



Amplifying positive bias: a case for positive-biases in children-AI relations

Dr. Karolina La Fors
Post-doc researcher in Responsible Design
DesignLab, University of Twente
k.lafors@utwente.nl

Outline

Problem description as to bias

Challenges to mitigate negative & amplify positive biases

Opportunities to mitigate negative & amplify positive biases

Recommendations for how in children-AI relations

Why distinct between
positive (pro-diversity) &
negative (discriminatory) biases
in children-AI relations?

Artificial Intelligence (AI)

Following Russel & Norvig:

AI systems are *intelligent agents that are able to communicate with reasoning*. (2020)

Network of systems & humans (Latour, 1987)

Children can consider AI as moral authority, even if no (wish for) legal personhood (Wynsberghe, Robbins, 2019).

Developmental Psychology

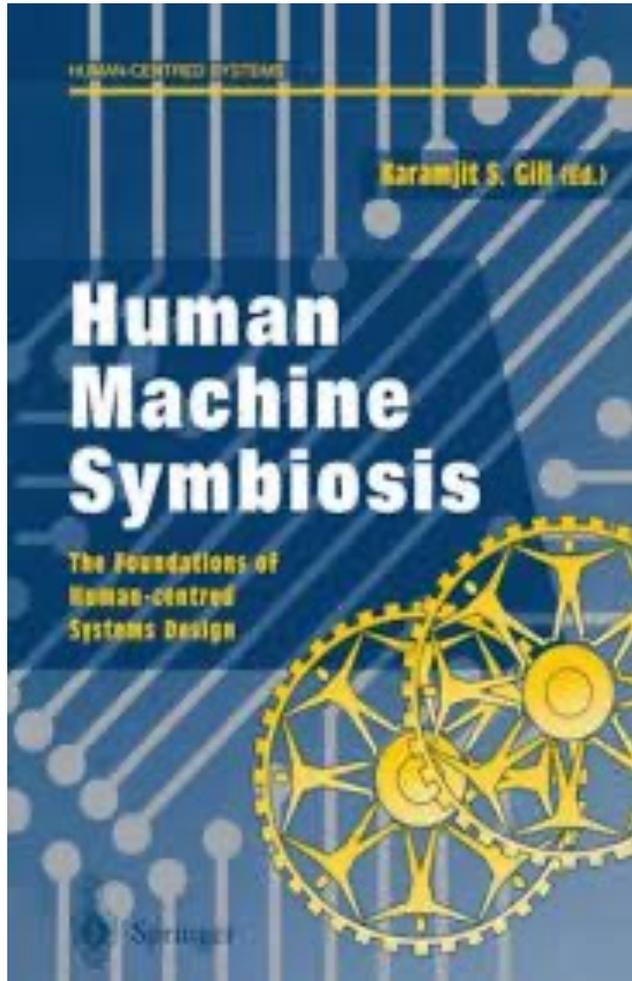
Discriminatory bias most profound detrimental, mental health impacts on children (Anderson, K. F.; 2013)

Identity, diversity and inclusion through education most profound effect to cultivate children's well-being (Keys Adair, J.; 2011)

AI Ethics and Responsible AI discourse

Scarce on how to amplify identity, diversity and more inclusion against AI-mediated discrimination of children

Human-Centred AI



1. Human-Machine (HM) value alignment
2. Human-Human (HH) value alignment
(Gill, K.; 1996)

Currently **larger emphasis**
on HM then on HH value alignment
Both are needed!

Challenges & Opportunities?

1. Challenge

Bias in AI ethics & responsible AI discourse is only negative

Confirmation bias (Mitchell, M., 1997; O'Neill, 2017);
Selectional bias (Enzenbach, Wicklein, Wirkner, 2019);
Allocative bias (Barocas, Crawford, Shapiro, Wallace, 2017);
Representational bias (Barocas, Crawford, Shapiro, Wallace, 2017);
...

From children's views:

Essential but not enough to ONLY focus on negative biases...

2. Challenge

Legal mitigation mechanisms of negative bias

Emphasis on

- online discrimination against adults rather than on children (UNICEF, 2020)
- children's right to privacy (Art. 16, UNCRC)

e.g.: *S. and Marper v. UK*, App. No. 30562/04 and 30566/04 [2009] 48 EHRR 50; *X. and Y. v. Netherlands*, App. No. 8978/80 [1985] 8 EHRR 235

Less on

- children's rights to identity (Art. 8, UNCRC); right to non-discrimination (Art. 2, UNCRC):
- ECtHR court cases on non-discrimination against children are from offline contexts

e.g.: *Oršuš and Others v. Croatia* [GC] App. no.15766/03 (ECtHR 2010); *D.H. and Others v. the Czech Republic* [GC], App. no. 57325/00, ECHR 2007-IV

How about?

- Enforcability of **UNCRC General Comment 25?**
- EAIA promising for children: *'reasonably foreseeable misuse'* :
 - 1) *redefinition of high-risk AI from children's perspectives?*

From children's views: **Awaiting court cases too slow**

3. Challenge

AI systems unexplainable from children's perspective

Proliferation of AI devices in general

Growth in unsupervised learning algorithms

Circular life cycle and context-overarching character of big data (La Fors, Custer, Keymolen, 2019)

Complicates filtering out children's data and preventing discriminatory effects

From children's views:

Improve explainability (can it/should it/if not, what then?)

Elicit values for children in contexts (e.g.: forgiveness in care/law enforcement; La Fors, 2020)

1. Opportunity

Implementation of (Dutch) Code of Children Rights

Specifying when & in which context desirable to invite children for co-design

Mapping AI parameters for children's pro-diversity mentality

Learning from reactions to the UK Code of Children's Rights (since yesterday official)

2. Opportunity

Input from children, parents & members of quadruple helix

EU HLEG Guidelines on Trustworthy AI (“diversity, non-discrimination and fairness” (EC, 2019)

UNICEF Policy Guidance on AI and children (UNICEF, 2020)

Inform the Proposed European AI Act

3. Opportunity

For Responsible *Humans In The Loop & AI*:
How to 'train' towards a 'benefit of the doubt'

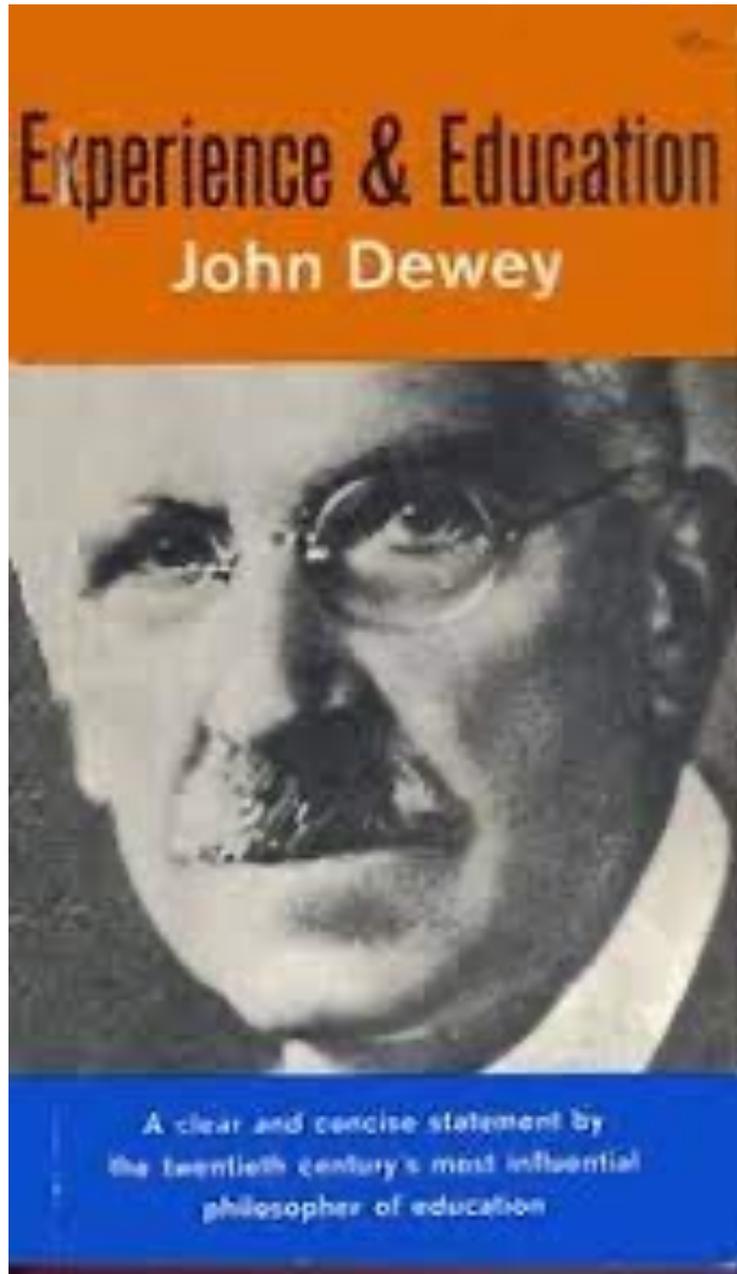
Value-sensitive design for AI (Umbrello, Van der Poel, 2021)

Cultivation of identity and respect for diversity from childhood onwards

Recommendations: Open Educational & Co-creational Spaces

How?

HH + HM value alignment methods



“Education as growth (being an end)”

Dewey critiqued:

“immaturity (child-) vs. maturity (adulthood)”

Dewey: immaturity is a positive characteristic

Why?

- 1) Immature’s dependence on others
- 2) Immature’s capacity to learn from experience (plasticity) => conditions for growth
- 3) Purpose of education & humans is growth
positive biases requirement for growth

A Human-Human value alignment and method



“The laws of morals and the laws of music are the same.” Zoltán Kodály

- Ethnomusicologist, philosopher
- Composer, developer of music educational method (since 2017: UNESCO world heritage)
- Educator of individual & collective identity for children
- Main motivation: **public debate is not broad & early enough**
- Also true for Responsible AI & AI Ethics

A Human-Machine value alignment method

GUIDANCE ETHICS APPROACH



Guidance ethics approach

By Peter-Paul Verbeek and Daniël Tijink

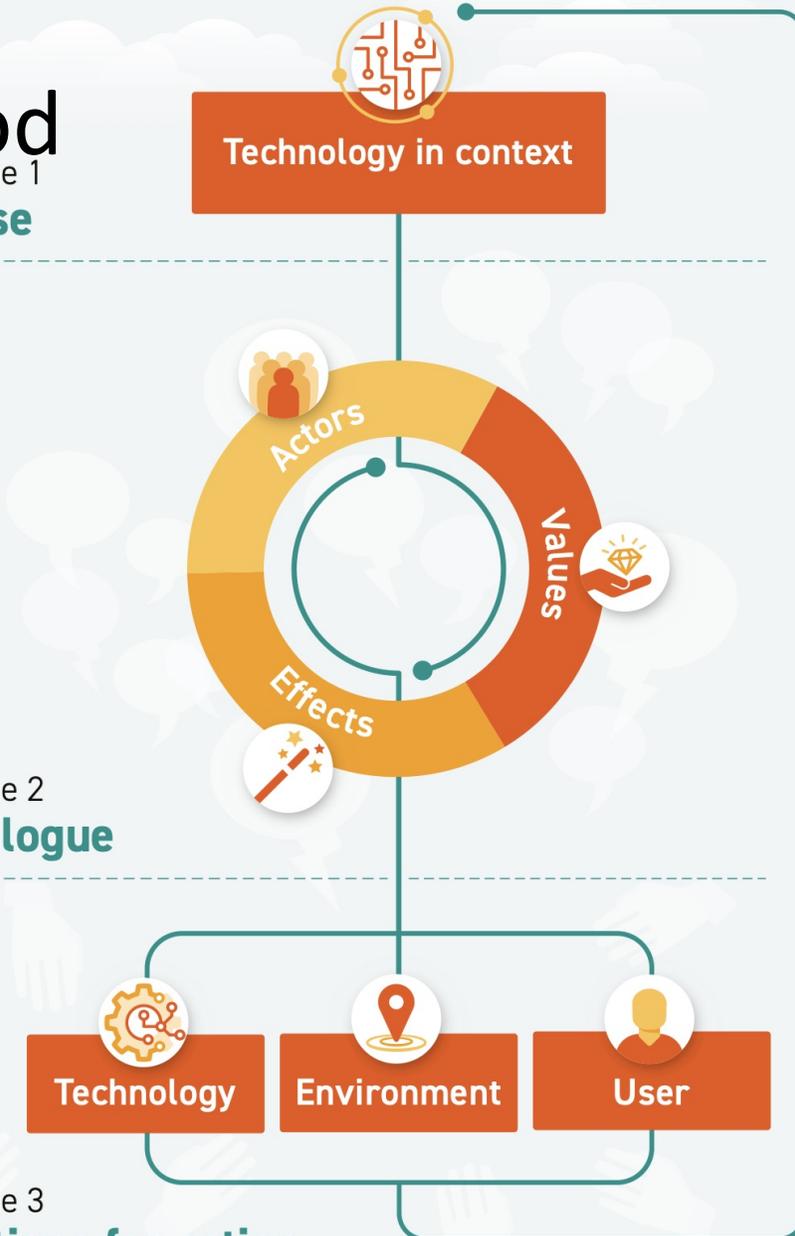
An ethical dialogue about technology with perspective on actions

Stage 1
Case

Technology in context

Stage 2
Dialogue

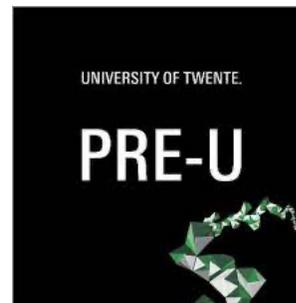
Stage 3
Options for action



ELSALab (Ethical, Legal & Societal Aspects of AI)

Initiative for *AI4Youth ELSALab*

just launched by DesignLab in The Netherlands:





Thank you very much!

Questions