World Wellbeing Panel

February 2025

Artificial intelligence and wellbeing

Artificial intelligence and wellbeing

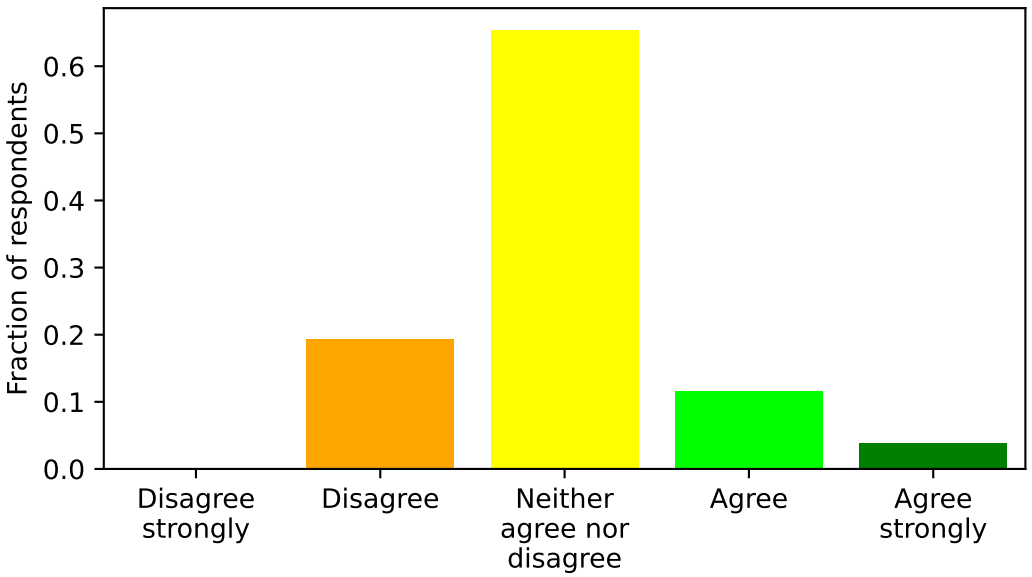
Report by Chris Barrington-Leigh (McGill University)
with Paul Frijters
(Academia Libera Mentis; and London School of Economics, visiting),
Tony Beatton (Queensland University of Technology), and
Arthur Grimes (Victoria University of Wellington)

In February 2025, members of the World Wellbeing Panel were asked for their views on two statements relating to the impact of innovations in artificial intelligence (AI) on human wellbeing.

Statement *1:*

Advances in artificial intelligence since 2000 have, overall, improved human wellbeing.

We received responses distributed as follows:



The large majority of respondents (17 of 26) were in the middle, choosing "Neither agree nor disagree", yet most respondents had plenty to say. Some who were on the fence nevertheless focused on the downsides of AI (Heffetz, Binder, Fabian, Helliwell,

Sarracino, Smith). Five respondents chose Disagree, while only four replied Agree or Completely Agree, making the answers almost perfectly balanced.

Enhanced productivity of bad things was the main fear: AI was predicted to increase crime, warfare, fraud, marketing, surveillance, propaganda, pornography, and manipulation through social media. In this direction, Rojas and Lepinteur pointed out that currently AI is created for profit rather than being designed specifically to support psychological wellbeing. Relatedly, Barrington-Leigh sees the "dark side of capitalism" being boosted by AI. Feared negative long-run impacts on the social and political system included reducing human intelligence, creativity, and connections, whilst increasing inequality and the concentration of social power (Ferrer-i-Carbonnell, Mahadea, O'Connor). Some mentioned the ecological cost of AI technology (Fabian, Andrén).

Given the broadness of the question, respondents had to define AI themselves. It was considered by most to encompass machine learning technologies far older than the recent large language model revolution. AI was thus given credit for learning (first at Facebook) that divisiveness increases engagement on social media (Sarracino). Social media algorithms were mentioned frequently, for instance for worsened youth mental health, eating disorders, and lower autonomy (Andrén, Smith).

Interestingly, some cited personal experience in their academic work as evidence of negative effects of AI: the recent homogeneity of style in research grants (Ferrer-i-Carbonell), automatically generated theses and research paper reviews (Binder), or the unwanted summaries at the top of Google search results (Fabian).

Commonly cited benefits from AI related to productivity and economic growth, drug and vaccine discovery, personalized education (Rojas, Mahadea, Jantsch, Lepinteur, Greyling, Andrén, Rossouw), research tools and data (Greyling), traffic management (Rojas, Andrén), and agriculture (Rossouw).

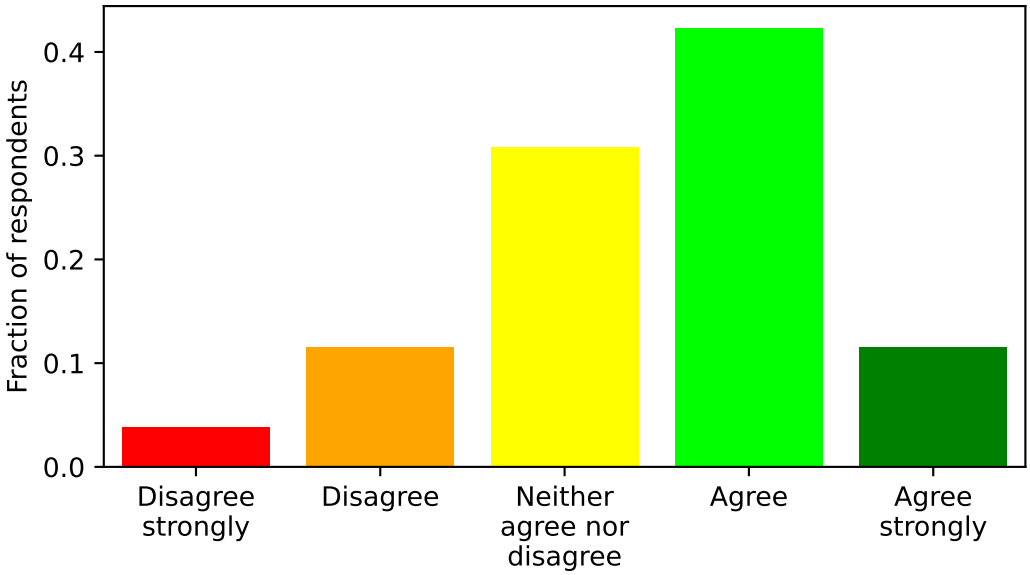
A number of respondents expressed the belief that AI is already in wide use in health care and diagnosis, beyond roles in drug discovery (Hendriks, Proto, Lepinteur), along with offering private personal advice in support of physical health. Chatbots focused on mental health support and promotion were also cited as a positive application (Andrén), even while the negative mental health impacts of AI were prominent in many responses.

While AI almost surely increases productivity in many places, Frijters, Foster, and Grimes mentioned that consumer convenience plays a relatively small role in human wellbeing, in comparison to the quality of human relationships of various kinds, thus again stressing that it is the negative effects on the social system that drive overall impacts.

Statement *2:*

Statement 2: Strong government regulation of AI is critical to ensuring positive future impacts on wellbeing.

We received the following responses:



For this question, the Agree (11) and Completely Agree (3) camp held the majority, though eight were on the fence again. Only four disagreed.

Market failures were cited often as a rationale for regulation. Rojas argues that in the domain of AI, assumptions of consumer sovereignty, wellbeing-enhancing individual choice, and the association between profit and consumer surplus, are all invalid in light of the enormous asymmetry of information and the complexity of the product. O'Connor and Smith both point out that due to network effects, AI provision is also not a naturally competitive market and may lead to concentration over time. Moreover, among those few most successful AIs, the private entities will not be incentivized to support well-being (O'Connor).

Jantsch suggests that government steering of AI may be more feasible if it is framed positively. That is, the objective could be to ensure that AI secures and expands human capabilities and opportunities, rather than framing it as defense against a threat.

Notwithstanding the market failures, Foster suggests that governments do not know better than markets, and prefers to rely on societies figuring out positive uses of AI without a role for government.

Regardless of the rationale or need for regulation of AI, many respondents did not believe that it is feasible. Many cited the rapid rate of technological change, implying that most regulation will be solving problems of the past (Binder, Grimes). "It is too late" wrote Ferrer-i-Carbonell. Also, the global technology landscape makes regulatory success dependent on an unlikely degree of international cooperation (Lepinteur, Frijters).

Not all respondents gave examples of what kinds of regulations they had in mind to achieve the ends for which they advocated, or which they feared as overreach. In fact, while Heffetz and Fabian asked what regulation would look like, only one respondent gave any example of a specific policy prescription: Barrington-Leigh advocates for a ban on any advertising that targets individuals using profiles built from their past activity.

List of Contributors:

Professor Ori Heffetz
Professor Mark Fabian
Professor Mohsen Joshanloo
Professor David Blanchflower
Professor Martin Binder
Professor Mariano Rojas
Professor Gigi Foster
Doctor Tony Beatton
Professor Daniel Benjamin
Professor Paul Frijters
Professor Martijn Hendriks
Professor Eugenio Proto
Professor Arthur Grimes
Professor Ada Ferrer-i-Carbonell
Doctor Kelsey J O'Connor
Professor Darma Mahadea
Doctor Antje Jantsch
Doctor Anthony Lepinteur
Professor Jan Delhey
Doctor Francesco Sarracino
Professor Talita Greyling
Professor Stephanié Rossouw
Doctor Conal Smith
Professor John Helliwell
Professor Daniela Andrén
Professor Chris Barrington-Leigh

For any questions, concerns, or comments related to this survey, please contact:

Chris Barrington-Leigh
Director, *World Wellbeing Panel*
Associate Professor, McGill University
Contact info at:
`https://wellbeing.research.mcgill.ca`