Why Do We Trust Nuclear Safety Assessments?

Failures of Foresight and the Ideal of Mechanical Objectivity

John Downer

- CISAC, Stanford -

August 2011

CISAC STANFORD | CENTER FOR INTERNATIONAL SECURITY AND COOPERATION

We are poor at planning for major technological disasters.

Fukushima illustrates this:

- No plan for using sea water as an emergency coolant.
- Few insights into the effects of sea-water on reactor.
- Geiger-counters maxed-out at insufficient levels.
- Insufficient stocks of 'Prussian Blue' for treatment of cesium 137. Very poor distribution of available stocks.

Fukushima is far from exceptional.

Western societies regularly make technological choices that are illogical if considered in terms of worst cases.

In respect to nuclear power, for example:

- Reactors are 'clustered' in close proximity, allowing failures to propagate.
- US 'Severe Accident Management Guidelines' are voluntary. The NRC doesn't require that operators demonstrate knowledge of them.

or in Aviation:

- No backward-facing seats / smoke hoods / anti-misting kerosine.

Our 'risk-assessment bureaucracies' *deny* certain hazards.

One of the main reasons for these failures of foresight lies in a misplaced, but institutionally deep-rooted, confidence that such failures ***will not happen.***

Result of a shift to risk regulation by *probabilistic analysis*.

If assessments find the likelihood of worst cases to be negligible, then there is **no incentive to plan for them**. Indeed, it creates **institutional problems** to doing so.

Once proofs that something will not happen have become official, it becomes very difficult -- **bureaucratically, rhetorically and legally** -- to require expenditure on planning for it.

This is impeccably rational (at least in principle).

Government is about *choices*.

- Societies cannot justify spending time, money and resources preparing for events that will not happen.
- There is a potentially infinite list of unlikely events that citizens might worry about, and potentially infinite amounts of money that governments could spend preparing for them.
- If we are comfortable that the likelihood of a disaster is almost non-existent, then we ought to be comfortable with a regime that ignores its implications.

But for this logic to work, risk assessments have to be very *credible*.

- We have to trust in assessments that assert a nuclear accident will probably never happen.

And technology assessments ***are*** very credible (or look it).

- They are elaborate, esoteric, formal, quantitative, calculative, '*objective*'.
- They are performed by scientists and engineers: professions that are institutionally (and often culturally) construed as dealing in '*hard facts*' rather than '*mere judgement*'.

Technological risk is construed as a *calculable variable,* with a *correct* and *obtainable* solution.

- This is "the bureaucratic vision of safety"; or "**The Ideal of Mechanical Objectivity**".

The Ideal of Mechanical Objectivity

Idea that you can measure the safety of a complex system, with socio-organizational dimensions and a stochastic operating environment, as you would (eg) the tensile strength of an iron bar.

Critical element of our 'civic epistemology' of technological risk.

Caricature but not a straw man.

Deeply embedded in our discourse, rules and institutions.

"The math is the math"

"It's not an opinion, its a calculation"

It is a treacherous ideal

1. We fail to plan for disaster because we trust in risk assessments [at least institutionally].

And...

2. We trust in risk assessments because we believe in the Ideal of Mechanical Objectivity.

Yet...

3. There are good reasons to believe the Ideal of Mechanical Objectivity **is misleading**. It's projections easily convey a false sense of surety.

So why is it so prevalent?

There is little interest in making its shortcomings explicit.

Engineers may not believe it literally, but it serves many institutional interests:

- Enhances the authority of engineers & regulators. (Who may be conditioned to communicating with a public with little technical understanding.)

- Defends policymakers against political, legal & reputational risks of overt political judgement.

So it has largely fallen to social scientists to highlight the inherent limits of technology assessments and undermine the ideal of Mechanical Objectivity.

Normal Accident Theory

Is the one major effort to make this argument *a priori* (ie: to explain why it *must* be true):

- Safety assessments dismiss fateful 'billion-to-one' coincidences.

Yet...

- Complex, tightly-coupled systems create the potential for billions of such events.

Thus..

- The kinds of fateful coincidences that safety assessments ignore are statistically *probable*.

Argument From Epistemology

- Compelling philosophical / sociological arguments for why even the most 'rigorous' knowledge-claims are always *logically incomplete* and *unproven*.
- Tests, experiments & models are inevitably 'theory laden' and this creates ambiguity.
- It is possible for expert beliefs to be wrong, even though they are *entirely logical and rigorous* (in relation to the best understandings of the time).

Thus:

- Some accidents that arise from engineering errors are epistemologically unforeseeable. *[Epistemic Accidents]*

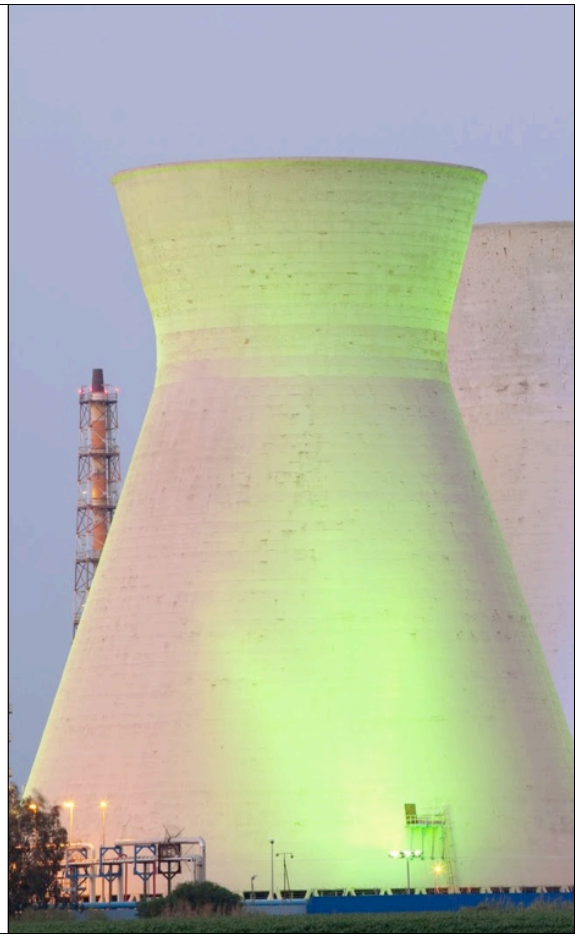
Why should it be necessary need to prove the fallibility of safety assessments in *principle* when they keep failing *in practice*?

Safety assessments keep proving to be wrong. Even in the (statistically tiny) nuclear industry there are far too many accidents and near accidents.

This should raise fundamental questions.

> "[Fukushima] means that assurances from the industry in Japan and overseas that the reactors were robust is now blown apart. [...] It raises fundamental questions on all reactors..." [*The Atlantic*]

With every accident in a safety-critical highly-regulated system -- be it a meltdown, plane crash, or exploding oil platform -- then some 'objective' calculation of safety is shown to be erroneous.



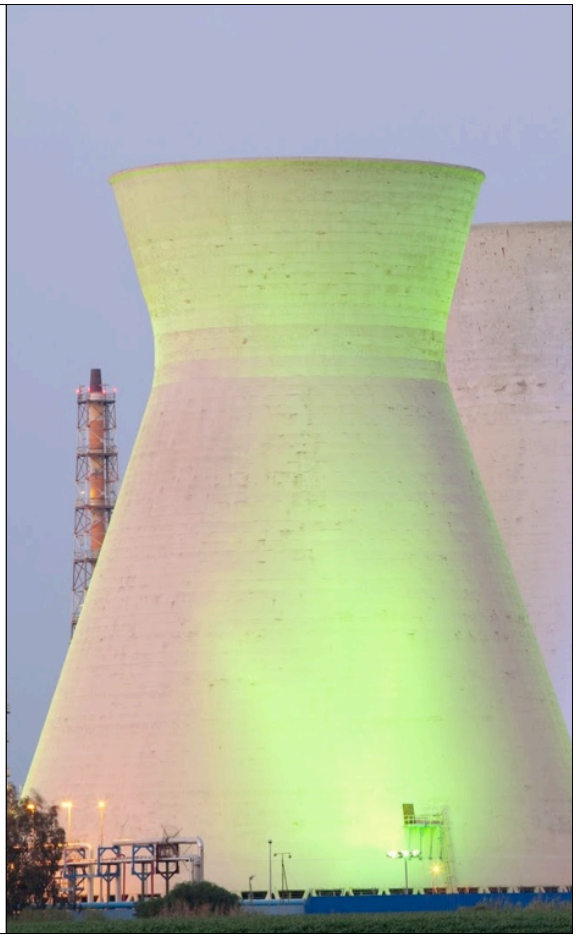
How does the Ideal of Mechanical Objectivity retain its institutional credibility?

Disasters fail to discredit the ideal of objective, calculative safety assessments.

So...

How does our civic epistemology withstand failure?

What happens when calculation fails?



1. Deny The Assessments Failed.

Either:

i. If accident is contained, then can say that the assessments took the failure into account.

- eg: (very early Fukushima): 'intact' plant + unusual circumstances = proof of assessment.

ii. Can paint assessments not as *calculations of the probability of failure* but as *calculations of risks to health,* then exploit the contested health effects.

- The effects of accidents are very *theory-laden*.
- Radiation kills indirectly, after long delays, in ways that are visible only through statistical inference. So by making conservative assumptions about health effects it is possible to redeem assessments.

2. Contest the 'Relevance' of the accident and its assessment:

Emphasize *differences* between the failing system and others. (eg: in design; circumstance; regulatory regime, etc).

eg: Chernobyl regulated very differently & by different people. It's failure could have implications for the USSR without having wider implications for assessment.

In Japan the nuclear regulator, NISA, was also responsible for *promoting* nuclear power.

Also see this with the technology itself:

Fukushima's design, from 1960's, is *different* from those that followed.

- "Using a plant built 40 years ago to argue against 21st-century power stations is like using the Hindenburg disaster to contend that modern air travel is unsafe." [Atlantic]

3. Concede a flaw but suggest a remedy.

Just as we identify and remedy flaws in the technologies themselves, so we can identify and remedy flaws in our assessment calculations and practices.

Assessments might be flawed because they failed to account for a specific unforeseen possibility (earthquake, tsunami, etc). So it is possible to adjust assessment processes, reassess all the plants, and move on.

- eg: NRC recently announced that US Nuclear safety assessments "...do not adequately weigh the risk of a single event that would knock out electricity from the grid and from emergency generators, as a quake and tsunami recently did in Japan". It is reframing assessments accordingly.

4. Blame an implementation error.

Claim that assessments were sound but there were errors, transgressions and deviances in its application.

- eg: reports TEPCO covered-up a series of failings, including data about cracks in critical circulation pipes, which may have proved fateful.

This is to say that the rules weren't being followed. All safety assessments come with the caveat (and hidden assumption) that the technology will be operated according to procedure.

Can redeem credibility by reforming the techniques of compliance, adding layers of oversight, and punishing wrongdoers.

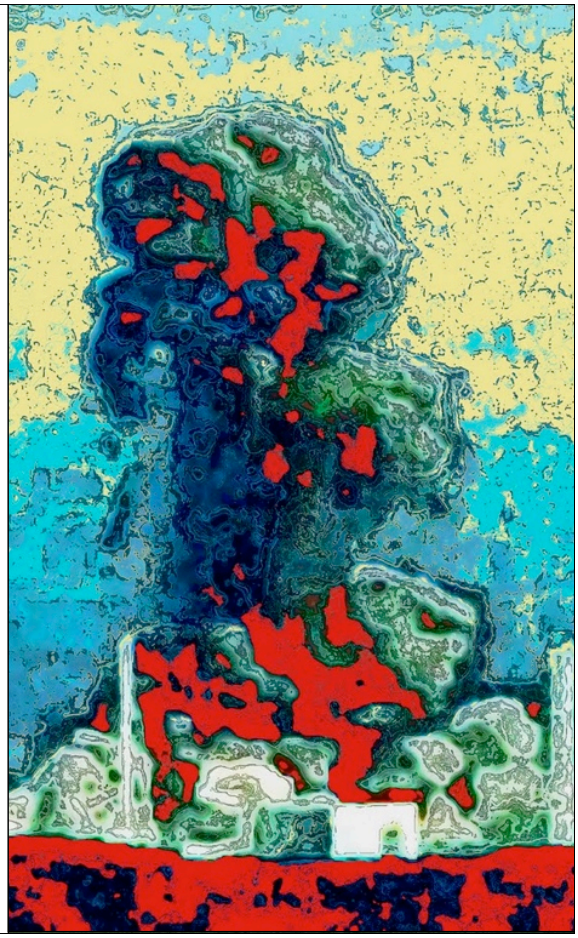
- eg: IAEA has outlined a five-point plan to strengthen the bureaucracy of nuclear oversight.

Can also make this claim of nature! (ie: earthquake was a freak)

Summary:

When the credibility of assessment practices is challenged, it is possible to say:

1. The assessments are fine, this failure was expected and predicted.
2. Assessment practices differ, and just because one was wrong, doesn't mean all are.
3. The assessments were wrong, but they are correct now.
4. The assessments are fine, but people failed to follow the rules. It will never happen again.



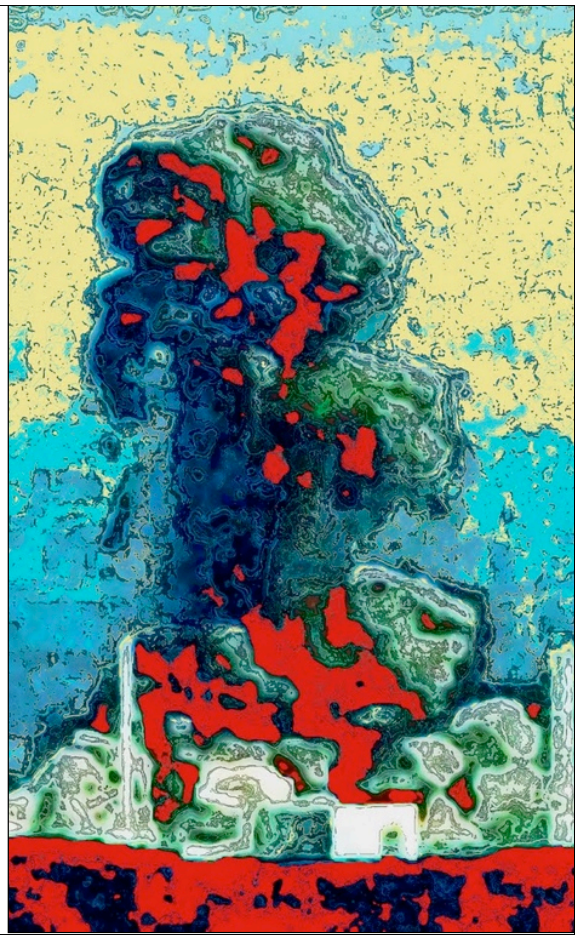
All these claims are credible, logical, and often valuable...

1. Risks might be overblown.

2. Assessment practices *do* differ, and it is not strictly logical to say the failure of one is the failure of all.

3. Assessments undoubtedly improve as we learn more.

4. People do break rules.



...but they hide the larger point:

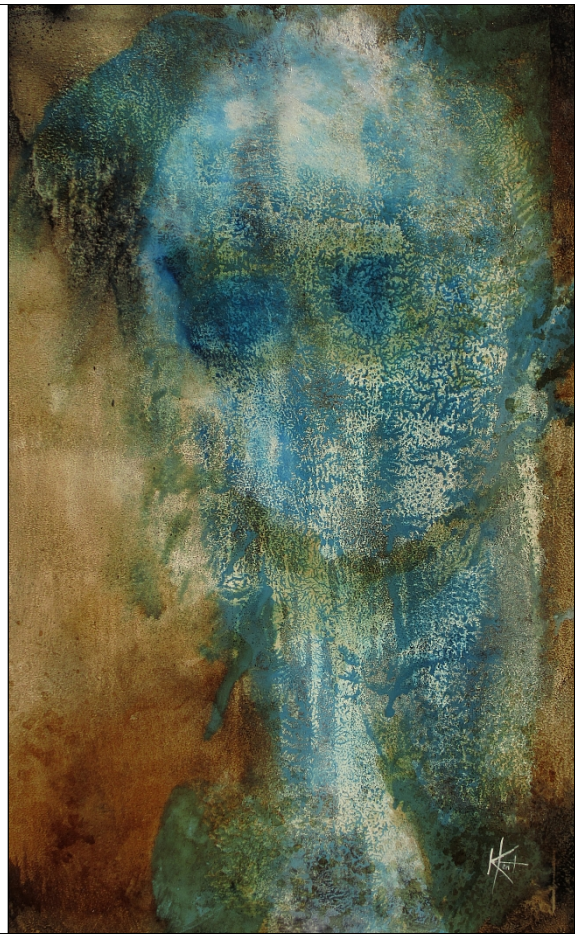
When we make these arguments, we are quietly reifying the mechanical ideal of objectivity.

We are implicitly saying that when assessments prove misguided it is because of an *aberration*, and not because of their *nature*.

In learning the close lessons, we are obscuring the larger lesson: that the safety of complex systems is not a property that can be accurately predicted with objective rigor.

Assessments are *theories*, not *facts*: imperfect reflections of reality.

This lesson, if it could be learned, would help us make better choices. (eg: We would be institutionally better able to plan for failure.)



Thank You