Trustworthy Nanotechnology: Risk, Engagement and Responsibility

Bjorn K. Myskja

Introduction

The papers in this special edition are based on contributions to the symposium “Trustworthy nanotechnology” in Trondheim, Norway in June 2010. This symposium was initiated by the research project NANO\textsc{trust},\textsuperscript{1} funded by the Research Council of Norway’s Programme for nanotechnology and new materials. The motivation for the project was a realization that the question of trust in nanotechnology has been raised at a stage when the material, conceptual and social body of nanotechnology has been in its infancy. As trust or distrust in social and technological systems generally is based in previous experiences with these systems, nanotechnology raises new conceptual and social challenges. A number of public engagement exercises and research projects on the societal aspects of nanotechnology have, due to the relative novelty of the nano field, been fairly general in their involvement with the technology. One of the potential strengths of this project was its collaboration with a project investigating the use of polymeric micro- or nanoparticles to enhance vaccine formulations for farmed salmon. Thus we raised the question of trust in relation to a concrete case of nanotechnology development, integrating ELSA research with the technology research. Reflecting on trust in collaboration with technologists at this early stage has both advantages and drawbacks. It provides a unique opportunity in creating a forum for ethical reflection, opening the possibility for making a positive impact on technology development. By engaging technology researchers at this stage one can avoid the historical role of ELSA research as an exercise in negative ethics, i.e. ethics as border guard against real and perceived risks and harm to environment and society. On the other hand, being too far upstream may imply that the engagement is reduced to hypothetical discussions or conceptual clarifications.

Engaging with nanotechnology at this stage means that attempts to foster public trust would need to focus less on technological products and more on the complex process of science. While a products focus draws attention to the public’s choices, understanding, knowledge and acceptability of the product, a process focus draws attention to the visions and goals as well as organization and governance of nanotechnology. Thus the focus of the project and the symposium shifted from the conditions for public trust in general, to the body of scientific practitioners and practices that is to be trusted. This shift implies a corresponding shift

\textsuperscript{1}This project should not be confused with its Austrian namesake, starting at the same time, which has a more specific risk-research agenda: http://www.nanotrust.ac.at/

B. K. Myskja (✉)
Department of Philosophy, Norwegian University of Science and Technology, NO 7491 Trondheim, Norway
e-mail: bjorn.myskja@ntnu.no
from the concept of trust to the concept of trustworthiness and from what conditions trust to what conditions trustworthiness. With this focus, we aim at addressing to what extent nanotechnology is providing a sound basis for fostering public trust.

One important feature of recent approaches to technology studies is that they draw attention to extraordinary cases where there is much at stake, values are disputed and high risks are involved. These cases are often accompanied with high degree of controversy, either scientifically or publically, calling attention to post-normal strategies that widens the circles of engaged stakeholders. Controversies function as a mobilizing factor. The normal case however is not the one of controversies, but more like the case of exploring PLGA particles for delivery of vaccines in aquaculture. Projects like this one, that run under the heading of nanotechnology display low degree of substantial controversy. When the mechanisms of society “talking back” to science [3] are staged through initiatives like the NANO-TRUST project there is little visible and unusual scientific or public questioning—neither of individual projects of nanotechnology nor of nanoscience initiatives in general. Nanoscience and -technology differs from genomics in this respect: the “talking back” is staged in ways that may suppress internal scientific conflicts, conflicts on specific products, engage NGOs or media coverage in the ways that we have witnessed in connection with genomics.

Therefore the symposium sought to explore how the challenges related to trustworthiness can be raised in projects where there are few or no controversies to spark increased scrutiny and scientific self-reflection. What do such normal cases require from the working scientist in terms of their responsibility to display and be able to normatively account for co-produced societal changes? How can and should nanotechnologists meet the challenge of societal interaction under such presumably low risk circumstances? In this issue, three of the papers are products of the reflections of the NANO-TRUST research group on the trustworthiness of a technology application subject to uncertainty and ignorance regarding potential harms to health and environment. In addition, we include papers related to two of the invited presentations that contextualized the issue of trustworthiness based on other nanotechnology governance research projects.

**Technoscience, Complexity and Trust**

The issue of trust has arguably been at the centre of academic analyses of the complexities of modern society. Several sociological works have focused on risk as an inherent trait of life within a technologically dominated democracy, and have seen trust as a strategy to handle the complexities [9] and as a lubricant making society function and a key to progress [2]. Trust has a different role now compared to relations in the transparent village communities of the past [6]. The complexities of modern society have several sources, but the interdependent technological and social changes including rapid rise in living standards, education and democratization of policy processes are important elements. Given the central role of technology development in providing the material basis for these social changes, we see the enhanced significance of trust as interlinked with technology development in general.

Trust as the interrelation of technology and modern society is at the core of a more specified issue concerning the development of new technologies and technological practices. As Gibbons [3] famously stated:

> Modern science has until recently flourished partly because of a stable, underlying agreement between its practitioners and the rest of society. In other words, there has been a social contract between science and society, an arrangement built on trust which sets out the expectations of the one held by the other, and which—in principle—includes appropriate sanctions if these expectations are not met.

He goes on to state that a new contract is developing due to the complexities and accompanying uncertainties of modern society, leading to a new understanding of the basis for trust. Science under these new conditions are elsewhere called mode 2 in contrast to the old mode 1 research.

Before we turn to how trust is re-emerging as a key issue under these relatively new interrelations of science, technology and society we can briefly mention that trust even under this dying paradigm was not primarily based on a belief that the scientists were faultless and that nothing went wrong in the research. The beneficial knowledge produced was judged on a background of relative poverty, where
scientifically based technology was the vehicle to improve life and bring welfare. Any improvement in basic living conditions such as new medicines and therapies, improvements in food production and food security, reduction of harmful and painful working conditions would be significant. Many negative side-effects, such as toxic herbicides, dangerous machinery or therapies would be acceptable. When the starting conditions are bad, the benefits will be significant.

But the basis for public trust under a regime of the unchallenged authority of experts is stronger than these merely pragmatic grounds. In a society with an asymmetrical relation between an educated elite and a relatively uneducated public, there is a paternalistic pattern to the relation between trustors and trustee, paradigmatically expressed in the doctor-patient relationship. Just like young children accept parental authority and trust that their choices are for the best, the lay public trusted science. In this way we can say that the Age of Enlightenment came late to the field of sciences, as compared to other areas of society:

Enlightenment is man’s emergence from his self-imposed immaturity. Immaturity is the inability to use one’s understanding without guidance from another. This immaturity is self-imposed when its cause lies not in lack of understanding, but in lack of resolve and courage to use it without guidance from another. Sapere Aude! “Have courage to use your own understanding!”—that is the motto of enlightenment. [7]

Speaking of mode 1 as a paternalistic mode of science is certainly a simplified account, as scientific authority was probably never completely unquestioned. Compared to the present role of science in modern society, however, this account can illustrate the altered relationship between science and society. When children grow up, the basis for asymmetry in the trust relation is changed and, hopefully, a symmetrical relation of trust based on dialogue develops. The children “talk back” to the parents, who listen to them. But emergence from immaturity is not the same as always talking back. It is rather a matter of freely deciding when to talk back and when to trust the authority, be it parents or scientific expertise. Onora O’Neill [10] has pointed out how this altered relation between expertise and lay people has changed the field of medicine, and a similar analysis is relevant for the role of expertise in modern society in general.

Under this new paradigm for the relation between science and society, the positive end results of the risk-benefit equation is not self-evident any more, if we accept that it ever was. In an affluent society, the promised benefits of research are not sufficient to accept the risks because we perceive scientific research differently, and we know that the scope of expert knowledge is limited. This calls for new ways of conceiving of the conditions for societal trust in technology. Within this picture, nanotechnology raises some special challenges, as it is presented as a new field of science and technology. The basic story promises that new materials and methods will appear with new properties opening up for a number of positive effects. Experience tells us that new, powerful, beneficial technologies usually have been accompanied by more or less unexpected harm in a wide range of fields such as energy production, transportation technology, the development of pesticides, medicines and food products to mention some. If we believe the promises of the powers of technologies employing the nanoscale, we should suspect that there are comparable risks involved. We also know from recent events within genomics and medical technologies that when research results are applied as technologies in environment and society, science based risk assessments tend to be inadequate. The new relation between science and society and the lack of clarity about the nature of nanotechnology calls for an analysis of the conditions for trust under these circumstances.

Trust

There are many ways to understand the altered role of trust in modern society. Functional accounts of trust in institutions dominate sociology literature, but the paradigmatic instances of trust are not about institutions but concerns asymmetrical relationships between people. The Danish theologian Logstrup [8] points out that the human condition is one of vulnerability and holds that this is the basis for trust. Trusting someone involves an appeal that they take responsibility for our well-being without any guarantee that they actually will. We recognize the asymmetry of a paternalistic relationship in his account.
But it is not necessarily a one-sided paternalism, as two equal parties can trust one another in a relation of reciprocal asymmetry. Still, this is quite different from cognitive accounts that see trust as reasonable expectations about how the trustee will act. One example is Hardin’s theory of encapsulated interest, where the basis for the relation is the trustee’s belief that the trustor has integrated the interests of the trustor in her own interest ([6]: 17). But although Hardin’s theory seems to cover many central cases of trust, there are some important and relevant instances that cannot be explained by his account. Sometimes we trust people that we do not really believe care the least about our interests, simply because we have no better choice. Refugees around the world do that every day in order to escape. A cognitivist could reply that this is desperation, not trust, but it is typical of the way people in vulnerable situations tend to act. A theory that cannot account for the situations where we leave our fate in the hands of people we have no knowledge about, does not really capture essential aspects of trust. When Luhmann says that the function of trust in modern society is to reduce complexity, the implication is that we do not need to know the intentions of those we trust. If we knew, we would not need to trust them. Another case not explained by cognitivist theories is that we trust people in order to induce trustworthiness ([1]: 350 ff). In these cases we trust people despite inadequate knowledge sufficient to predict their action.

Such cases pose problems for cognitivist theories, and it is probably more correct to classify trust as an attitude, which is in line with Løgstrup’s account that we should regard interpersonal trust as a basic fact of the human condition, a way we perceive the world. A related observation is that in many instances “[t]rust can emerge when I choose to act as if I trust, that is, choose to make myself vulnerable to others, and find that nothing ill happen to me.” ([14]: 113). On this account, trust is neither as a belief nor a feeling but can be compared to a Kuhnian paradigm, a way of “seeing as” (Ibid, 102 ff). Distrust is not the normal state, but a consequence of failure or repeated failures of the trustee to live up to the expectations placed in them. We can also postulate a kind of distrust by analogy, in the sense that if we have negative experiences with someone, we tend to distrust similar people, groups or institutions.

This latter point can explain why trust, or rather distrust, has become a focus for discussions of technology and science in modern society, despite trust being the default position and an integral aspect of life in modern society. Several episodes involving technologists and other forms of expertise have been widely published as grounds for distrust in scientific expertise. Likewise, scientific controversies may contribute to this distrust. Even the new regimes of accountability and transparency, which are meant to enhance trust, may paradoxically lead to distrust [10,11]. The reception of genetically modified food in Europe has been an indication that distrust may represent a significant obstacle to development within science and technology. One explanation is that the scientific community acts as if they live under the conditions described above as mode 1, and fail to see that the paternalistic expert role is outdated. If they want to live up to the new task of science, scientists must engage in a dialogue with society.

**Trust in Nanotechnology**

Nanotechnology is a peculiar case for the study of the interaction between scientific expertise and society. We do not know whether it has been “contaminated” by the distrust of genomics for example in Europe, but the way the nano field is developed indicates a strong wish to avoid similar controversies. One example is the Royal Society & Royal Academy of Engineering [13] report with its focus on risk and public dialogue. But nanotechnology is a confusing theme for dialogue, because the technology has no focus in objects and methods. Nano is defined by size and socioeconomic factors, and consists of a collection of technologies from different scientific disciplines using different methods involving materials with very different properties. This is not a problem in most contexts as research is not dependent on what it is called, although funding to some extent is. When the issue is trust, however, the inclusiveness of the nanotechnology label becomes problematic. What exactly do we ask people to relate to if we ask them whether they trust nanotechnology? We may trust the research institutions or the people involved in it, or the research groups. Or we put our faith in the governmental control systems which can be described as systematic, institutionalized distrust that is a
condition for trust \[4,9\]. We know that misconduct appears, so the authorities have systematized control routines that operate on a generalized distrust. Because of this distrust, we can trust the systems without knowing the researchers personally. But these types of trust presuppose a kind of unity to the institutions we trust that we cannot expect from nanotechnology. This is the background for the NANOTRUST project.

The NANOTRUST Project

This project started out from the assumption that trust is a necessary prerequisite for social robust development and application of nanotechnology. Nanotechnology is framed by scientific uncertainty and ignorance both regarding positive and negative potentials of the technology, and we assume that independent research and development requires that politicians and the public have faith in the capabilities of science to handle these uncertainties. This requires that the scientists understand the impact of their activity as well as the limits to their knowledge and that this knowledge is communicated in a way that enhances trust. In addition, it is required that the scientists understand and accept public values and expectations for their activity. Therefore the aim of the project has been to gather perspectives on risk, social values and expectations with regard to introduction of nanotechnology, using nanoparticles for delivery of vaccines as our case study. To do that we cooperated with a technology project using nanoparticles as vaccine carrier in fish farming. We wanted to combine three strands of ELSA studies in one project: (1) A scientifically based risk and uncertainty analysis. (2) Engagement of the nanotechnology researchers in discussions of how they conceive of their own research and its societal implications. (3) A normative analysis of the basis for trust in this particular nanotechnology application, in order to study the relation between responsibility and trust connected to risk management and communication.

While carrying out the project, we realized that in designing the project we made the common mistake of talking about trust while being mainly concerned with trustworthiness \[5\]. Accordingly, we turned our focus from products to the complex process of science and the visions and goals as well as organization and governance of nanotechnology. The strength and weakness of our project is that it is an ordinary research project where, arguably, there is little at stake and we would not expect controversies which often function as a mobilising factor for engagement. Therefore we arranged a symposium in order to explore how the challenges related to trustworthiness can be raised in circumstances where there are few or no controversies to spark increased scrutiny and scientific self-reflection. We wanted to explore what such normal cases require from the researchers regarding their responsibility to display and normatively account for co-produced societal changes. How can and should nanotechnologists meet the challenge of societal interaction under such presumably low risk circumstances?

Trustworthy Nanotechnology?

One of the main issues in the governance of new technologies has been how to handle risks and uncertainties. The case of nanovaccine in aquaculture is not a high risk case, judging from studies of the use of these particles in other animals. But there are differences between fish and mammals, and the possible release of nanoparticles may raise new challenges in a marine environment. In the article ‘Mapping Uncertainties in the Upstream’ \(\text{Kåre N. Nielsen, Børge N. Fredriksen and Anne I. Myhr}\) suggest how to handle these uncertainties. The article is a result of the cooperation between the fish vaccine researchers and ELSA researchers, and exemplifies one of the results of the joint learning processes in integrated research. By using an Upstream Oversight Assessment modified with elements from an uncertainty analysis framework they seek to identify required areas of research based on literature study and the experiences of the nanovaccine project.

The novelty of the project raises particular challenges for a study of this kind, and this demonstrates one of the problems in relation to trustworthiness understood as transparency. For example, if the researchers communicate uncertainty regarding the scenario that these nanoparticles could act as “Trojan horses” making harmful substances in the environment enter the human body, while stating that this is highly implausible, the result can still be rejection of the project on very weak grounds. How to conduct
transparent communication of uncertainty in a responsible way calls for reflective judgment. The article suggests that the UOA is useful for an early mapping of benefits and concerns, and for identifying areas in need of further research. The cooperation between ELSA and technology researchers indicates one possible path towards early stage evaluations of possible negative side effects of concrete research projects, in order to facilitate a more robust nanotechnology development.

When there is scientific uncertainty regarding risks to environment and health, the precautionary principle is suggested as basis for regulation in a range of international treatises and policy documents. The principle has been formulated in a number of ways, but it is probably correct to say that all attempts have been subject to criticism from a variety of angles. Precaution is a regulatory principle that is simultaneously accepted and controversial. Some years ago, an UNESCO expert group suggested a new and more comprehensive approach to the implementation of the principle. In the paper ‘Precaution or integrated responsibility approach to nanovaccines in fish farming?’ Anne I. Myhr and Bjørn K. Myskja analyze the advantages and disadvantages of the UNESCO version by employing it to the use of nanoparticles for delivery of vaccines in aquaculture.

This version of the principle has some clear advantages as compared with earlier formulations by suggesting a procedural approach drawing on reflective judgment and public engagement in order to determine how to handle concrete cases. The authors do, however, criticize the UNESCO version for vagueness of key concepts and some weakly founded normative assumptions. They suggest that the precautionary approach should be supplemented by combining the ideal of best scientific practice with the insights of Jonas’ ethics of responsibility to ensure a sound basis for political decisions on technology regulation. On their account, scientific and political responsibility can be an alternative framework for the discussion of how to handle potentially harmful use of nanotechnology. Without discarding the need for public engagement, they emphasize that technologists and ELSA researchers are responsible for making a thorough and sound assessment of the technology and the societal issues involved, and share their best opinions with the public and political authorities.

We have pointed out earlier that the nanovaccine project is an example of nanotechnology projects that appear to be uncontroversial, and can be regarded as far from the major controversies dominating the academic literature on the challenges of this novel technology field. Kåre N. Nielsen, Trond G. Åm and Rune Nydal suggests an interpretation of the nanotechnology research field as a centre-periphery axis where this project is clearly situated on the periphery. Their paper ‘Centre and Periphery of Nano- A Norwegian Context’ points out that the scale-based definition of nanotechnology has been accompanied by controversies regarding the value of the research field. Claims of emptiness and hype, accompanied by the suspicion that the main function of the nano label has been to secure new sources of funding, raise the question whether the field is “worthy of our trust”.

The authors instead suggest a mode 2-approach, understanding nanotechnology as context driven research. Under this perspective they build their discussion on Rocco & Bainbridges definition of nanotechnology as novel nanoscale technologies which are conditioned by a convergence of traditional disciplines with potentially great socioeconomic impact. This standard vision is at the centre of nanotechnology, and here we find strategic research initiatives that seek to reorganize existing research according to this vision. In their interviews with researchers, the Nanolab at the Norwegian University of Science and Technology serves as a paradigmatic example of this work at the centre. Peripheral nanotechnology research activity, such as the nanovaccine project, is not marked by being committed to this vision. Such projects are oriented towards solving particular problems, and their acceptance of the nano label can be regarded as incidental. Still, their activity provides a crucial arena for discussing and validating what is to be achieved through the work of reorganization that takes place at the centre.

This reflection represents a widening of the perspective on trustworthy nanotechnology compared to the first two articles, from a project based focus on the nanovaccine case, to a process based focus on nanotechnology as a research field. One crucial general nanotechnology question concerns how to do research in a way deserving of public trust. As mentioned above, one common strategy in modern society is to regulate the field and establish control routines based on the principle of systematic, deper-
sonalized distrust. In this way the practitioners are forced to behave in a way that ensures that the institution or practice is trustworthy. But they themselves are not trustworthy in a normative sense. If I am forced to act morally right, my act is only in accordance with morality, not morally good, as Kant points out. Besides, regulation of morality may turn out to be counterproductive, if the rules and control mechanisms themselves produce incentives to act contrary to the intention of the regulation. In this way, a tight regulation with strict control mechanisms may create loopholes that stimulate people to act contrary to the ideals inherent in the normative system ([12]: 69 ff). Stimulating reflection on responsible practices may be a better solution. One such tool can be self-regulatory mechanisms such as codes of conduct and other forms of soft law.

In the article ‘Conversations about responsible nanoresearch’, Kamilla L. Kjølberg and Roger Strand address the question of responsible nanoresearch simply by asking the researchers themselves how they conceive of responsible nanotechnology. Taking the ‘European Commission recommendation on a code of conduct for responsible nanosciences and nanotechnologies research’ as point of departure, they engaged in conversations with a group of researchers from different fields of nanoscience. They knew the researchers from earlier collaborations and, perhaps more important, the researchers knew the authors and their perspective on nanotechnology governance. Thus there was already some basis for reciprocal trust, and some familiarity with the challenges identified with mode 2-science. On the basis of these conversations, the authors distinguish six idealized types of reactions to the document. They found a plurality of viewpoints both within the group and within the individual researcher, in the sense that the interviewees did not conform to one ideal type.

In the final part of the paper, Kjølberg and Strand present three conceptions of responsible nanoresearch. Even if the concept of responsibility developed in the code of conduct is a sound basis for reflection, it is not fully adequate, they argue. In order to develop a full notion of responsibility, we need increased awareness of moral choices, on the model of Arendt’s moral judgment as “thinking”. In this way, the individual responsibility of researchers is emphasized.

In the final paper in this special issue, trust as glue in nanotechnology governance presents a reflection on an exercise in stakeholder dialogue. Here the group of participants is extended beyond the researchers themselves, and include representatives of government, industry, trade unions and non-governmental organizations taking part in the German NanoKommission. The main focus of this study is how a group of stakeholders with diverse interests develops internal bonds of mutual trust based on their cooperative dialogue. These bonds lead to the formation of limits to what it is allowable to think and express in nanotechnology governance.

The article presents three vignettes about the work of the German NanoKommission illustrating that governance in practice can create social units where mutual recognition and trust serve as the glue holding the complex structure together. The participants develop mutual responsibility for the cooperative work. This kind of collaboration between different interest groups is certainly beneficial if the goal is reduction of public controversies in order to secure undisturbed technology development, with the appearance of public legitimacy. Åm argues that this is not adequate if we evaluate the process against the ideals of deliberative democracy, which are generally accepted as the normative basis for public engagement. Deliberative democracy emphasizes that all affected parties should be represented in the dialogue in order to secure a legitimate basis for policy decisions. So if psychological group processes of bonding between representatives of different interests create new patterns of loyalties, the basis for trusting the forum is undermined. Trust between the members serves as the glue ensuring good conditions for internal cooperation, but it weakens the trustworthiness of the group externally.

Acknowledgements The three first articles of this special issue are results of the NANOTRUST project funded by the Research Council of Norway. I would like to thank the participants of the ‘Trustworthy nanotechnology’-symposium for their contribution to the dialogue within the project, and reviewers for their contributions to improving this collection of papers.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.
References

1. Elster J (2007) Explaining social behaviour: more nuts and bolts for the social sciences. Oxford University Press, Oxford
2. Fukuyama F (1995) Trust: the social virtues and the creation of prosperity. Free Press, New York
3. Gibbons M (1999) Science’s new social contract with society. Nature 402:C81–C84
4. Grimen H (2009) Hva er tillit? Universitetsforlaget, Oslo
5. Hardin R (1996) Trustworthiness. Ethics 107:26–42
6. Hardin R (2006) Trust. Polity, Cambridge
7. Kant I (1983) What is enlightenment. In: Kant I (ed) Perpetual peace and other essays. Hackett, Indianapolis, pp 41–46
8. Løgstrup KE (1956) Den Etiske Fordring. Gyldendal, København
9. Luhmann N (1968) Vertrauen: Ein Mechanismus der Reduktion sozialer Komplexität. Ferdinan Enke, Stuttgart
10. O’Neill O (2002) Autonomy and trust in bioethics. Cambridge University Press, Cambridge
11. O’Neill O (2006) Transparency and the ethics of communication. In: Hood C, Heald D (eds) Transparency—the key to better governance? Oxford University Press, Oxford, pp 75–90
12. Pogge T (2002) World poverty and human rights. Polity, Cambridge
13. Royal Society & Royal Academy of Engineering (2004) Nanoscience and nanotechnologies: opportunities and uncertainties. Royal Society, London, Retrieved March 12, 2011, from http://www.nanotec.org.uk/finalReport.htm
14. Weckert J (2005) Trust in cyberspace. In: Cavalier RJ (ed) The impact of the internet on our moral lives. State University of New York Press, Albany, pp 95–117