Paternalism and the Right to be Left Alone

Luciano Floridi

Professor of Philosophy and Ethics of Information
Director of Research
① Premise: the political problem of design

② Problems with the principles of design

③ Paternalisms

④ Conclusion
① Premise: the political problem of design

② Problems with the principles of design

③ Paternalisms

④ Conclusion
① Premise: the political problem of design
② Problems with the principles of design
③ Paternalisms
④ Conclusion
① Premise: the political problem of design
② Problems with the principles of design
③ Paternalisms
④ Conclusion
Information Societies are global

Global challenges require "big" political agents

Nation States are (were) THE information agents

Nation States rely on ICT = Information Societies

Information Societies empower information MAS

Information MAS become the new political agents
How to design political agents as multiagent systems in a post-Westphalian and post-Bretton Woods information society.
adding together consequents of the same antecedent

contraposition true only if original implication true
Tolerance is the first virtue of social institutions, as truth is of systems of thought. A theory however elegant and economical must be rejected or revised if it is untrue; likewise laws and institutions no matter how efficient and well-arranged must be reformed or abolished if they are intolerant. Each person possesses an inviolability founded on tolerance that even the welfare of society as a whole cannot override. For this reason tolerance denies that the loss of freedom for some is made right by a greater good shared by others. It does not allow that the sacrifices imposed on a few are outweighed by the larger sum of advantages enjoyed by many. Therefore in a tolerant society the liberties of equal citizenship are taken as settled; the rights secured by tolerance are not subject to political bargaining or to the calculus of social interests. The only thing that permits us to acquiescence in an erroneous theory is the lack of a better one; analogously, an intolerance is just only when it is necessary to avoid an even greater intolerance. Being first virtues of human activities, truth and tolerance are uncompromising.
Justice is the first virtue of social institutions, as truth is of systems of thought. A theory however elegant and economical must be rejected or revised if it is untrue; likewise laws and institutions no matter how efficient and well-arranged must be reformed or abolished if they are unjust. Each person possesses an inviolability founded on justice that even the welfare of a whole cannot override. For this reason justice denies that the loss of freedom for some is made right by a greater good shared by others. It does not allow that sacrifices imposed on a few are outweighed by the benefits or advantages enjoyed by many. Therefore in a just society the liberties of equal citizenship are taken as settled; the rights secured by justice are not subject to political bargaining or to the calculus of social interests. The only thing that permits us to acquiescence in an erroneous theory is the lack of a better one; analogously, an injustice is tolerable only when it is necessary to avoid an even greater injustice. Being first virtues of human activities, truth and justice are uncompromising.
① Premise: the political problem of design
② Problems with the principles of design
③ Paternalisms
④ Conclusion
THE POLITICAL PROBLEM OF DESIGN


Justice

Tolerance

Two design principles not only one.
THE POLITICAL PROBLEM OF DESIGN

Problems with the principles.

Justice

Value of Uncertainty

Paternalism

Tolerance

Paradox
Too much T.

If information about T then abuse of T.
Too much I
then no
fairness.

More I less T.

PROBLEMS WITH THE PRINCIPLES OF DESIGN

Justice

Paternalism

Value of Uncertainty

Tolerance

Paradox
PROBLEMS WITH THE PRINCIPLES OF DESIGN

Too much P then no freedom no tolerance.
① Premise: the political problem of design
② Problems with the principles of design
③ Paternalisms
④ Conclusion
TOLERATION
An agent A ψ (acts, behaves, does, omits doing or designs something, ...) tolerantly towards another agent B’s φ-ing (believing, thinking, acting, behaving, ...) iff:
① A is informed about B’s φ-ing;
② A disapproves of B’s φ-ing;
③ A could ψ to interfere with B’s freedom (liberty or autonomy) to φ;
④ A refrains from ψ-ing.

For example, Alice is informed about Bob’s smoking, she disapproves of it, and could forbid him to smoke in her house, but refrains from doing so.
PATERNALISM
An agent $A \psi$ (acts, behaves, does, omits doing or designs something, …) paternalistically towards another agent $B$’s $\phi$-ing (believing, thinking, acting, behaving, …) if and only if

1. $A$ is informed about $B$’s $\phi$-ing
2. $A$ is informed about $B$’s $\phi$-ing not improving $B$’s well-being
3. $A$ could $\psi$ to interfere with $B$’s freedom (liberty or autonomy) to $\phi$
4. $A$ does $\psi$
5. $A$ does $\psi$ because $B$’s $\phi$-ing does not improve $B$’s welfare
6. $A$ does $\psi$ without $B$’s consent.

Alice knows and is concerned that Bob smokes and forbids him to do so in her house for the sake of his health and against his will.
Paternalism: X acts paternalistically towards Y by doing (omitting) Z: [if and only if]
① Z (or its omission) interferes with the liberty or autonomy of Y
② X does so without the consent of Y
③ X does so just because Z will improve the welfare of Y (where this includes preventing his welfare from diminishing), or in some way promote the interests, values, or good of Y.

Stanford Encyclopedia of Philosophy
The dilemma of tolerant paternalism:

Is tolerant paternalism possible?
[...] compelled \textit{ἀναγκάζοιτο} to stand up suddenly and turn his head around and walk and to lift up his eyes to the light, and in doing all this felt pain and, because of the dazzle and glitter of the light, was unable to discern the objects whose shadows he formerly saw. Plato, \textit{Republic}, 7.515c
Then the master told his servant, “Go out to the roads and country lanes and compel ἀνάγκασον them to come in, so that my house will be full”.

*Luke 14:23*

New International Version.
The dilemma of tolerant paternalism:

Is tolerant paternalism possible?

Or: how can A both respect B’s freedom and interfere with it for B’s sake, without B’s consent, on the basis of a more privileged epistemic position?
FSA as a simple scheme of action (System/Society) that “consumes” actions as inputs:

1. Finite set of **states** $S$: \{S$_1$, S$_2$, S$_3$, S$_4$\};
2. Finite set of input/actions $A$ (**alphabet**): \{A$_a$, A$_b$, A$_c$\};
3. A **transition function** $f$: $S \times A \rightarrow S$;
4. S$_1 \in S$ is the **start state**; and
5. F $\subseteq S$ is the set of **accept states**.
System states

<table>
<thead>
<tr>
<th>Actions</th>
<th>S₁</th>
<th>S₂</th>
<th>S₃</th>
<th>S₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aₐ</td>
<td>S₂</td>
<td>S₃</td>
<td>S₂</td>
<td>S₄</td>
</tr>
<tr>
<td>Aₐ</td>
<td>S₁</td>
<td>S₁</td>
<td>S₂</td>
<td>S₄</td>
</tr>
<tr>
<td>Aₖ</td>
<td>S₁</td>
<td>S₁</td>
<td>S₄</td>
<td>S₁</td>
</tr>
</tbody>
</table>

Scheme, not Model, Blueprint, TE (constraints).
### System states

<table>
<thead>
<tr>
<th>Agents</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_a$</td>
<td>$S_2$</td>
</tr>
<tr>
<td>$S_3$</td>
<td></td>
</tr>
<tr>
<td>$S_2$</td>
<td></td>
</tr>
<tr>
<td>$S_4$</td>
<td></td>
</tr>
<tr>
<td>$A_b$</td>
<td>$S_1$</td>
</tr>
<tr>
<td>$S_1$</td>
<td></td>
</tr>
<tr>
<td>$S_2$</td>
<td></td>
</tr>
<tr>
<td>$S_4$</td>
<td></td>
</tr>
<tr>
<td>$A_c$</td>
<td>$S_1$</td>
</tr>
<tr>
<td>$S_1$</td>
<td></td>
</tr>
<tr>
<td>$S_4$</td>
<td></td>
</tr>
<tr>
<td>$S_1$</td>
<td></td>
</tr>
</tbody>
</table>

### Virtue Ethics – Micro-Macro system states.
Consequentialism, Deontologism (why, not what).
System states

<table>
<thead>
<tr>
<th>Agents</th>
<th>Actions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_a$</td>
<td>$S_2$</td>
</tr>
<tr>
<td>$A_b$</td>
<td>$S_1$</td>
</tr>
<tr>
<td>$A_c$</td>
<td>$S_1$</td>
</tr>
</tbody>
</table>

Environmental Ethics (ontocentric).
System states

<table>
<thead>
<tr>
<th>Agents</th>
<th>Actions</th>
<th>S₁</th>
<th>S₂</th>
<th>S₃</th>
<th>S₄</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aₐ</td>
<td>S₂</td>
<td>S₃</td>
<td>S₂</td>
<td>S₄</td>
<td></td>
</tr>
<tr>
<td>Aₖ</td>
<td>S₁</td>
<td>S₁</td>
<td>S₂</td>
<td>S₄</td>
<td></td>
</tr>
<tr>
<td>Aₗ</td>
<td>S₁</td>
<td>S₄</td>
<td>S₁</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Ethics as axiology indicates what moral value is and which state is morally preferable (has value).
1) Do not allow going from $S_x$ to $S_y$. Paternalism as **prohibitionism**. Incompatible with Freedom/Toleration. Compatible with shaping behaviour.
3) Make going from $S_x$ to $S_y$ more/less difficult. Paternalism as **ethics by design** (policy of disincentives).

Not very compatible with Freedom/Toleration. Compatible with shaping behaviour.
4) Make $S_z$ a better option than $S_y$. Paternalism as nudging (persuasion). Not very compatible with Freedom/Toleration. Compatible with shaping behaviour.
5) Place R (reflection) between $S_x$ and $S_y$. Paternalism as pro-ethical design. Compatible with Freedom/Toleration. Compatible with shaping behaviour.
In a complex information society we may need to use all kinds of paternalism to shape the topology of the network of possible actions whenever harm would be irreparable.

However, when possible, we should adopt the form of paternalism most consistent with the fostering of freedom and toleration, i.e. pro-ethical design.
“You need the haystack to find the needle”, General Keith Alexander, Director of the National Security Agency (NSA).
**Intercepting** is a form of paternalism by an agent A (the *paternaliser*) towards a paternalised B that is
① meant to shape the behaviour of an agent B
② by interfering with B’s freedom,
③ for B’s own good,
④ without B’s consent.

Paternalism assumes that:
A knows better than B what is better for B and can shape B’s behaviour.
PATERNALISM
An agent A ψ (acts, behaves, does, omits doing or designs something, …) paternalistically towards another agent B’s φ-ing (believing, thinking, acting, behaving, …) if and only if
① A is informed about B’s φ-ing
② A is informed about B’s φ-ing not improving B’s well-being
③ A could ψ to interfere with B’s freedom (liberty or autonomy) to φ
④ A does ψ
⑤ A does ψ because B’s φ-ing does not improve B’s welfare
⑥ A does ψ without B’s consent.
Alice knows and is concerned that Bob smokes and forbids him to do so in her house for the sake of his health and against his will.
A acts paternalistically towards B by intercepting if and only if:

① A’s intercepting interferes with B’s freedom
② A intercepts without the consent of B
③ A intercepts just because intercepting will improve the welfare of B (where this includes preventing B’s welfare from diminishing), or in some way promote the interests, values, or good of B.

Stanford Encyclopedia of Philosophy
Problems with the “interference” condition.
① influences behaviour: not in practice and not meant to in theory
② interferes with freedom: ditto
③ withdraws metainformation hence interferes/influences: no, disanalogy with medical case
④ other?
⑤ either spying not paternalistic
⑥ or spying redefines paternalism, cf. positive and negative freedom (freedom to vs. freedom from)
⑦ positive and negative paternalism.
Negative paternalism: any strategy by an agent A (the *paternaliser*) that is
① meant to check that a negative behaviour of an agent B (the *paternalised*) does not occur
② by monitoring B’s exercise of freedom,
③ for B’s own good,
④ without B’s consent.

Paternalism assumes that:
A knows better than B what is better for B and can shape B’s behaviour.
5) Check going from $S_x$ to $S_y$ and the not-going from $S_x$ to $S_z$.

Negative paternalism.

Incompatible with Freedom/Toleration.

Compatible with shaping behaviour.
CONCLUSION

Tolerant paternalism is possible. So objection against its adoption as a design principle removed.
Paternalism is not a behaviour/action, but a way of behaving/acting: paternalistically.
Second order property of relations.
Trust is also a second order property.
If paternalism then not trust, if trust, then no paternalism.
Paternalism has the effect of shaping PD’s behaviour or actions, but its definition is the contrary of trust. So ultimately it is a matter of control.
Lack of trust is paternalism’s opportunity cost.
ACKNOWLEDGEMENTS. This research was supported by an AHRC grant.

COPYRIGHT DISCLAIMER. Texts, marks, logos, names, graphics, images, photographs, illustrations, artwork, audio clips, video clips, and software copyrighted by their respective owners are used on these slides for non-commercial, educational and personal purposes only. Use of any copyrighted material is not authorized without the written consent of the copyright holder. Every effort has been made to respect the copyrights of other parties. If you believe that your copyright has been misused, please direct your correspondence to: luciano.floridi@oii.ox.ac.uk stating your position and I shall endeavour to correct any misuse.