AI, virtual reality and robots are the future of behavioural public policy

Even though technology is largely underused by policymakers, the behavioural public policy of the future will be high tech. Artificial intelligence, virtual reality and robots hold the potential to help people change their behaviour for the better. **Dario Krpan, Fatima Koaik, Pujen Shrestha** *and* **Robin Schnider** *envision a future in which policymakers use technological interventions via various personal devices and appliances, from virtual reality glasses to IoT devices.*

The year we've just witnessed has been nothing short of a whirlwind in the realm of technological advancements. One of the most noteworthy developments has been in the field of artificial intelligence (AI), where models like GPT-4 have reached new heights, pushing the boundaries of what is possible. Many of these cutting-edge technologies hold immense potential in helping people change their behaviour for the better, and yet they remain underused by policymakers. Here we explore technological developments that have the largest potential to reshape behavioural public policy and examine how they can be used by policymakers to implement or test behavioural interventions.

Virtual reality

Many important human-created issues affecting the world, from climate change to conflicts, can be explained by two psychological barriers:

> 1. People cannot teleport themselves into the future to directly experience the long-term consequences of their present actions.

2. People cannot easily put themselves in someone else's shoes to experience how their actions affect others and make them feel.

Virtual reality can tackle these challenges. It can be used to transport people into the future (to virtually recreate the future in which the consequences of their actions can be experienced) and to "embody" people as someone else (to create a virtual body that

people can operate and from whose perspective they can see the world). Research showed that these types of interventions can have various positive consequences. For example, a white person being embodied as black in virtual reality [decreased implicit racial bias](#).

For these reasons, virtual reality is a medium that policymakers should explore to create behavioural interventions that have a potential to tackle the world's biggest challenges. Beyond its promising effectiveness, the technology is ripe for policy use because virtual reality glasses have become [widely available](#) and are more affordable than ever. Given these trends, it will be important for policymakers to form partnerships with companies that make virtual reality apps to create powerful behavioural science interventions for the future.

Robots as messengers

It is estimated that humans and robots will live together in a not-so-far-away future. In fact, [NEOM](#) is already building cognitive cities where robots and humans are expected to coexist in harmony. Robots represent a large opportunity for behavioural public policy because they can be used as messengers to help people accomplish various important goals. For example, robots can prompt individuals to exercise, warn them to avoid eating unhealthy snacks, or give them personalised feedback in any situation.

One advantage that robots have over human messengers is that they may be less likely to incite psychological reactance, which refers to feelings of anger that people can experience if they feel that someone is trying to take away their freedom by telling them what to do. Preliminary [research](#) indicates that reactance is typically instigated by various social cues that people use in communication (facial expressions, eye contact, gestures, etc) but that may be absent or present to a lesser degree in certain robots. Because robots can be designed in an infinite number of ways, it is plausible that they can be built to evoke minimal resistance and easily motivate humans to achieve their goals. To build behavioural interventions for the future, policymakers should start collaborating with roboticists to develop interventions that robotic messengers can communicate to humans depending on their goals and preferences.

Behavioural informatics

Behavioural informatics refers to the application of the internet of things (IoT) devices (such as smart home appliances and wearable devices) for the purpose of behavioural change. Currently, it is estimated that around 15 billion IoT devices exist, which means that on average there are two such devices per each person in the world. These devices have a significant potential for behavioural change because many people have them and because they can generate large amounts of data to produce various behavioural insights. For example, IoT devices can capture real-time emotional and motivational states of the person, provide insights about the barriers that prevent the individual from undertaking certain behaviours, and so on. This knowledge can be used for timely behavioural interventions in numerous ways. For example, if the relevant IoT devices show that the person is in an energetic state, the person can be nudged to exercise. If the devices show that the person is in a sad mood, positive content can be provided to cheer them up. These are only some examples, and infinite other options are possible.

There are currently two barriers to using the IoT devices for the purpose of behavioural change. First, this requires building complex algorithms that can infer behavioural insights from the data and analyse these insights to provide timely interventions. Second, collecting such data may invade people's privacy and in some cases clash with laws and regulations from different countries. However, behavioural informatics is the future of behavioural public policy. Capturing real-time data and automatically transforming it into timely interventions, assuming this is aligned with the person's preferences and privacy recommendations, is a hyper-personalized approach that goes beyond the traditional "one size fits all" policymaking. The task of policymakers will be to collaborate with both computer scientists and regulators to build such interventions in an ethical way that is consistent with best practices and regulations.

Personal behavioural assistants

For citizens who want to change their behaviour to achieve personal goals, it would probably be most useful to have a personal behavioural science assistant whom they can constantly ask for advice. Policymakers may not have the resources to act in such a capacity. In this context, large language models have reached a stage of development in which they can be trained on numerous behavioural science articles and reports to provide evidence-based advice that can help people who want to transform their lives.

Initiatives that could be adopted for this purpose are already under way. For example, scientists collaborating on the Human Behaviour-Change Project (HBCP) have developed "an online 'knowledge system' that uses artificial intelligence, in particular natural language processing and machine learning, to extract information from intervention evaluation reports to answer key questions about the evidence." Various platforms are already commercially available. One example is scite.ai, which can answer a range of questions, including ones about human behaviour, by referring to the relevant scientific literature. The main challenge of such platforms is that in some cases they still tend to misinterpret the references they consult and provide inaccurate advice. In this context, developing specialised platforms that provide behavioural change advice will require introducing "safety mechanisms" that will ensure that users critically evaluate what they read before they implement it. However, such platforms can be hugely helpful to policymakers because they can be used to provide personal support to citizens who may otherwise struggle to change.

Testing the impact of policy interventions

What if the impact of behavioural interventions could be tested on LLMs such as ChatGPT, instead of on human participants? This would allow policymakers to develop evidence-based policies even in cases when it is not possible to conduct field studies for various practical or other reasons. Recent research has shown that testing policies and policy interventions without the need for human participants may soon become possible, considering that LLMs and humans have similar psychological characteristics. For example, researchers asked GPT-3.5 to judge the ethics of various scenarios that were previously appraised by humans (such as having an affair with one's best friend's spouse). The responses by GPT-3.5 and humans were almost identical and had an unusually strong correlation of 0.95.

When used as research participants, LLMs can improve diversity because they can be prompted to play different roles. For example, they can be asked to respond from the perspective of minority groups members, people from specific age groups, and so on. However, it is important to know that LLMs have a certain disadvantage when asked to act as research participants – their responses are most similar to individuals from Western, Educated, Industrialised, Rich, and Democratic (WEIRD) countries because many of these models were trained on their data. Therefore, LLMs are not yet a perfect

substitute for human participants when it comes to testing behavioural policy interventions, but they have an enormous potential to be used for this purpose. Policymakers should focus on perfecting these models and collaborating with model builders. If models are trained with data from non-WEIRD countries, it may soon be possible to seamlessly test many interventions to develop more effective policies. This will usher a new age of behavioural public policy.

Final word

The behavioural public policy of the future will ultimately be driven by new developments in technology. In this blog post, we have outlined a vision in which the policymakers of the future will develop technological interventions for various personal devices and appliances, from virtual reality glasses and robots to IoT devices. In doing so, they will not need to test these interventions on actual humans but on LLMs that are psychologically like humans. People will choose which of these devices and interventions to use to change their behaviour for the better, accompanied by LLMs as their personal behavioural assistants.

- *The post represents the views of its author(s), not the position of LSE Business Review or the London School of Economics.*
- *Featured image provided by Shutterstock*
- *When you leave a comment, you're agreeing to our Comment Policy.*

Dario Krpan is an Assistant Professor of Behavioural Science at the Department of Psychological and Behavioural Science, LSE.

Fatima Koaik is a Behavioural Science Lead at the Ideation Center, Strategy& Middle East, PwC, and a PhD Candidate at the Department of Psychological and Behavioural Science, LSE.

Pujen Shrestha is a Behavioural Economist and Associate at the Ideation Center,

Strategy& Middle East, PwC.

Robin Schnider is a Behavioural Economist and Senior Associate at the Ideation Center, Strategy& Middle East, PwC.