



INSTITUTIONS YOU CAN TRUST

Felipe Morales Carbonell

OUTLINE

- i) Some comments on the 'big picture'
- ii) Trust and understanding
- iii) Institutions we can trust need understanding
- iv) Institutions we can trust need to manifest understanding
- v) Institutions we can trust are understandable

NO INSTITUTION CAN BE TRUSTED (?)

ART. X.—RESISTANCE TO CIVIL GOVERNMENT.

I HEARTILY accept the motto,—“That government is best which governs least;” and I should like to see it acted up to more rapidly and systematically. Carried out, it finally amounts to this, which also I believe,—“That government is best which governs not at all;” and when men are prepared for it, that will be the kind of government which they will have. Government is at best but an expedient; but most governments are usually, and all governments are sometimes, inexpedient. The objections which have been brought against a standing army, and they are many and weighty, and deserve to prevail, may also at last be brought against a standing government. The standing army is only an arm of the standing government. The government itself, which is only the mode which the people have chosen to execute their will, is equally liable to be abused and perverted before the people can act through it. Witness the present Mexican war, the work of comparatively a few individuals using the standing government as their tool; for, in the outset, the people would not have consented to this measure.

Like Thoreau, I *also* start from the rough intuition that institutions are not to be trusted.

But what I mean is that in many cases I have no reason to believe that *existing* institutions can be *fully* trusted.

So I think in many cases institutions deserve our distrust.

SOME INSTITUTIONS CAN BE (PARTIALLY) TRUSTED

This position still leaves it open if possible or even existing institutions could be fully or even partially trusted, which opens a bunch of interesting questions.

Under what conditions can institutions be trusted?

Can those conditions be engineered?

A WICKED PROBLEM

These questions seem to presuppose that there is a way for institutions to be trustworthy. I don't want to challenge this, but it is worth pointing out that the problem falls under the class of *wicked problems* that Rittel & Webber (1973) described as problems that can have no neat solutions because they cannot be formulated in a determinate way, have no definitive solutions, indeed are non-soluble, consequential and unique.

Wicked problems require that we raise awareness, accept their insolubility and adapt around them.

Policy Sciences 4 (1973), 155–169

© Elsevier Scientific Publishing Company, Amsterdam—Printed in Scotland

Dilemmas in a General Theory of Planning*

HORST W. J. RITTEL

Professor of the Science of Design, University of California, Berkeley

MELVIN M. WEBBER

Professor of City Planning, University of California, Berkeley

ABSTRACT

The search for scientific bases for confronting problems of social policy is bound to fail, because of the nature of these problems. They are “wicked” problems, whereas science has developed to deal with “tame” problems. Policy problems cannot be definitively described. Moreover, in a pluralistic society there is nothing like the undisputable public good; there is no objective definition of equity; policies that respond to social problems cannot be meaningfully correct or false; and it makes no sense to talk about “optimal solutions” to social problems unless severe qualifications are imposed first. Even worse, there are no “solutions” in the sense of definitive and objective answers.



Quality
you can trust

An offer of Coca-Cola is more than just an invitation to refresh. It's a way of saying with assurance, "here's something wholesome and delicious for you to enjoy." So for pleasure or refreshment—or both—have a Coke!

"COKE" IS A REGISTERED TRADE-MARK.

COPYRIGHT 1952, THE COCA-COLA COMPANY



SOME (NEGATIVE) CHARACTERISTICS OF TRUSTWORTHY INSTITUTIONS

Some general ideas:

- Institutions you can trust do not presuppose that you trust them to function properly (*robustness*).
- Institutions you can trust do not assert that you can trust them (*no self-advertisement*).

Why?

ROBUSTNESS

Trust in institutions allows institutions to work more efficiently, and distrust in institutions can unravel the functioning of institutions.

So there is an incentive for institutions to desire being trusted (if they value efficiency), and an incentive for institutions to be more or less insensitive to distrust.

This kind of robustness leads to *reliability*—we can predict the behavior of the institution under perturbations.

But we do not want complete insensitivity to distrust.

We need to find a balance.

NO SELF-ADVERTISEMENT

Given that there is an incentive for institutions to be trusted, we would expect there to be an incentive to present themselves or appear as trustworthy.

So there is an *incentive* to appear as trustworthy, and as a consequence of this, to self-advertise as trustworthy.

But this is easily exploitable. So the better policy is for the institution to act in a trustworthy manner without advertising itself as trustworthy.

Such form of non-self-advertisement is a form of transparency.

TRANSPARENCY AND TRUST

There is an idea that transparency breeds trust.

In a game, giving information to other players may weaken one's position and improve theirs. So giving information signals that you are not worried about their position relative to yours.

Institutional transparency makes it so that the stakeholders (the public) interacting with the institution are able to gain knowledge about the workings of the institutions, which is presumably linked to *accountability*: presumably, if you know what the institution does, and how it does it, you can trace the effects of its actions to the places/parts in the institutions that are responsible for those effects.



WORRIES ABOUT TRANSPARENCY

Recently, some authors have put into question the value of transparency.

Transparency can encourage people to be less honest so increasing deception and reducing reasons for trust: those who know that everything they say or write is to be made public may massage the truth. Public reports may underplay sensitive information; head teachers and employers may write blandly uninformative reports and references; evasive and uninformative statements may substitute for truth-telling. Demands for universal transparency are likely to encourage the evasions, hypocrisies and half-truths that we usually refer to as 'political correctness', but which might more forthrightly be called either 'self-censorship' or 'deception.'

— O'Neill 2002, p. 73

More recently, Nguyen (2021) has argued that transparency is a form of surveillance:

EPISTEMIC INTRUSIVENESS

- i) Public transparency demands that the experts justify their actions to non-experts.
- ii) Many of the actual reasons for expert action are not available to non-experts.
- iii) The demand for transparency creates pressure for experts to offer justifications for their action in terms available to the non-expert.
- iv) This pressure distorts, to varying degrees, expert action and communication

INTIMATE REASONS

- i) Some reasons are only available to people in some community or group.
- ii) So many of the actual reasons for action for members of that community are not available to outsiders.
- iii) The demand for transparency placed on that community—and its associated institution—asks that community members justify their actions in terms available to outsiders.
- iv) That demand distorts the actions and justifications of those community members.

UNDERSTANDING AND TRANSPARENCY

So much for stage setting. Now, onto the main point...

I want to take a certain epistemological perspective on these issues.

In particular, I want to discuss the link between *understanding* and trustworthiness, through the link between understanding and transparency.

The thought is: *only if you can understand someone, you can trust them.*

Why? Understanding them makes their reasons transparent. To understand a person requires (Grimm 2016) that we can grasp why they are motivated to act in the way that they do.

If I have a good enough grasp of someone's reasons, their actions cannot be too surprising for us.

INSTITUTIONS AND UNDERSTANDING

Now, I want to suggest a similar story in the case of trusting institutions.

Roughly: *only if you can understand an institution, you can trust it.*

But in this case, I think we also need to have stronger requirements for trustworthy institutions.

Namely:

- *Trustworthy institutions need to understand the problems that they solve.*
- *Trustworthy institutions act in ways that manifest this understanding.*
- *Trustworthy institutions act in ways that make themselves understandable.*

(As a matter of fact, analogue requirements may also apply to trustworthy individuals and groups; I omit this point here).

INSTITUTIONS THAT UNDERSTAND

What does it mean to say that an institution understands?

Roughly: it means that it has a certain cognitive grasp of an object that is the target of its understanding, through the use of representational devices of that target.

More broadly: we *say* that an institution understands a target if it behaves as if it had grasp of the target.

What does it mean to say that an institution *grasps* an object? That the structure of the institution affords it the possibility to act in certain salient ways with regard to it.

So, for example, an institution that aims to reduce the effects of crime in a society grasps crime in so far as it can perform certain actions with regard to crime, including cognitive tasks that aim at predicting the behaviour of crime, and so on.

INSTITUTIONS SHOULD UNDERSTAND

I said that *trustworthy institutions need to understand the problems that they solve.*

That is, we need to be able to say that these institutions do have a grasp of the problems that they try to solve.

A misled institution cannot be trustworthy because it is not appropriately connected to the ground truth of the problems that it tries to address. A misled institution will try to intervene into a problem in a way that does not address the problems that it is supposed to address; it is ineffective and thus unreliable.

HOW INSTITUTIONS SHOULD MANIFEST UNDERSTANDING

It is not enough that institutions understand, they need to *manifest* their understanding, make it manifest to possible observers.

This is part of why we characterized grasp in terms of the possibility of performing certain actions: we can say that they understand only on the basis of what we can come to know of their performances.

It should be part of what an institution does that it must display its understanding of the problems that it solves. It cannot just solve those problems in a way that is hidden to any observer (most institutions display themselves to at least some observers).

To this degree, institutions should be transparent (concerning their understanding).

That an institution is transparent with regard to any information does not entail that it is transparent about its understanding.

HOW INSTITUTIONS ARE UNDERSTANDABLE

By making their understanding manifest, institutions also make themselves understandable: they allow observers to understand them. The structure of an institution is rationalized by the way in which the institution solves problems.

By manifesting its own understanding of those problems, an institution makes it possible to understand the functioning of the institution.

IN FAVOUR OF OPEN INSTITUTIONS

The kind of transparency that I have described here does not constitute a nocive form of surveillance, as Nguyen worries.

It is a form of being *open* to people, in the way that it integrates institutions to observers, stakeholders, and other institutions.

Institutions are usually understood as stable objects that are constituted and built to stand without concern for interacting with agents that are not assigned roles in their structure (this contributes to their robustness). Open institutions account for interacting with agents which are not assigned roles in their structure.

Open institutions are less stable and efficient by design. That is OK.

VALUE ALIGNMENT

There is an additional dimension to the issue of trust: whether institutional goals are aligned or at least compatible with the values of those who may trust them.

If I don't align myself to the values of an institution I understand, can I trust it? **No.**

I can, however, *rely* on the institution to the extent that I can predict that it will act in ways that are misaligned to my goals.



SOME ADDITIONAL PESSIMISM

Even if open institutions are trustworthy, are other kinds of institutions trustworthy? Openness is a matter of degree.

How open should institutions be so that we can trust them? Given that strictly open institutions are rare, how far are we from living in an environment where we can mostly trust institutions?

Can we even come closer to such an environment?

REFERENCES

- Alfano, Mark & Huijts, Nicole (forthcoming). Trust and distrust in institutions and governance. In Judith Simon, *Handbook of Trust and Philosophy*. Routledge.
- Budnik, Christian (2018). Trust, Reliance, and Democracy. *International Journal of Philosophical Studies* 26 (2):221-239.
- Elliott, Kevin C. (2022). A Taxonomy of Transparency in Science. *Canadian Journal of Philosophy* 52 (3):342-355.
- Grimm, Stephen (2016). How understanding people differs from understanding the natural world. *Philosophical Issues*, 26:209–225.
- Lauret, Pierre (2018). Why (and how to) trust Institutions? Hospitals, Schools, and liberal Trust. *Rivista di Estetica* 68:41-68.
- Kelsall, Joshua (2024). Towards a Non-Reliance Commitment Account of Trust. *Journal of Value Inquiry*, 1-17.
- Marušić, Berislav (2017). Trust, Reliance and the Participant Stance. *Philosophers' Imprint* 17.
- Metzen, Hanna (2024). Trust in Scientific Expertise and the Varying Demands of Value Transparency. *Canadian Journal of Philosophy* 54 (5):364-379.
- Nguyen, C. Thi (2021). Transparency is Surveillance. *Philosophy and Phenomenological Research* 105 (2):331-361.
- O'Neill, Onora (2002). A Question of Trust: The BBC Reith Lectures 2002. Cambridge University Press.
- Railton, Peter (2014). Reliance, Trust, and Belief. *Inquiry: An Interdisciplinary Journal of Philosophy* 57 (1):122-150.

¡Gracias!



fmoralesc.github.io



ef.em.carbonell@gmail.com



okf@scholar.social