – researchers, their organizations, research funders, professional bodies and academies. A particular responsibility for research performing organizations is the provision of training in ethics and research conduct for all involved in the research cycle including students, researchers and administrators.

The EC, the study "Standard Operating Procedures for Research Integrity" (SOPs4RI) collected examples of good practice and made recommendations for Research Producing Organisation to implement policies for RI. (Mejlgaard, N. et al, 2020). With pertinent examples the areas for development are the research environment; supervision and mentoring; research integrity training; research ethics; data management, research collaboration, and procedures to deal with integrity breaches.

Responsible Research and Innovation

The principle of Responsible Research and Innovation (RRI), introduced in 2010, was built on 25 years of experience with science and innovation funded by European institutions

The European Commission's White Paper on Governance and the Science and Society Action Plan emphasised openness, responsibility and the inclusion of stakeholders and the public. The explicit link to European fundamental principles is a new feature in RRI.

The most prominent way of transferring the ideas of RRI into practice is the mutual mobilisation and learning (MML). It aims at enhancing the interaction between scientists, stakeholders and the public, mobilising them in questions thrown up by a particular technology or a societal problem with a strong technology link. RRI aims to help innovators pick amongst options those that are more likely to be socially acceptable. How this can be achieved in practice is still an open question.

Concluding remarks

Science is a complex multi-actor and multi-level institution. Notwithstanding questions about aspects of research integrity reflected in the behaviour of scientists and the absence of commitment in their organization, public trust in science and scientist is in evidence. Contemporary challenges to the perceived trustworthiness of science include right wing populist political ideology and some predictable but unintended consequences of the neo-liberal shift movement for public accountability, auditing and funding of universities.

It is apparent that there are no magic bullets, no quick fixes, leading to a cornerstone in the retention of trust in research and innovation. It is the responsibility of multiple actors – researchers, their organizations, research funders, professional bodies, academies. And in the discharge their duties ensuring that the public voice plays a role in choosing the direction of the translation of science into technologies.

2.7 Take home messages

How can trust in science (in individual scientists, in institutions, in science as a system, in scientific outcomes) be strengthened? What are the mechanisms to build trust in these different areas?

Is there a crisis of public trust in science? Overall, science and the veracity of science are trusted by more than 8 in 10 Europeans. College and University scientists are viewed as having a greater commitment to the public good – fiduciary responsibility - than scientists working in companies. In parallel, transparency of funding is perceived to be less evident in company than in college/university scientists, raising questions of conflicts of interest. According to the Wellcome Global Monitor 2020, only 7% in EU27 say they have 'not much' or 'not at all' trust in science and 10% say 'not much' or 'not at all' trust in scientists to find out accurate information about the world. However, 41% in EU 27 say they have 'not much' or 'not at all' trust in the national government in this country.

Is there a crisis of research integrity? After decades promoting ethics and integrity in research conduct, a survey of over 40,000 active researchers finds integrity deficits in both researchers' behaviours and in how they judge their organization's practices.

The unintended consequences of the neo-liberal university: In the neo-liberal university market economy the performance of universities, departments and individual researchers is audited with a variety of metrics that feed league tables. Research quality and impact are converted into funding levels. Productivity and entrepreneurial flair are

rewarded, job insecurity is rife and those who fail are discarded. Has this contributed to the replication crisis, the withdrawal of academic articles is not infrequent and, by researchers own admission, questionable research practices are not uncommon.

Policy recommendations

There are no panaceas: The way people act is to a great extent shaped by broader cultural frameworks of societal or institutional norms and values. 'Coleman's boat' provides an apt illustration of these interactions between macro-level phenomena such as organisational culture and practices and micro-level phenomena such as individual attitudes and behaviour. If the aim is to change attitudes and behaviours, it is more effective and efficient to target the normative environment than to focus on the individual level.

Promoting research integrity: Research integrity is the responsibility of multiple actors – researchers, their organizations, research funders, professional bodies and academies. A particular responsibility for research performing organizations is the provision of training in ethics and research conduct for all involved in the research cycle including students, researchers and administrators. The pro-active commitment of the NSF in the US shows that funding organizations have a vital role in setting the agenda regarding research integrity and for monitoring compliance.

Ethical partnerships: add criteria to support 'ethically credible partnerships' between university and commercial sector research partnerships. ECPs added to protocols for the Declaration of Interests would focus on public transparency of the partnership agreement, the maintenance of academic standards and research independence.

The honest issue advocate: The role of scientific evidence for policymaking is becoming crucial yet contested. The issues addressed at the science-governance interfaces are characterised by complexity, uncertainty, scientific dissent and disputes on values, yet decisions are often urgent and stakes high. The honest issue advocate would adopt a commitment to both scientific rigor and the defence of democratic principles.

Beware of metrics: when performance metrics are linked to individual and/or institutional systems of rewards changes are likely to follow to achieve desirable results. The changes may well be dysfunctional.

3. Trust and truth (T3.3)

Task 3.3: *Methodology and approach: participatory dimension of trust* (M4–M12) [EUR, AUTH, LSE, CEUT]

3.1 Philosophical backdrop: truth and its vicissitudes

It is worthwhile to explore how the challenge of growing distrust in science may be linked to the current crisis afflicting one of the core signifiers of Western thinking, namely truth (ἀλήθεια). Truth is philosophy's primal word as it were, and in the so-called post-truth era, we are facing a crisis of truth. The fact that the *Oxford Dictionary* elected "post-truth" as "word of the year for 2016" seems symptomatic of this predicament. A chronic disparity, rather than correspondence, has arisen between what we claim to know (knowledge) and what is happening out there (reality).

The tension as such has always been there of course, as a stimulus for reflection and research. Plato already distinguished ἐπιστήμη (validated knowledge) from δόξα (opinion). And when Jesus claimed that He had come into this world to testify the truth (John 18:37), Pilate (a Roman scholar, well-versed in philosophy) famously retorted "What is truth?" (Τί ἐστιν ἀλήθεια; John 18:38), thereby entering into a dialogue which pointed to a disparity between the truth of this world and the truth of faith, – a disparity which medieval scholasticism aspired to supersede.

This disparity has also been present in modern science. Philosopher of science Gaston Bachelard for instance distinguished technoscientific validity (*fabricated*, literally, in laboratories) from conceptions and experiences emerging in the lifeworld, not yet affected by the cathartic operations of technoscience.

According to Heidegger, writing in the 1930s, however, the crisis of truth radically aggravated, giving rise to what he referred to as a "collapse" of truth ("Einsturz der ἀλήθεια", 2014, p. 224; Cf. Zwart 2022). For Heidegger, this event was closely related to the radical instrumentalization and mobilisation of knowledge by what he referred to as the wave of meta-politics during the 1930s (Nazism, Americanism, Communism): the industrialisation of research and the mobilisation of brain power by state power, which he not only witnessed in Germany, but also in other countries. This trend was exemplified by the Manhattan Project in the New Mexican desert, but also advocated by Vannevar Bush, in his report *Science: the Endless Frontier*, written on behalf of the National Science Foundation (NSF). The creation of NSF as a federal agency was proposed in 1945 by Vannevar Bush as head of the *Office of Scientific Research*, a governmental organisation, which managed several big science projects (including the Manhattan Project).[1] Bush presented his proposal in his famous report to the President of the United States, calling for an expansion of government support for post-war scientific research. The report advocated a "big science" approach (Rasmussen 2002) to overcome the divide

[1] https://www.nsf.gov/od/lpa/nsf50/vbush1945.htm

between "basic" and "applied" research.[2] The purpose of science, Bush argued, is to deliver useful scientific knowledge to society, although to achieve this, science should be granted sufficient freedom of research. His report entails a top-down big science initiative, as is exemplified by its military terminology, e.g. "frontier", "mobilisation", "war against disease", "national security", "targets", etc. Recently, the NSF presented a similar call in the arena of science politics, by presenting a *Convergence Accelerator* to promote use-inspired convergence research via "deep" interdisciplinary collaboration and partnerships between academic and non-academic stakeholders. NSF defined convergence research as application-oriented, but fed by basic discovery-oriented research, adopting a convergent approach by integrated teams including industry, academics, non-profit organisations, government entities, and others. Convergence initiatives should be proactively and intentionally managed, should include multiple societal stakeholders and should involve intensive education and mentorship. In March 15, 2019, the NSF Convergence Accelerator was presented via a "Dear Colleague" letter published on the NSF website and addressing the U.S. research community at large. It was indicated that the NSF convergence initiative would focus on two tracks, namely: (1) Harnessing the Data Revolution and (2) Future of Work at the Human-Technology Frontier. The political context has drastically changed, of course, including the context of global neo-liberalism (as a contemporary version of meta-politics), but the question raised by Heidegger is still relevant: what are the implications for truth in an era of big science politics, linking research with strategic ambitions? Could there be a link between the science-politics entanglement and the current post-truth era?

3.2 Top-down versus bottom-up science: tensions

When science is directed by major granting agencies, coordinated institutions, and international research programs, the Overton window of appropriate questions to query narrows. The notion that one can trust or distrust science as a whole, is relevantly recent, with a coordinated notion of science as a collective research project transcending national borders in the nineteenth and twentieth centuries. Since the Enlightenment, science as a universal and universalizing project of bringing scientific knowledge together has entailed both notions of developing expertise while at the same time democratizing inquiry. The image of the lay scientist, coming from the older image of the natural philosopher, is available to anyone in theory.

That there are no intrinsic barriers to becoming an amateur scientist or intellectual, and such an educated polity provides necessary checks and balances against groupthink. This decentralization of inquiry from the clergy or government allows in theory a proliferation of research programs. However, the loss of coordination, and the possibility that a research program may develop out of control of political power has historically posed a danger. Thus, the advantages of opening up the firehose of potentially deputized scientists can be clarified, including the educational fundament it requires. In *The Panda's Thumb: More Reflections in Natural History*, co-architect of the Evolutionary Synthesis Stephen

[2] https://www.srainternational.org/blogs/srai-news/2020/07/09/sciences-75-year-pursuit-of-the-endless-frontier

Jay Gould (1992, p. 126), puts it thusly: "I am, somehow, less interested in the weight and convolutions of Einstein's brain than in the near certainty that people of equal talent have lived and died in cotton fields and sweatshops."

At the same time, the participatory dimension of trust in science has been made more difficult by ever-increasingly specialized technical knowledge. The knowledge production of our contemporary age mobilizes huge teams of scientists, in scenarios where few people actually understand fully what is going on at each step of the process (Princen 2005). Technoscience driven by large governmental and/or corporate labs linked together often seeks very specific expertise, requiring technical training to standardize techniques (such as pipetting) which bring us closer to possibilities of reproducibility (Knorr-Cetina 1981, 1999).

In many institutions, there is a palpable tension between official research programs, and the native interests of scientists. The questions that scientists pursue, or wish to, may implicitly challenge the frameworks and paradigm of top-down research programs (Kozlov 2023). These discrepancies may lead to agonistic, synergistic research programs, but they can also equally lead to antagonistic ones working in different epistemic paradigms as well. Thus, diverse research questions, which often also employ divergent methods, may lead to different epistemic cultures.

3.3 The post-truth era

The post-truth era is characterised by a pandemic of disregard for truth. The global public environment has become a data-sphere, where terabytes of data are relentlessly circulating. The data deluge threatens to make the concept of scientific truth obsolete, to the extent that anything can be verified by adapting algorithms to desired outcomes. This results in an erosion of the credibility of technoscientific expertise, relentlessly interpellated by public discontent.

Although many academics currently deplore this disregard for truth in the post-truth era, up to the point of launching marches for science (marches on behalf of scientific truth), they must be aware of their own involvement, to prevent becoming entrapped in the position of the Beautiful Soul (bemoaning the current crisis while overlooking how they themselves are deeply *involved* in what they deplore). To some extent, the erosion of truth has been actively promoted by academic scholars themselves. Researchers in the field of Science and Technology Studies (STS) could be mentioned here, for instance, intent on exposing how scientific facts are socially constructed. According to hard-core STS, scientific truth is determined by experts in the context of power games. Truth is the outcome of social processes and political negotiations, it is what certain self-serving coalitions of experts temporarily present as truth. For STS, the adoption of post-truth politics by politicians like Donald Trump evidently became a source of embarrassment, of trauma even (Zwart 2022). In 2004, Bruno Latour (one of the founding-fathers of STS) already criticised his own field for spreading the message that scientific facts are to be distrusted and that there is no such thing as truth. In retrospect, Latour deplored how right-wing "extremists", by questioning expert views concerning climate change for

instance, appropriated STS strategies, so that their conspiracy theories seemed uncannily similar (in terms of argumentative structure) to former STS ideas. Latour now considered it a mistake that he had moved *away* from "matters of fact" (2004, p. 231) and that he had contributed to "debunking science" (p. 232). According to Latour, post-truth politics is like critical radicalism gone mad, as if the STS virus of critique had escape from the scholarly laboratory, so that its deleterious effects could no longer be contained; as if the virus of criticism had mutated into a right-wing mutant and is now gnawing everything up (p. 231). Similar retractions were published by Sheila Jasanoff,[3] Sergio Sismondo (2017) and other TS protagonists.

3.4 RRI: restoring trust in science?[4]

Within the European context, the *Responsible Research and Innovation* (RRI) concept was launched more than a decade ago as part of the *Science With and For Society* (SWAFS) programme. RRI was launched in a top-down manner, in the context of EU funding programmes (Schomberg 2013; De Saille 2015, Owen et al 2021), but adopted and further developed by a thriving academic Social Sciences and Humanities community ("bottom-up"), giving rise to several salient RRI accounts (Timmermans and Blok, 2018), notably resulting in the AIRR concept, which focuses on *anticipation*, *inclusion*, *reflexivity* and *responsiveness* as key components of the "process dimension" of RRI (Owen, Macnaghten, and Stilgoe 2012; Stilgoe, J., R. Owen, and P. Macnaghten. 2013).

Currently we notice a shift in RRI discourse from conceptual work towards developing tools towards implementing and institutionalising RRI through mutual learning. One of the projects involved in this pragmatic turn was RRI Practice. The key objective of the RRI-Practice project (RRI-Practice - Official Website) was to understand the barriers and drivers to the successful implementation of RRI. Also, RRI Practice aimed to promote reflection on organisational structures and cultures of RPOs and Research Funding Organisations (RFOs) to identify and support best practices to facilitate the uptake of RRI in organisations and research programs, based on 12 case studies conducted in the context of the project. The project resulted in a series of recommendations to the EC to foster RRI uptake in different kinds of organisations and national cultures, in Europe but also worldwide. The project resulted in a series of recommendations to the EC to foster RRI uptake in different kinds of organisations and national cultures, in Europe but also worldwide.

One of the key outcomes of RRI Practice was a *Handbook for Organisations Aimed at Strengthening Responsible Research and Innovation* (Wittrock & Forsberg 2019). According to the authors, RRI calls for inviting society into the research and innovation processes in order to align such processes and their outcomes with the values and needs of society. It calls for democratising research and innovation in the sense of broadening the scope of actors, acknowledging a diversity of scientific and value-related perspectives, and accepting that opening up means drawing on a broader range of capacities and

[3] http://first100days.stsprogram.org/2017/03/28/what-should-democracies-know/

[4] Part of this section is also used D2.3 of the joinus4health project, also devoted to reflection on RRI.

competencies, increasing both the quality and impact of the research (p. 7). They emphasise that RRI is a response to current challenges, e.g., **mistrust** of science, to some extent triggered by scandals related to research misconduct, but it may also involve questioning the quality and **trustworthiness** of research due to industrialised scientific production. RRI, according to the authors, aims to take such challenges seriously. Structural change was seen as a particularly difficult objective in research practices and research organisation with little experience in working with RRI. In contrast, practices, and organisations where RRI had already gained a foothold were generally more responsive to further RRI implementation. In short, structural change is hardest to bring about where it is most needed (p. 6).

How to foster public trust in science instead of discontent? In a recent survey Novitzky et al (2020) signalled obstinate "discrepancies" between RRI ambitions as stated both in formal EC documents and in academic discourse, and their "implementation", i.e., the actual uptake and awareness of RRI objectives among researchers and innovators actually working on EC-funded projects. Ten years after the introduction of the RRI concept, which gave rise to an impressive series of RRI projects funded by the EC, a reflection on its key ambitions is called for, and was already initiated by scholars closely involved in RRI discourse and RRI projects, – for instance in a recent paper entitled "An unfinished journey? Reflections on a decade of responsible research and innovation" (Owen et al 2021). Moreover, self-reflection is an ongoing process, an inherent dimension of the RRI endeavour as such.

3.5 Take home messages

Several authors, including Wittrock and Forsberg (2019) mentioned above, have pointed to the link between responsible research and trust in science. Knowledge societies are facing multiple global challenges, aggravated by a climate of distrust and an omnipresent erosion of trust: in representative bodies and governmental organisations, in media and public information, but also in science, scientific expertise and innovation. Many voices in contemporary societies are questioning whether scientific information is sufficiently valid, disinterested and objective.

At the same time, in order to effectively address the global disruptions facing Europe and the rest of the world, evidence-based insights and potential solutions provided by science are indispensable. Therefore, the credibility, reliability and trustworthiness of science is an issue of crucial importance (Oreskes 2019).

In order to restore a trust relationship between science and society, science and society must be brought closer together. And to achieve this, it is necessary to look at both sides, considering both the scientific and the societal pole of the trust relationship. Distrust per se is not the issue, as fostering an attitude of scepticism and criticism is an intrinsic part of the scientific method, but how to distinguish between warranted and unwarranted trust? The importance of participatory research and open science for fostering trust in science in widely acknowledged, notably in an era when disruptive technologies (AI, robotics, etc.), disruptive events (the Corona pandemic) and the growing dependence of

research on private funding have contributed to distrust in scientific expertise. The IANUS project builds on the conviction that warranted trust in scientific research can be strengthened by fostering participation in research as a co-creative and inclusive process, sensitive to societal values, concerns and needs.

4. Ethical aspects of trust in science (T3.4)

Task 3.4: *Ethical aspects of trust in science* (M8–M16) [CEUT, EUR, RCL]

4.1 Ethics as Intrinsic Value: Trustworthiness and Integrity

Trust in science hinges upon the perception of trustworthiness (Resnik, 2015). Ethical conduct, characterized by transparency, honesty, and accountability, is foundational to trustworthiness (Shamoo & Resnik, 2015). Integrity, an ethical cornerstone, encompasses the commitment to truthfulness and adherence to ethical norms (Steneck, 2006). It is central to building and sustaining trust in scientific endeavours.

Good science embodies both high-quality research and ethical integrity. Ethics in research comprises a set of core values and principles that guide scientists. Researchers must adhere to these throughout their work, ensuring alignment with research ethics' values, principles, and legal standards, and securing necessary permissions from ethics committees.

Where do these moral values and ethical principles stem from?

Moral values and ethical principles in science are rooted in its goals and the expectations of the scientific community. Science serves as both a social institution and a profession, intertwining activities of various stakeholders for common purposes, such as knowledge, trust, and societal benefits (Resnik, 1998). It shapes our understanding of the world and forms a cultural and educational foundation. Achieving these objectives requires adhering to epistemological, moral, and social norms to ensure trust and cooperation. Unethical behaviours jeopardize this trust, so it is vital that scientists act honestly, objectively, and responsibly, considering potential impacts of their work. Science, as a profession, comes with specific values and standards, often outlined in ethics codes. Scientists wear many hats, sometimes leading to conflicting roles. Such conflicts can risk ethical breaches, emphasizing the importance of self-awareness and balance (Werhane & Doering, 1997).

Furthermore, research institutions play a crucial role in maintaining ethical standards, considering not just professional researchers but also the broader concept of citizen science (Iphofen & O'Mathúna, 2022). Their leadership often shapes the culture of integrity within these institutions (Boulter, 2020; Forsberg et al, 2018). Research funding sources also influence ethical standards, and both researchers and institutions must be discerning about the financing they accept, ensuring it aligns with broader ethical values.

4.2 *Telos* of Trust: Societal Goals and Ethical Imperative

Trust in science is instrumental in realizing crucial societal goals, including public health, environmental protection, and innovation (Sachs et al., 2019). Achieving societal goals through science entails ethical responsibility, emphasizing that trust is not merely a means but a moral end (United Nations, 2015). Ethical considerations guide the alignment of scientific research with societal values.

Building, restoring, and maintaining well-placed trust in science is imperative for the successful attainment of the Sustainable Development Goals (SDGs) and the realization

of the European Green Deal targets. Such trust underpins evidence-based decision-making, essential for sustainability policies. It encourages public support for green technologies and sustainable practices. Innovations in addressing environmental problems often rely on this trust, promoting investments in green technologies (Juma, 2016). Global challenges like environmental issues benefit from international cooperation built on mutual trust in science (Sachs et al., 2019). This trust also ensures research and policy accountability, enhancing scientific credibility (National Academies, 2017). Moreover, sustained trust is vital for the long-term commitment required in environmental solutions and influences funding decisions for impactful research. Central to sustainable development is the ethical conduct, which trust in science fortifies (United Nations, 2015). Lastly, societies that trust science can better adapt and remain resilient against environmental challenges.

4.3 Risk-Taking and Incomplete Evidence: Ethical Dimensions

Science often involves navigating uncharted territory, necessitating decisions in the face of incomplete evidence (Elliott, 2017). Ethical deliberation is essential in assessing risks, ensuring responsible experimentation, and minimizing harm (Resnik, 2015). Ethical dilemmas arise when considering the risks of novel research. Balancing principles such as autonomy beneficence, non-maleficence, and justice pointed out by Tom Beauchamp and Jim Chidldress in their landmark book “Principles of Biomedical Ethics” (Beauchamp & Childress, 2019). As none of these principles is overriding and they can be ranked differently depending on the specific context, they leave considerable room for judgement about individual cases and policies, as well as for negotiation and compromise (Beauchamp, 2010: 154).

Though widely regarded as foundational in bioethics, some scholars, such as Clouser and Gert (1990), have critiqued its vagueness in terms of normative guidance and lack of theoretical unity. Alternate ethical principles have been proposed, notably by the BIOMED II Project, which highlighted autonomy, dignity, integrity, and vulnerability (Rendtorff, 2002). In relation to scientists’ professional conduct, a significantly longer list of principles has been proposed by Adil E. Shamoo and David B. Resnik, who outline fifteen principles: honesty, objectivity, carefulness, credit, openness, confidentiality, collegial respect, honouring of intellectual property, freedom, protection of experimental animals, protection of human subjects, stewardship, respect for the law, professional responsibility and social responsibility (Shamoo & Resnik, 2015: 18-19).

In addition, ethical frameworks vary, with liberalism emphasizing individual rights and communitarianism the common good. While liberalism prioritizes the individual over society, communitarianism does the opposite. Though liberal individualism, focusing on autonomy, dignity, and privacy, was dominant in research ethics, recent shifts lean towards community values, dubbed the ‘communitarian turn’ in bioethics. Efforts have emerged to balance individual rights with the common good, yet mainstream research ethics often favours individual rights over broader societal interests (Sutrop 2011a, Sutrop 2011b).

To sum up, ethics underscores trust in science in profound ways: it serves as the intrinsic value of trust, emphasizing trustworthiness and integrity; it positions trust as a *telos* to achieve crucial societal goals, underlining the ethical imperative of responsible conduct; and it guides ethical considerations when taking risks in the pursuit of scientific advancement. Recognizing and embracing these ethical foundations is pivotal for fostering trust in science, thereby advancing knowledge and societal well-being.

4.4 Ethical frameworks for promoting ethical research: RE, RI and RRI

To ensure trust in science, researchers and other stakeholders have to fulfil their moral obligations. These obligations have been outlined and justified by three different fields of study: Research Ethics (RE), Research Integrity (RI), and Responsible Research and Innovation (RRI), which have been established and developed through a combination of historical events, ethical discussions, institutional efforts, and evolving societal needs. Here is a brief overview of these concepts.

4.4.1 Research Ethics (RE)

Research ethics, a branch of applied ethics, scrutinizes the moral challenges emerging from research activities and outlines foundational principles for responsible research conduct. At its core, research ethics prioritizes the protection of research subjects, both human and animal. In human research, this involves ensuring respect for participant autonomy, voluntary participation, and the procurement of informed consent. A fundamental tenet of research ethics is safeguarding participant privacy and ensuring the confidentiality of research data. When animals are involved, the Three R's principle – reduction, refinement, and replacement – is central. Beyond these protections, research ethics emphasizes the importance of managing potential risks and ensuring that the research yields benefits, either immediately for the participants or for future beneficiaries in comparable situations. Systems for ethical review, such as research ethics committees (REC) or institutional review boards (IRB), have been instituted to uphold these principles. Currently, many academic journals and research funding bodies necessitate approval from these committees.

Historical Perspective: Research ethics as an organized domain has its roots in the post-WWII era. Reflecting on its evolution, it is evident that the field primarily emerged as a countermeasure to the malevolent employment of scientific discoveries, such as nuclear weaponry, and the unethical research experiments on vulnerable populations without their informed consent. Notable examples include the Nazi medical experiments, the Imperial Japanese experiments in China, the Tuskegee syphilis study, the Milgram experiments, and the Tearoom Trade study (refer to Emanuel et al, 2011; Israel, 2015; Shamoo & Resnik, 2015; Briggle & Mitcham 2012). The inception of foundational ethical codes, namely the Nuremberg Code (1947) and the Declaration of Helsinki (1964), were reactions to these notorious breaches of ethical boundaries, serving as cornerstones for contemporary research ethics.

The broader field of bioethics, with medical ethics, in particular, significantly influenced the development of research ethics providing key insights. Bioethics explored the ethical dimensions of medical research and practice, foregrounding principles such as autonomy, beneficence, non-maleficence, and justice.

The formalization of RE saw the creation of institutional review boards (IRBs) and ethics committees within academic and research institutions. These entities are tasked with scrutinizing research methodologies to guarantee the ethical handling of human and animal subjects.

4.4.2 Research Integrity (RI)

Research integrity, rooted in professional ethics, emphasizes the ethical standards intrinsic to the researcher's profession. Institutional rules fortify it further, stipulating specific behavioural guidelines. For instance, researchers are expected to uphold honesty, refrain from falsifying, fabricating, or plagiarizing data, and sidestep conflicts of interest during peer-review processes. The scope of those upholding research integrity has expanded. Beyond individual researchers, research institutions, outcomes of research, and the broader scientific community are now recognized as stakeholders of research integrity (Meriste et al., 2016).

Historical Perspective: Research integrity gained traction in the latter half of the 20th century, spurred by increasing concerns about scientific misconduct such as data fabrication, plagiarism, and authorship controversies. Its origins can be traced back to incidents where researchers were found guilty of data falsification, fabrication, or plagiarism— notable examples include the Baltimore affair, the cold fusion debate, and the Hwang case (Resnik, 1998; Resnik, 2009; Shamoo & Resnik, 2015). Prominent cases like the controversy surrounding Robert Millikan's oil-drop experiment further underscored the significance of upholding research integrity. While its emergence was initially reactive, responding to instances of fraud, falsification, and plagiarism (FFP), research integrity should not be narrowly equated to misconduct alone. It should encompass a broader perspective, emphasizing the aspirational facets of research, or the ideals researchers strive for in their work. David Shaw differentiates between scientific integrity and research integrity, suggesting that the former encompasses the entirety of the research enterprise, inclusive of insights gleaned from past research (Shaw, 2019: 1086).

***The formalization of RI*:** As concerns about research misconduct grew, academic institutions and organizations responded by establishing dedicated offices and committees to oversee research integrity. These entities formulated policies and processes to manage and address allegations of unethical behaviour. For instance, the Office of Research Integrity (ORI) was founded in the United States, with equivalent bodies being set up in other countries, to guide and ensure responsible research practices. After the inaugural World Conference on Research Integrity in Lisbon, Portugal, in 2007, the European Network of Research Integrity Offices (ENRIO) was inaugurated. Presently, ENRIO boasts 31 members across 23 countries. In 2020, ENRIO formalized its status as

a legally recognized association, positioning itself to devise a sustainable strategy for both the network's future and the broader promotion of research integrity in Europe.

4.4.3 Research ethics and research integrity. A comparison

The definitions of 'research integrity' and 'research ethics' often intertwine, and a clear distinction between the two is not universally agreed upon. While some scholars view 'research integrity' as the broader category, others use 'research ethics' as the encompassing term, especially when referencing it as a distinct academic subfield akin to bioethics or environmental ethics. When 'research integrity' is perceived as the more comprehensive term, it typically alludes to the ideal ethos that researchers should embody. Integral to this ethos is the characteristic of having 'integrity' (Sutrop, Parder, Juurik, 2020: 75).

Historically, research ethics and research integrity operated as distinct domains. However, recent discourse suggests these areas are increasingly intertwined due to their shared concerns (Iphofen, 2020: 17). Initiatives from the European Commission, such as the EnTIRE and VIRT2UE projects, seek to address jointly research and scientific integrity. Drawing from Braun, Ravn, & Frankus (2020), one can interpret the distinction between the two in operational terms: ethics committees operate proactively, evaluating projects in advance, while research integrity offices address potential misconduct post-factum. Research ethics generally concerns project planning and design, whereas research integrity emphasizes the actual conduct of research. For ethical research practices, both domains are crucial, with their relevance oscillating based on the research's stage and trajectory.

4.4.4 Responsible Research and Innovation (RRI)

Responsible Research and Innovation (RRI) aims to ensure that scientific and technological advancements align with societal well-being, uphold ethical principles, and address pressing global challenges. It offers a comprehensive perspective on the purposes and impacts of research and innovation, transcending mere economic interests to advocate for an equitable and sustainable future. This framework encompasses the entire spectrum of technology assessments, from research inception to product marketing. Within this paradigm, ethics are seamlessly integrated into new technologies, exemplified by methods like the "privacy by design" approach. Central to RRI is the engagement of various societal actors, including the public, policymakers, civil society organizations, and marginalized groups, ensuring diverse perspectives and values are incorporated.

RRI stresses proactive anticipation of potential repercussions, interdisciplinary collaboration, and the inclusion of varied stakeholders. It emphasizes adaptability, ethics, transparency, rigorous governance, and the fortification of ethical education in research sectors. Ultimately, the approach seeks to cultivate ethical and beneficial outcomes in research and innovation. A hallmark of RRI is its promotion of sustainable practices, urging researchers to reduce environmental detriments, prioritize resource conservation, and align with broader sustainable development objectives.

The historical perspective: RRI evolved to address the intertwined ethical, societal, and environmental aspects of research and innovation. The ascendance of Science and Technology Studies (STS) during the mid-20th century significantly influenced this evolution. STS researchers stressed the importance of a thorough examination of the societal impact of scientific and technological initiatives, focusing on the production, dissemination, and societal integration of scientific knowledge. Throughout the latter part of the 20th century, spirited public discussions on matters such as nuclear energy and genetic engineering heightened consciousness about the nuanced consequences of these innovations. These discussions played a pivotal role in shaping ethical guidelines for scientific and technological endeavours, further emphasizing the need to embed ethical deliberations within research methodologies.

The formalization of RRI. The establishment of ethical review boards, the proliferation of research conduct guidelines, and the adoption of codes of ethics became increasingly prevalent in academic and research entities. The European Commission was instrumental in championing and structuring RRI. Inaugurated in 2014, the European Union's Horizon 2020 program integrated RRI as a foundational element of its research and innovation trajectory (Von Schomberg, 2011). Central to RRI are the tenets of ethical governance and the principled regulation of research and innovation. This encompasses the formation of regulatory blueprints and supervisory systems to encourage ethical behaviour and responsibility.

Although some critics label RRI as predominantly European, with perceptions in regions like the US gravitating more towards "broader impacts" (Davis & Laas 2014), the tenets of RRI have garnered global acknowledgment. Its influence has permeated research and innovation strategies across various global locales, transcending its European origins.

4.5 Take home messages

Some scholars argue that the prevailing ethics framework is overly influenced by biomedical sciences, suggesting its incompatibility with non-medical disciplines (Haggerty 2004; Schrag 2010; Schrag, 2011). They criticize the ethics review system as overly procedural, verging on a mere checkbox exercise. A notable examination of ethics review within the Commission Horizon 2020 revealed that ethics experts identified ethical issues at twice the rate of applicants across various funding schemes (Buljan, Pina & Marušić, 2021). Moreover, the emergent research integrity mechanisms, designed to address misconduct, risk casting all scientists in a suspect light, necessitating their continuous demonstration of compliance. Consequently, what was intended to foster trust between scientists and society, and amongst scientists themselves, is sometimes viewed as indicative of growing mistrust and increased bureaucratic encumbrance.

There is a growing consensus emphasizing the alignment of research's ethical and legal norms with the intrinsic values held by scientists. Ensuring mutual trust and adherence to ethical research standards necessitates that these norms resonate with scientists' core beliefs. A disconnect between institutional norms and personal values might predispose scientists to overlook or contravene these norms, thereby increasing the potential for

misconduct. Adherence to norms purely out of obligation or fear of repercussions rather than genuine belief undermines research integrity (Sutrop & Lõuk 2022).

In constructing frameworks for research ethics and integrity, it is imperative to establish a foundation of trust. Scientists should be perceived not merely as potential transgressors, but as professionals deeply familiar with ethical standards and committed to upholding professional ethics. Trust and cooperation are indispensable elements for the effective conduct of scientific research.

One should exercise caution and not place blind trust in scientists, as various violations reveal some may be unaware of or unwilling to adhere to ethical standards. The focus should be on judiciously assessing the trustworthiness of scientists and the potential risks of misconduct. Oversight of ethical behaviour should be balanced, emphasizing the responsibility of scientists and research institutions rather than excessive control.

5. Legal and regulatory aspects of trust in science (T3.5)

Task 3.5: *Governance and regulatory aspects* (M8–M16) [ISI, CEUT, SKU, KIT, RCL]

5.1 Introduction

Well aware that the concept of 'trust in science' used in this research project has an interpersonal property based on relationships between a truster (e.g. actors from society) and a trustee (actors from science), this task analyses exclusively the science domain, therefore addressing the aspect of trustworthiness of science. In this task, we aim to develop a concept of governance of trustworthiness in science.

Besides mere legal and regulatory aspects, we also want to include scientific standards and virtues that actors of the science system want to promote. In this sense, the concept follows a holistic governance approach that considers all relevant aspects and actors of science governance.

In view of the complexity and multi-layered nature of the concept of trust in science, the various starting points regarding the governance of science shall be clarified with the help of a simple micro-, meso-, macro-level model, to which the various actors and aspects of trust in science are assigned.

This task further fosters dialogue and interaction between WP3 and WP6, since a deeper analysis of science governance aspects is a fundamental prerequisite of developing recommendations for key recipients.

5.2 Concept operationalization for the governance of trust in science

5.2.1 The micro-level of governance

The micro-level of the governance of trust in science refers to the individual researcher or research projects and the research conducted in by individual researchers. The relevant question here is, what individual researchers can do to make research (more) trustworthy?

Here we basically see two areas of action - legal and regulatory aspects on the one hand and research values, virtues and standards on the other hand.

The **legal and regulatory aspects** refer to the fundamental freedom of science, legally enshrined in laws and constitutions in many democracies. This freedom is undisputedly a very high good. At the same time, freedom of science is not an absolute right of freedom. In fact, the freedom of science has limits, which is particularly evident in the case of the dual use dilemma. **As already discussed in WP5 (T5.3)**, there are limits to the freedom of science. In all areas of science, albeit to varying degrees, there is the danger that findings may be misused by third parties for harmful purposes. This possibility of using research findings for both beneficial and harmful purposes makes it difficult in many fields to differentiate between 'good' and 'bad' research, defensive and offensive

research, and research for peaceful or criminal purposes. Weighing the risk of potential misuse of research findings versus their benefits presents special challenges for the responsibility and self-control of researchers (German National Academy of Science Leopoldina & German Research Foundation, 2022). The dilemma refers to the risk of potential misuse of research, as Selgelid (2009) states: "The so-called "dual use dilemma" arises in the context of research [...] as a consequence of the fact that one and the same piece of scientific research sometimes has the potential to be used for harm as well as for good". Examples for dual use dilemmas are:

- In defense technology, materials research and nanotechnology can lead to the development of offensive weapons;
- Research results on pathogenic microorganisms and toxins can also be used for new bioweapons and for terrorist attacks;
- Analyses in molecular plant genetics can lead to bio attacks on seeds;
- Nuclear energy can be used for more than just peaceful purposes;
- In computer science, work to protect against computer viruses can promote not only their prevention but also their spread and new forms of cyberwar;

Researchers stress shortcomings in the debate about the governance of the so-called dual-use dilemma. One shortcoming is that framing dual-use concerns about individual and specific pieces of science and technology (for example research in biotechnology) as a single dilemma requiring a governance solution potentially distracts from the need for more fundamental changes in the politics and practice of dual-use governments. According to this argument, the research process has multiple stages, from research design, research funding decisions, research conduct and publication. The dual-use issue is therefore relevant to all actors involved in any of these stages. This includes researchers, funders, universities, private industries as well as publishers. It also goes deeper as it involves institutionalized research norms and practices, including national and international models of scientific governance and funding (Dando, 2011; Edwards et al., 2013; Whitman, 2010). Some guidelines therefore call for a wider governance of the dual-use dilemma, addressing the following aspects (German National Academy of Science Leopoldina & German Research Foundation, 2022):

- Risk analysis: Researchers should take account of the consequences and opportunities for application and misuse of their work and its controllability. In doing so, they should also consider the risks of not conducting the research in question

- Minimising risk: Research actors should minimise the risks associated with the implementation or use of their research. Measures on risk minimisation should be assessed and carried out both before and during an ongoing research project. Evaluating publications: The possible consequences of publishing results in high-risk research areas should be evaluated even before the start of the project. This applies, in particular, in cases where research results can lead to specific dangers or significant damages.

- Forgoing research as a last resort: Responsible decision-making by researchers may in individual cases lead to a high-risk project only being carried out at a later point

in time, following a research moratorium, or perhaps, not at all, even when the project is not prohibited by law.

- Documentation and communication of risks: If research entails risks of negative outcomes for society, scientists should document these risks, how they weigh up against possible benefits, and the measures taken to minimise them both before and, in the event of changes, during their work. Researchers should bring this documentation to the attention of the KEF responsible for these problems or the head of their institution before the research begins.

- Training and information: in university teaching and research training, researchers should communicate the principles of a responsible approach to research risks.

- Responsible Person: Evaluating whether research complies with legal provisions, self-regulatory measures and ethical principles is, in the first instance, the task of the researchers responsible for the project. In addition, the researchers' superiors bear responsibility, in particular within the scope of their legally required duty of supervision. The persons involved in the research should primarily inform the researchers responsible for the project, but if necessary also those researchers' supervisor and the responsible KEF, of legal violations that have occurred or could occur, as well as any ethical reservations.

The aspects of **research values, virtues and standards** refers to the discussion **of Research Integrity and Responsible Research and Innovation (RRI)**. Research Integrity incorporates the foundations of high-quality research and/or research excellence, that are fundamentally linked to the perception, development, establishment and incorporation of Responsible Conduct in Research (RCR) and Research Ethics (RE) standards and/or guidelines. It means conducting research in a way which allows others to have trust and confidence in the methods used and the findings that result from this. The ultimate goal is to produce scientific knowledge that underlies principles of accuracy, honesty, reproducibility and transparency. Research Integrity relates to all the stages of the research process, namely from the study design through the methodology applied, the data collected and analysed, the interpretations made and the way results are presented, where reflections on the social impacts of the research should always be taken into account (H2020 INTEGRITY, 2020). The European Code of Conduct for Research Integrity (ALLEA, 2017) defines four fundamental principles that aim to guide researchers in their work along with their engagement with the practical, intellectual and ethical challenges arising during their research: Reliability, Honesty, Respect and Accountability. Scientific misconduct is most clearly understood as (1) fabrication, (2) falsification or (3) plagiarism, following the ALLEA-code by All European Academies (2017). The reasons for scientific misconduct may vary and occur in various fields. Therefore, efforts have been coalescing around the idea of responsible research and innovation (RRI) to describe scientific research and technological development processes that take into account effects and potential impacts on the environment and society. RRI involves holding research to ethical standards, taking into account gender equality in

science, investing policy-makers with the responsibility to avoid harmful effects of innovation, public engagement and ensuring that societal actors have the knowledge neccessary to understand the implications by promoting science education and Open Access (Stilgoe et al. 2013, von Schomberg 2013).

5.2.2 The meso-level of governance

The meso-level of science governance addresses scientific organisation like universities, universities of applied science, research performing organisations, but also groups of research organisations (like Fraunhofer Society), professional associations (such as IEEE) or academic associations (like the European Sociological Association).

The relevant question here would be what actors on the meso-level can do to make research (more) trustworthy? Here too, we shed light on legal and regulatory aspects as well as standards or virtues, for example the implementation of governance structures like ombudspersons for good scientific practice and for research integrity, ethical committees, development of (binding) guidelines for researchers, codes of conducts, or institutional procedures for dealing with scientific misconduct. As analysed already in T3.5 Godlee et al. (2011) for example propose an expansion of research ethics committees within research organisation. They further suggest these organizational structures should not only approve ethics proposals but also implement a system that allows the committee to check if a study is doing what it set out to do in the proposal. They further argue that committees should be able to implement sanctions if, after publication, it turns out that a study did something other than what it was approved to do. On the other hand, this significantly increases the workload on these committees, redirecting the discussion back to funding issues focused on maintaining such committees within research institutions. Alternative solutions range from increased punitive measures as put forward by Godlee et al. (2011) to enforce codes of conduct, which all require continued and increased oversight. Clearly one of the most direct forms of change possible is a reduction in workload pressure put on researchers, and a focus on alternative forms of support in relation to grants and publications. The regards to organizational codes and conducts and guidelines Bonn et al. (2017) recognize that these elements are only a small portion of the efforts needed to promote research integrity and trust in science. However, given their value and preeminence among research institutions and also the potential for inconsistencies they must be carefully crafted to optimally promote a sustainable culture of good research. Given the discrepancy between various guides and codes of conducts of universities in Europe, the researchers call for an improved harmonization of future guidance.

Besides governance structures or codified schemes of securing research integrity and preventing research misconduct, science organisations can also offer mandatory trainings for researchers on Research Integrity, good scientific practice as well as on corruption in order to support a culture for Research Integrity at the institutional level. Zwart and Van der Meulen (2019) for example claim that a bottom-up approach is needed to ensure quality and good science, rather than top-down, through everyday practices on integrity. Research is often conducted through team-work and researchers are dependent on one

another. Hence an environment should be fostered in which researchers and research institutions together stimulate a research culture for reflection on integrity and shared responsibilities. In the Bonn PRINTEGER Statement (Forsberg et al., 2018) similar guidelines are proposed to minimise misconduct and stimulate integrity by focusing on strengthening its research culture. The statement proposes, among other things, proper instruction and training for (early career) researchers on integrity, fostering stronger attachment of researchers to institutions (such as by avoiding temporary contracts), stimulation of open dialogue between researchers and staff across hierarchical levels; a diversified incentive management by including a range of performance indicators, a good work environment and work satisfaction, and transparency about misconduct cases.

5.2.3 The macro-level of science governance

The macro-level of science governance refers to a systemic view on research. The question here would be what systemic actors can do to make research (more trustworthy), for example research funders or research policy actors.

While societal demands for research have increased massively in the last years research funding and research assessment processes, have been under growing scrutiny (Bonn and Bouter 2023). In particular the concept of research excellence and the application of too narrow criteria and indicators of research quality or impact that distort incentives are often criticized as they create unsustainable pressures on researchers. According to this critique, the narrowing of criteria and indicators reduces the diversity of research missions and purposes, leading researchers to adopt similar strategic priorities or to focus on lower-risk, incremental work. Furthermore, the systemic biases against those who do not meet these narrowed indicators of quality or impact, or who do not conform to certain career pathways, reduce the diversity and representative legitimacy of the research community (Curry et al. 2020). Therefore, efforts are coalescing around the idea of responsible research assessment (RRA), an umbrella term for assessment approaches that incentivize, reflect and reward the plural characteristics of high-quality research, support inclusive research cultures and are sensitive to local and particular contexts (Curry et al. 2020, Stilgoe et al. 2013).

6. Conclusions and Recommendations

6.1 The concept of trust and distrust

What is trust. Trust is a complex mental state which involves vulnerability and risk-taking by consciously giving up (some) control and being ready to feel betrayal. It may also involve a positive attitude towards the trustee and a belief that the trustee will have goodwill/integrity and fulfil or at least be able to fulfil the trustor's expectation (in moral relationships one expects the trustee to fulfil her commitment). Trust can involve an emotion, a belief, or both.

Public trust in science can mean several things: trust in what scientists say (epistemic trust) and what they do – trust in a scientific method (reproducibility and replication of their work), in research findings, in individual scientists, in research institutions, in products of science and innovation, in science as a system.

Trust and distrust. Trust is valuable when placed in trustworthy agents and activities. Distrust is justifieid when placed in untrustworthy agents and activities. Both trust and distrust have a gradual nature. Mere lack of trust does not automatically translate into active distrust, it could mean a neutral stance.

Complexity of Trust in Science: The foundation of trust is science lies in the competency of the trustee and the assurance that the trustee fulfils her fiduciary obligations. However, trust in science doesn't solely stem from the trustwortiness of science or scientists. It depends also on the social status, education, values, personal character traits of the trustor as well as on the level of trust in society, socioeconomic, political conditions in which trusting takes place.

Beyond the Cognitive Deficit Model: While earlier models might have suggested that educating the public is the straightforward path to gaining trust, it's clear now that the dynamics are far more intricate. Trust in science is influenced by a myriad of factors – from personal experiences to religious beliefs and political inclinations. Hence, while education is essential, focusing solely on it might not address the deeper-rooted reasons for distrust.

Making Science Trustworthy may not Secure Trust. Trust and trustworthiness are not always related. We can trust those who are untrustworthy and distrust those who are trustworthy. Thus trust in science does not depend solely on the trustworthiness of science. At the same time, efforts to bolster the trustworthiness of the science and scientists through ethical and responsible practices remain paramount.

6.2 Ethical Aspects of Trust in Science

Integrity as the Pillar of Trust: Trust in scientific processes and outcomes heavily relies on the perceived trustworthiness of the researchers and their work. At the heart of this trustworthiness is integrity, defined by a commitment to truthfulness and adherence

to ethical norms. It is only when science operates with utmost integrity that it can foster and maintain public trust.

Interdependence of Quality and Ethics in Science: The hallmark of robust science is not only its quality in terms of research rigor but also its ethical grounding. For research to be deemed trustworthy, it must align with the core values and principles of research ethics. This encompasses securing necessary permissions and ensuring all activities uphold legal and moral standards.

Moral Obligations Spanning Multiple Fields: Ensuring trust in science isn't limited to just adhering to research ethics. The broader spectrum of trustworthiness encompasses responsibilities derived from Research Ethics (RE), Research Integrity (RI), and Responsible Research and Innovation (RRI). Each of these fields has emerged from a combination of historical precedents, ethical debates, institutional initiatives, and evolving societal requirements, highlighting the multifaceted nature of the moral obligations that researchers and stakeholders must uphold.

Challenges and Implications of the Dual-Use Dilemma: Scientific findings, especially in certain fields, carry the risk of being misused for harmful purposes, creating a challenging situation for scientists and policymakers. Addressing this requires comprehensive risk analyses, potential measures to minimize risks, thoughtful evaluation of research publications, and occasionally even forgoing certain research endeavors. It's imperative to understand that while the dual-use issue is primarily associated with individual research findings, its implications reverberate throughout the broader research ecosystem, affecting stakeholders from researchers to funders, universities, industries, and publishers.

Need for Collective Efforts and Shared Responsibility: Whether at the individual, organizational, or systemic level, fostering trust in science requires concerted efforts from all stakeholders involved in the research process. It's not just about adhering to ethical standards but also about creating environments where researchers can openly communicate, reflect on integrity, and share responsibilities. From offering proper training on research integrity to ensuring transparency in research processes and outcomes, all steps contribute to building a more trustworthy scientific community.

6.3 The Multifaceted Nature of Trust Governance

Ensuring the trustworthiness of science requires attention at micro, meso, and macro levels, each with distinct challenges and solutions.

· **Micro-level:** The trustworthiness of individual researchers and research projects is bound by both legal regulations and inherent research values, virtues, and standards. While the freedom of scientific research is fundamental and cherished, it is not unlimited, especially when it comes to the dual-use dilemma where scientific findings might be used for both good and harmful purposes. Responsible research incorporates principles of

accuracy, honesty, reproducibility, and transparency, ensuring the generation of trustworthy scientific knowledge.

· **Meso-level:** Trustworthiness at the organizational or institutional level involves the implementation of governance structures such as ethical committees, guidelines for researchers, and codes of conduct. These structures, while essential, need consistent oversight and harmonization to prevent inconsistencies. Research institutions play a pivotal role in cultivating a culture of integrity by providing mandatory trainings, encouraging open dialogues, and promoting shared responsibilities among researchers.

· **Macro-level:** The systemic perspective on research governance reveals the increasing societal demands on research. With the growing scrutiny on research funding and assessment processes, there is a need to re-evaluate and broaden the criteria for what is considered "research excellence." A narrowed view on excellence might lead to biases, reduced diversity in research agendas, and undue pressures on researchers. The movement towards responsible research assessment (RRA) seeks to recognize and reward diverse high-quality research endeavors and support more inclusive research cultures.

6.4 Perception of Trust and Mechanisms to Build Trust

Scientists working in commercial companies are less trusted. Public trust in science remains relatively high in the European context, with the majority trusting the veracity of science and university-based scientists. However, there is a discernible difference in trust when it comes to scientists working in companies, potentially due to perceived conflicts of interest and less transparency in funding sources.

Mechanism to Build Trust: Emphasizing transparency in all scientific endeavors, particularly when it comes to funding sources, can reduce skepticism. For company-affiliated scientists, making research methodologies, data, and potential conflicts of interest public can help alleviate suspicions.

Trust in science is affected by the vicissitudes of the concept of truth. Contemporary knowledge societies are witnessing a widespread erosion of trust in various establishments, including representative bodies, media, and notably in scientific research and expertise. The questioning of the objectivity, validity, and impartiality of scientific information has become prevalent.

Mechanism to Build Trust: To bridge the widening gap of distrust between science and society, there's a recognized need to make scientific research more participatory, inclusive, and sensitive to societal values and concerns. Engaging society in research as a co-creative process can bolster warranted trust in scientific endeavors.

Research Integrity is a Concern: Despite efforts over the decades to promote ethics and integrity, there are still significant challenges in the research community, evidenced by the findings from the survey of active researchers. The neo-liberal university's model,

with its emphasis on metrics and market-oriented outcomes, might exacerbate issues of research integrity.

Mechanism to Build Trust: To mitigate the challenges posed by the neo-liberal university model, it might be beneficial to de-emphasize purely metric-driven evaluations and instead promote a more holistic assessment of research quality, ensuring that integrity and ethics are core components. Training in research ethics should be ubiquitous and mandatory across all levels of the research community.

Policy Adjustments Can Foster Greater Trust: The broader cultural frameworks and institutional norms significantly influence individual behavior. While personal attitudes are important, targeting the normative environment can be more effective in bringing about positive change. The emphasis on metrics can be counterproductive if not appropriately managed, leading to potentially undesirable and dysfunctional outcomes in research.

Mechanism to Build Trust: Introducing policies that promote 'ethically credible partnerships' can pave the way for more transparent collaborations between universities and the commercial sector. Such policies can ensure the maintenance of academic standards and research independence. Furthermore, the role of the "honest issue advocate" can help bridge the gap between scientific rigor and democratic principles, particularly in policy-making where science plays a crucial role. Lastly, a critical examination and potential reevaluation of performance metrics can lead to healthier research environments that prioritize quality over quantity.

In summary, while trust in science remains high in some regions, it's evident that there are areas of concern, especially around research integrity and commercial influences. Addressing these through targeted policy recommendations and emphasizing transparency and ethical conduct at all levels can help strengthen trust in science.

References

Acar, O. A. (2023). Crowd science and science scepticism. *Collective Intelligence*, *2*(2), 1–5. https://doi.org/10.1177/26339137231176480

ALLEA. (2017). *European Code of Conduct for Research Integrity*. https://www.allea.org/wp-content/uploads/2017/05/ALLEA-European-Code-of-Conduct-for-Research-Integrity-2017.pdf

Allum, N.C., et al. (2023). *Researchers on research integrity: a transatlantic comparison* https://osf.io/preprints/metaarxiv/fgy7c/

Baier, A. (1986). Trust and anti-trust. *Ethics* 96, 231–260.

Barber, B. (1987). Trust in Science. *Minerva*, 25(1/2), 123-134.

Beauchamp, T. L. (2010). Principles and Other Emerging Paradigms in Bioethics. In T. L. Beauchamp (ed.) *Standing on Principles. Collected Essays*. Oxford: Oxford University Press, pp. 153–74.

Beauchamp, T. L. & Childress, J. F. (2019). *Principles of Biomedical Ethics* (8th edn). Oxford University Press.

Bonn, N. A.; Godecharle, S.; Dierickx, K. (2017). European Universities' Guidance on Research Integrity and Misconduct: Accessibility, Approaches, and Content. *Journal of Empirical Research on Human Research Ethics* 12(1), 33-44.

Bonn, N. A., Bouter, L. (2023). Research assessments should recognize responsible research practices. Narrative review of a lively debate and promising developments. In Valdés, E, Lecaros, J. A. (Eds.): *Handbook of Bioethical Decisions*. Springer International Publishing.

Boulter, L. (2020). What Research Institutions Can Do to Foster Research Integrity, Science and Engineering Ethics, 26(4): 2363–9.

Bouter, L. (2023). Why research integrity matters and how it can be improved. *Accountability in Research*, *00*(00), 1–10. https://doi.org/10.1080/08989621.2023.2189010

Braun, R., Ravn, T. & Frankus, E. (2020). What constitutes expertise in research ethics and integrity?, *Research Ethics*, 16(1-2): 1–16.

Briggle, A. & Mitcham, C. (2012). *Ethics and Science. An Introduction*. Cambridge: Cambridge University Press.

Bromme, R., Mede, N. G., Thomm, E., Kremer, B. & Ziegler, R. (2022). An anchor in troubled times: Trust in science before and within the COVID-19 pandemic. *PLoS ONE*, *17*(2 February), 1–27. https://doi.org/10.1371/journal.pone.0262823

Buljan, I., Pina, D. G. & Marušić, A. (2021). Ethics issues identified by applicants and ethics experts in Horizon 2020 grant proposals, F1000Research, 10:471 DOI: 10.12688/f1000research.52965.2. (Accessed: 30.09.2023)

Ciubotariu, I. I. & Bosch, G. (2022). Improving research integrity: a framework for responsible science communication. *BMC Research Notes*, *15*(1), 1–5. https://doi.org/10.1186/s13104-022-06065-5

Clouser, K. D & Gert, B. (1990). A Critique of Principlism. *Journal of Medicine and Philosophy,* 15(2): 219–36.

Cobern, W. W., Adams, B. A., Pleasants, B. A. S., Bentley, A. & Kagumba, R. (2022). Do We Have a Trust Problem? Exploring Undergraduate Student Views on the Tentativeness and

Trustworthiness of Science. *Science and Education*. https://doi.org/10.1007/s11191-021-00292-1
Curry, S., de Rijcke, S., Hatch, A. et al. (2020). *The changing role of funders in responsible research assessment: progress, obstacles and the way ahead*. Working Paper, Research on Research Institute (RoRI).
Dando, M. (2011). Did we make a huge mistake over dual use? *Bulletin of the Atomic Sciences*. http://www.thebulletin.org/web-edition/columnists/malcolm-dando/did-we-make-huge-mistake-over-dual-use
Davies, M. & Laas, K. (2014). "Broader Impacts" or "Responsible Research and Innovation?" A Comparison of Two Criteria for Funding Research in Science and Engineering. *Science and Engingeering Ethics,* 20, 963–983
D'Cruz, J. (2020). Trust and Distrust. In J. Simon (ed.) *The Routledge Handbook of Trust and Philosophy*. New York: Routldege, 2020, pp. 41-51.
De Saille, S. (2015). Innovating Innovation Policy: The Emergence of 'Responsible Research and Innovation'. *Journal of Responsible Innovation,* 2 (2): 152–168.
Domenicucci, J. & R. Holton (2017). Trust as a Two-Place Relation. In P. Faulkner, T. Simpson (eds.) *The Philosophy of Trust*. Oxford: Oxford University Press, pp. 149–160.
Edwards, B., Revill, J., & Bezuidenhout, L. (2014). From Cases to Capacity? A Critical Reflection on the Role of 'Ethical Dilemmas' in the Development of Dual-Use Governance. *Science and Engineering Ethics*, 20(2), 571–582. https://doi.org/10.1007/s11948-013-9450-7
Elliott, K. C. (2017). Ethical Reasoning in Big Data: An Exploratory Analysis. *Science and Engineering Ethics*, 23(5), 1247–1270.
Emanuel, E. et al (2011). *Ethical and regulatory aspects of clinical research*. Baltimore and London: John Hopkins University Press.
European Commission (2012). *Responsible Research and Innovation: Europe's Ability to Respond to Societal Challenges*. Brussels: European Union.
European Commission. Directorate-General for Research and Innovation (2023a). *Horizon Europe, Research & innovation reshaping democracies*. Publications Office of the European Union. https://data.europa.eu/doi/10.2777/697544
European Commission, Directorate-General for Research and Innovation (2023b). *Mutual learning exercise on the whole of government approach in research and innovation – First thematic report – Introduction and overview of the whole of government approaches in research and innovation*. Publications Office of the European Union. https://data.europa.eu/doi/10.2777/92395
Faulkner, P. (2017). The Problem of Trust. In Faulkner, Paul and Thomas Simpson (eds.) *The Philosophy of Trust*, Oxford: Oxford University Press, pp. 110–129.
Fonseca, C., Pettitt, J., Woollard, A., Rutherford, A., Bickmore, W., Ferguson-Smith, A. & Hurst, L. D. (2023). People with more extreme attitudes towards science have self-confidence in their understanding of science, even if this is not justified. *PLoS Biology*, *21*(1), 1–26. https://doi.org/10.1371/journal.pbio.3001915
Forsberg, E. M. et al (2018). Working with research integrity – guidance for research performing organisations. *The Bonn PRINTEGER Statement, Science and Engineering Ethics,* 24(4): 1023–34.

Fricker, M. (2007). *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780198237907.001.0001
Furman, K. (2020). Emotions and Distrust in Science. *International Journal of Philosophical Studies,* 28(5), 713-730. https://doi.org/10.1080/09672559.2020.1846281
Frost-Arnold, K. (2014). Imposters, Tricksters, and Trustworthiness as an Epistemic Virtue. *Hypatia*, 29(4): 790–807. doi:10.1111/hypa.1210
German National Academy of Science Leopoldina & Research Foundation (2022). *Scientific Freedom and Scientific Responsibility. Recommendations for Handling of Security-Relevant Research.* https://www.dfg.de/download/pdf/foerderung/grundlagen_dfg_foerderung/sicherheitsrelevante_forschung/empfehlungen_de_en.pdf
Godlee, F., Smith, J., & Marcovitch, H. (2011). Wakefield's article linking MMR vaccine and autism was fraudulent. *BMJ*, *342*(jan05 1), c7452–c7452. https://doi.org/10.1136/bmj.c7452
Gould, Stephen Jay (1992). *The panda's thumb: more reflections in natural history*, New York, NY: Norton.
Govier, T. (1992). Distrust as a Practical Problem. *Journal of Social Philosophy* 23(1), 52-63
H2020 INTEGRITY. (2022, June 23). *H2020 INTEGRITY - Research Integrity*. https://h2020integrity.eu/integrity/
Haggerty, K. D. (2004). Ethics creep: Governing social science research in the name of ethics. *Qualitative Sociology*, 27(4): 391–414.
Hardin, R. (1996). "Trustworthiness". *Ethics* 107, 26–42.
Hardwig, J. (1991). The Role of Trust in Knowledge. *The Journal of Philosophy*, 88(12): 693–708. doi:10.2307/202700

Haven, T., Gopalakrishna, G., Tijdink, J., van der Schot, D. & Bouter, L. (2022). Promoting trust in research and researchers: How open science and research integrity are intertwined. *BMC Research Notes*, *15*(1), 1–5. https://doi.org/10.1186/s13104-022-06169-y
Hawley, K. (2014). Trust, Distrust and Commitment. *Noûs* 48, 1–20
Hawley, K. (2019). *How to be trustworthy*. Oxford University Press.
Heidegger, M. (2014). *Überlegungen II-VI (Schwarze Hefte 1931 – 1938)*. Gesamtausgabe IV. Abteilung: Hinweise und Aufzeichnungen (Band 94). Frankfurt am Main: Klostermann.
Hertzberg, L. (1988). On the Attitude of Trust. *Inquiry*, 31(3): 307–322. doi:10.1080/00201748808602157
Holton, R. (1994). Deciding to trust, coming to believe. *Australasian Journal of Philosophy* 72, 63–76.
Intemann, K. (2023). Science communication and public trust in science. *Interdisciplinary Science Reviews*, *0*(0), 1–16. https://doi.org/10.1080/03080188.2022.2152244
Iphofen, R. (2020). Regulating research: governance and ethics across the spectrum. In R. Iphofen (ed.) *Handbook of Research Ethics and Integrity,* Springer, pp. 17–32.
Iphofen, R. & O'Mathuna, D. (2022) Introduction: ethical evidence and policymaking. In R. Iphofen and D. O'Mathuna (ed.) *Ethical Evidence and Policymaking. Interdisciplinary and International Research,* Bristol: Policy Press, pp. 1–12.

Israel, M. (2015). *Research Ethics and Integrity for Social Scientists* (2nd edn). London: SAGE.
Jacquet, J. (2022). *The Play Book: How to Deny Science, Sell Lies, and Make a Killing in the Corporate World*. Allen Lane: Penguin Books
Jasanoff, S. (2011). Constitutional Moments in Governing Science and Technology. *Sci Eng Ethics* 17, 621–638. https://doi.org/10.1007/s11948-011-9302-2
Jones, K. (1996). Trust as an Affective Attitude. *Ethics*, 107(1), 4-25.
Jones, K. (2019). Trust, Distrust, and Affective Looping. *Philosophical Studies*, 176(4): 955–968. doi:10.1007/s11098-018-1221-5
Juma, C. (2016). *Innovation and Its Enemies: Why People Resist New Technologies*. Oxford University Press.
Keren, A. (2014). Trust and belief: a preemptive reasons account. *Synthese* 191: 2593-2615.
Knorr-Cetina, Karin (1981). *The manufacture of knowledge: an essay on the constructivist and contextual nature of science*, Pergamon international library of science, technology, engineering, and social studies, Oxford ; New York: Pergamon Press.
Knorr-Cetina, Karin (1999). *Epistemic Cultures: How the Sciences Make Knowledge*, Cambridge MA: Harvard University Press.
Kozlov, Max (2023). 'Disruptive' science has declined — and no one knows why. *Nature*, 613 (7943), 225–225.
Lahno, B. (2001). On the Emotional Character of Trust. *Ethical Theory and Moral Practice*, 4(2): 171–189. doi:10.1023/A:1011425102875
Lahno, B. (2020). Trust and Emotion. In J. Simon (ed.) *The Routledge Handbook of Trust and Philosophy*. New York: Routldege, 147–159.
Latour, B. (2004). Why has critique run out of steam? From matters of fact to matters of concern. *Critical Inquiry* 30 (2): 225–248.
Luhmann, N. (1979). *Trust and Power*. Toronto: Wiley
Luhmann, N. (1998). Familiarity, confidence and trust: problems and alternatives. In D. Gambetta (ed.) *Trust: making and breaking cooperative relations*. Oxford: Basil Blackwell.
Mayer, R. C., Davis, J. H. & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709–734.
McGeer, V. (2011). Trust, Hope and Empowerment. *Australasian Journal of Philosophy*, 86(2), 237-254.
McLeod, C. (2023). Trust. In Edward N. Zalta & Uri Nodelman (eds.) *The Stanford Encyclopedia of Philosophy* (Fall 2023 Edition). URL = https://plato.stanford.edu/archives/fall2023/entries/trust/
Mejlgaard, N., Bouter, L. and Gaskell et al (2020). Research integrity: nine ways to move from talk to walk. *Nature*, Vol 586: 14 October 2020
Meriste, H. et al. (2016). *Normative analysis of research integrity and misconduct,* PRINTEGER deliverable 2.3, Available at: https://printeger.eu/wp-content/uploads/2016/10/D2.3.pdf (Accessed: 30.09.2023)
National Academies of Sciences, Engineering, and Medicine. (2017). *Fostering Integrity in Research*. National Academies Press.
Novitzky, P., Bernstein, M., Blok, V., Braun, R., Chan, T., Lamers, W., Loeber, A., Meijer, I., Lindner, R., Griessler, E. (2020). Improve alignment of research policy and societal

values. Science 369 (6499): 10.1126/science.abb3415
O'Neill, O. (2002). *Autonomy and Trust in Bioethics*. Cambridge: Cambridge University Press.
O'Neill, O. (2018). Linking trust to trustworthiness. *International Journal of Philosophical Studies* 26(1), 293–300.
Oreskes, Naomi (2019). Why trust Science? Princeton and Oxford: Princeton University Press.
Owen, R., Macnaghten, P., & Stilgoe, J. (2012). Responsible Research and Innovation: from Science in Society to Science for Society, with Society. *Science and Public Policy,* 39(6), 751–760. https://doi.org/10.1093/scipol/scs093
Owen, R., Schomberg, R. von, Macnaghten, P. (2021). An unfinished journey? Reflections on a decade of responsible research and innovation, Journal of Responsible Innovation, DOI: 10.1080/23299460.2021.1948789
Owen, R., Bessant, J. & Heintz, M. (2013). *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*. John Wiley and Sons Ltd.
Pew Research Center. (2022). *Americans' trust in scientists, other groups declines*. *15*, 1–25. https://www.pewresearch.org/science/2022/02/15/americans-trust-in-scientists-other-groups-declines/%0Awww.pewresearch.org.
Potter, N. N. (2002). *How Can I be Trusted? A Virtue Theory of Trustworthiness*. Lanham, Maryland: Rowman & Littlefield.
Princen, Thomas (2005). *The logic of sufficiency*, Cambridge, MA: MIT Press.
Rasmussen, N. (2002). Of 'Small Men', Big Science and Bigger Business: The Second World War and biomedical research in the United States," Minerva, Vol. 40, No. 2, pp. 115–146.
Rendtorff, J. D. (2002). Basic ethical principles in European bioethics and biolaw: autonomy, dignity, integrity and vulnerability – towards a foundation of bioethics and biolaw. *Medicine, Health Care and Philosophy*, 5(3): 235–44.
Resnik, D. B. (1998). *The Ethics of Science. An Introduction*. London: Routledge.
Resnik, D. B. (2009). *International Standards for Research Integrity: An Idea Whose Time has Come?* Accountability in Research, 16(4): 218–28.
Resnik, D. (2011). Scientific Research and the Public Trust. *Sci Eng Ethics*, 17, 399-409.
Resnik, D. B. (2015). *Research Ethics: Cases and Materials*. Indiana University Press.
Rolin, K. (2020). Trust in Science. In J. Simon (ed.) *The Routledge Handbook of Trust and Philosophy*. New York: Routldege, 354-366.
Rooduijn, M., Van Kessel, S., Froio, C., Pirro, A., De Lange, S., Halikiopoulou, D., Lewis, P., Mudde, C. & Taggart, P. (2019). *The PopuList: An Overview of Populist, Far Right, Far Left and Eurosceptic Parties in Europe*. http://www.popu-list.org
Rosman, T., Bosnjak, M., Silber, H., Koßmann, J. & Heycke, T. (2022). Open science and public trust in science: Results from two studies. *Public Understanding of Science*, *31*(8), 1046–1062. https://doi.org/10.1177/09636625221100686
Rowland, J., Estevens, J., Krzewinska, A., Warwas, I. & Delicado, A. (2022). Trust and mistrust in sources of scientific information on climate change and vaccines. Insights form Portugal and Poland. *Science and Education*, 31, 1399-1424. https://doi.org/10.1007/s11191-021-00304-0
Rutjens, B. T., van der Linden, S. & van der Lee, R. (2021). Science skepticism in times of COVID-19. *Group Processes and Intergroup Relations*, *24*(2), 276–283. https://doi.org/10.1177/1368430220981415

Sachs, J. D. et al. (2019). Six Transformations to Achieve the Sustainable Development Goals. *Nature Sustainability*, 2(9), 805–814.
Schrag, Z. M. (2010). *Ethical imperialism. Institutional Review Boards and the Social Sciences, 1965–2009*. Baltimore: John Hopkins University Press.
Schrag, Z. M. (2011). The case against ethics review in the social sciences. *Research Ethics*, 7(4): 120–31.
Selgelid, M. J. (2009). Governance of dual-use research: an ethical dilemma. B*ulletin of the World Health Organization*, 87(9), 720–723. https://doi.org/10.2471/blt.08.051383
Shamoo, A. E. & Resnik, D. B. (2015). *Responsible Conduct of Research* (3rd edn). Oxford: Oxford University Press.
Shaw, D. (2019). The quest for clarity in research integrity: a conceptual schema. *Science and Engineering Ethics*, 25(4): 1085–93.
Sheman, N. (2020). Trust and Trustworthiness. In J. Simon (ed.) *The Routledge Handbook of Trust and Philosophy*. New York: Routldege, pp. 28-40.
Simis, M. J., Madden, H., Cacciatore, M. A. & Yeo, S. K. (2016). The lure of rationality: Why does the deficit model persist in science communication? *Public Understanding of Science*, *25*(4), 400–414. https://doi-org.ezproxy.utlib.ut.ee/10.1177/0963662516629749
Simpson, T. W. (2012). What is Trust? *Pacific Philosophical Quarterly* 93, 550–569.
Sismondo, S. (2017). Post-truth? *Social Studies of Science* 47(1): 3–6 DOI: 10.1177/0306312717692076
Song, H., Markowitz, D. M. & Taylor, S. H. (2022). Trusting on the shoulders of open giants? Open science increases trust in science for the public and academics. *Journal of Communication*, *72*(4), 497–510. https://doi.org/10.1093/joc/jqac017
Steneck, N. H. (2006). *ORI Introduction to the Responsible Conduct of Research*, US Department of Health and Human Services.
Stilgoe, J., Owen, R., Macnaghten, P. (2013). Developing a framework for responsible innovation. *Research Policy* 43, 1568-1580
Strawson, P. (1962). Freedom and Resentment. *Proceedings of the British Academy* 48: 1-25.
Sutrop, M. (2011a). Changing Ethical Frameworks: from individual rights to the common good? *Cambridge Quarterly of Healthcare Ethics,* 20(4): 533–45.
Sutrop, M. (2011b). Viewpoint: How to avoid a dichotomy between autonomy and beneficence: from liberalism to communitarianism and beyond. *Journal of Internal Medicine*, 269(4): 375–9.
Sutrop, M., Parder, M.-L., & Juurik, M. (2020). Research Ethics Codes and Guidelines. In R. Iphofen (ed.) *Handbook of Research Ethics and Integrity*. Springer, pp. 67–89.
Sutrop, M. & Lõuk, K. (2022). Ethical Research in a global context: a dynamic tension between universal values, principles and contextual applications. In R. Iphofen and D. O'Mathuna (ed.) *Ethical Evidence and Policymaking. Interdisciplinary and International Research.* Bristol: Policy Press, pp. 15–39.
Ullmann-Margalit (2004). Trust, Distrust and In Between. In R. Hardin (ed.) *Distrust*. New York: Russell Sage Foundation.
United Nations. (2015). Transforming Our World: The 2030 Agenda for Sustainable Development. United Nations.

von Schomberg, R. (2011). Introduction: Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields. In R. von Schomberg (ed.) *Towards Responsible Research and Innovation in the Information and Communication Technologies and Security Technologies Fields. A Report from the European Commission Services*. Luxembourg: Publications Office of the European Union, pp. 7–14.

von Schomberg, R. (2013). A vision of Responsible Research and Innovation. In: R. Owen, M. Heintz and J Bessant (eds.) *Responsible Innovation*. London: John Wiley, 51-74.

Wagner, W., Kronberger, N. and Seifert, F. (2002). Collective symbolic coping with new technology: knowledge, images and public discourse. *British Journal of Social Psychology,* 41, 323-343.

Wellcome Global Monitor. (2020). *How Covid-19 affected people's lives and their views about science*. https://wellcome.org/reports/wellcome-global-monitor-covid-19/2020

Werhane, P. and Doering, J. (1997). Conflicts of Interest and Conflicts of Commitment. In D. Elliot. & J. D. Stern (eds.) *Research Ethics: A Reader*. London: University Press of New England, pp 165–89.

Whitman, J. (2013). Nanotechnology and Dual-Use Dilemmas. In B. Rappert & M. J. Selgelid (Eds.), *On the dual uses of science and ethics: Principles, Practices, and Prospects* (pp.13-28). ANU E Press.

Wittrock, C., Forsberg, E-M. (2019). *Handbook for Organisations Aimed at Strengthening Responsible Research and Innovation*, RRI-Practice project report. Deliverable 17.6. ISBN 978-954-9456-24-0

Wynne, B. (1991). Knowledges in Context. *Science, Technology, & Human Values*, *16*(1), 111–121. https://doi.org/10.1177/016224399101600108

Zwart, H. (2022). *Continental philosophy of technoscience*. Series: Philosophy of Engineering and Technology. Dordrecht: Springer