Immanuel Barshi

IN, ON, AND OUT OF THE LOOP

As automation evolves, we need to ensure that those overseeing critical systems can effectively transition from monitoring to taking control when it matters most. But some designs do not make obvious the need to intervene. Immanuel Barshi explores the quandary.

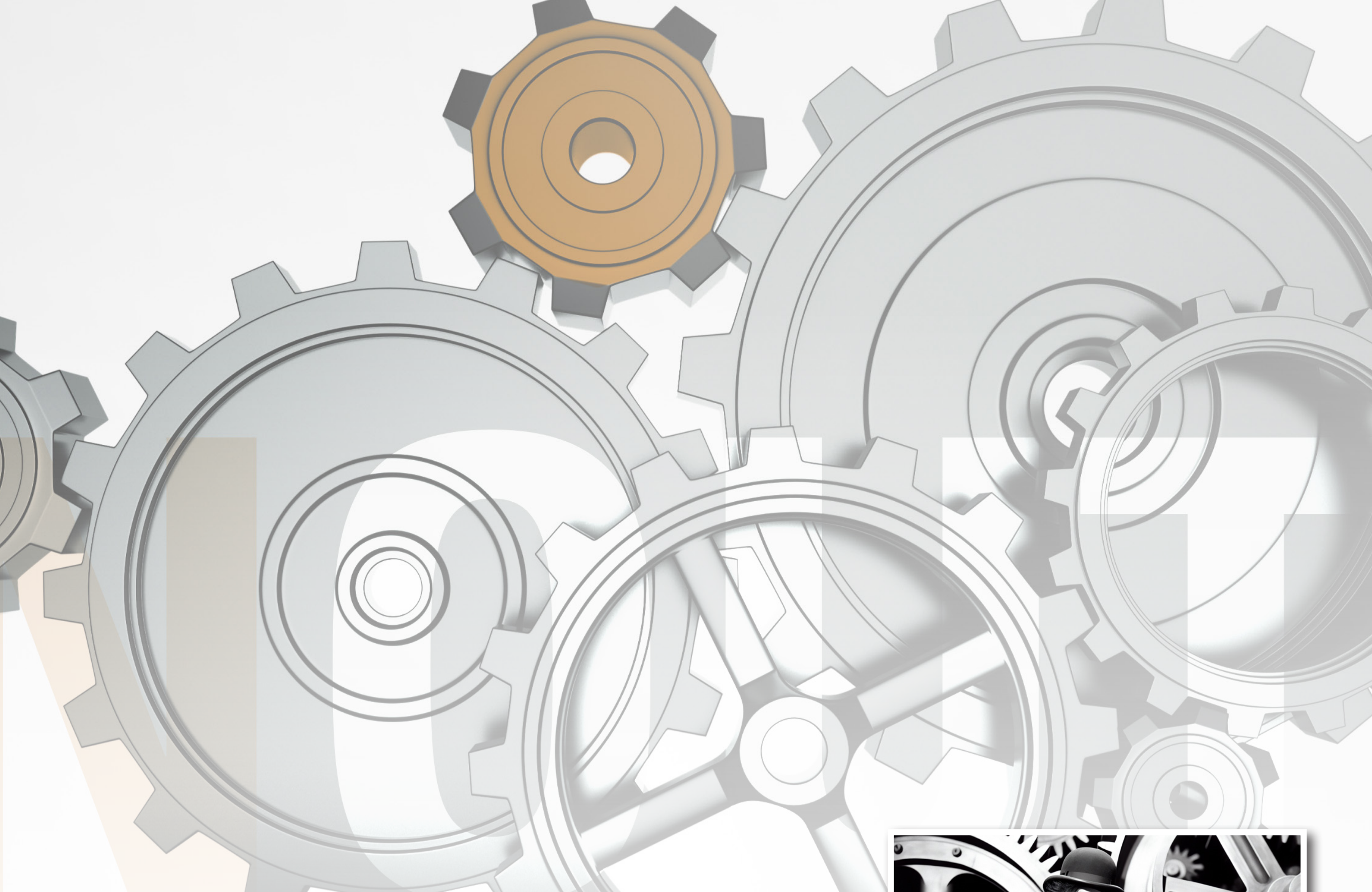
KEY POINTS

- **In the loop versus on the loop:** There is a difference between being 'in the loop', where a person has direct control over a machine, and being 'on the loop', where they supervise the automation without direct control. Effective supervision requires readiness to intervene if the automation fails.

- **Automation surprises:** Automation can lead to 'automation surprises', where systems behave in unexpected ways. These surprises can have serious consequences, as seen in historical aviation incidents.

- **Feedback:** There is a need for automated systems to provide clear indications of their states, transitions, and any potential issues. Systems that fail to communicate these effectively can lead to dangerous situations.

- **Training:** Controllers and pilots must receive thorough training to understand the logic and potential pitfalls of automated systems, and to be ready to step in and take control when automation does not perform as expected.



In his 1936 black and white part-talkie film Modern Times, Charlie Chaplin is struggling to stay in the loop with the new automation installed at the factory. He is constantly thrown out of the loop, and in some cases remains in the loop, in some respects, even after walking away from the production line. He is like an air traffic controller who continues to play the traffic in her head after the shift is over to figure out what could have been done better. It seems that not much has changed since the early introduction of automation. One small change did happen: Charlie Chaplin didn't have the experience of being on the loop.

Charlie Chaplin in Modern Times (AI generated)

We talk about being in the loop when we have direct control of the machine. Driving a manual transmission car in busy stop-and-go traffic during rush hour in the city keeps us in the loop of controlling the car. Executing a series of aerobatic manoeuvres in a 1945 Pitts Special keeps us in the loop of controlling the biplane. And constantly giving takeoff and landing clearances at a busy airport keeps us in the loop of controlling the traffic. We know exactly where every aircraft is and where we want it to go, and we keep monitoring to make sure it gets there.

> "We talk about being in the loop when we have direct control of the machine."

We talk about being out of the loop when we have no control of the machine, or even when we have no feedback on how the automated controller is managing the machine. We fill the coffee maker's reservoir with water, make sure there are coffee beans in the hopper, place a cup under the brewing head, set the timer to have a cup of coffee ready when we wake up the next morning, and go off to bed for the night's sleep. We have no direct control of the machine's internal algorithm. Hopefully, the smell of fresh coffee drifts into the bedroom just as the alarm clock goes off to wake us up. We have been sleeping peacefully out of the coffee-making loop all night. Some Tesla drivers have been sleeping peacefully out of the car-driving loop travelling at a constant distance behind the car in front of them, according to the setting of their adaptive cruise control, while their lane-centring algorithm gently follows the curves in the road, careful not to wake them up.

> "We talk about being out of the loop when we have no control of the machine, or even when we have no feedback."

Similarly, airline dispatchers increasingly rely on neural network or deep learning programs that spit out flight plans based on optimal winds, aircraft weight, temperature, traffic density, landing slot time, crew duty periods, passenger connection times, and many other variables. This is mostly opaque to the dispatcher and places them out of the loop.
When we know how to work the machine, staying in the loop is easy (if the machine is designed to allow us to do that).

When we don't need to care about the machine's behaviour, staying out of the loop is easy.

Charlie Chaplin may not have known it, but it was a good choice not to work into his film the experience of being on the loop. That's the tricky one.

We talk about being on the loop when we assume a supervisory role, when our only job is to make sure the automation does what it was designed to do. Of course, there is a minor catch to that seeming luxury. If the automation fails in some way, we need to be able to jump in and intervene so the operation can continue to flow smoothly.

> "We talk about being on the loop when we assume a supervisory role, when our only job is to make sure the automation does what it was designed to do."

Imagine having to supervise the coffee maker through the night. Chances are, you won't get much sleep. If the coffee-making machine was designed to be supervised, it would have to have some indication letting you know what state it is in. You would have to be trained to know the precise sequences of states it can go through in the process of making the coffee so you'd be able to predict when it is likely to transition to the next state and under what conditions it might fail to do so. You might be able to set your alarm clock to wake you up just before a transition takes place so you can confirm that it did it right. You might even suggest to the designers of the coffee maker to integrate such an alarm clock into the design of the coffee maker. In fact, the designers of aircraft's flight management systems (FMS) tried to do exactly that when they envisioned the role of the pilot as one of being on the loop rather than in the loop.
The FMS uses the FMA (flight mode annunciator) to indicate to the pilot the state that the aircraft automation is in, and in some cases when transitions occur between states (e.g., 'altitude capture' as the aircraft transitions from a climb to level flight at the assigned altitude). Airlines invest tremendous effort training pilots to understand and to know the sequences of states involved in flight and the transitions between them so pilots can successfully supervise the automation to maintain efficient and safe flights. As we know, that doesn't always happen.

For instance, the downing of Korean Airline flight 007 in 1983 was likely the result, in part, of the pilots not recognising the fact that their Boeing 747 remained in 'heading mode' and did not transition to 'navigation mode'. Unfortunately, this 'automation surprise' ended up in a far more serious surprise (see Degani, 2004, for an analysis of this accident as well as similar problems with various machines and different modes of transportation).

The aviation research literature is filled with discussions of such 'automation surprises'. These are situations in which the automation behaves in ways the operator – or rather, supervisor – did not expect and did not anticipate. From the early days of aviation's advanced automation (see, e.g., Wiener, 1989) through the mid-90s (see, e.g., Sarter and Woods, 1995) and all the way to the very present (see, e.g., Dekker and Woods, 2024), we continue to struggle with the supervisory role. In particular, we struggle with the timely transition from being on the loop to being in the loop.
Dekker and Woods (2024) describe some automation as being "strong, silent, and wrong". Their work is an important call for designers to at least eliminate the "silent" part, if they design the automation to be strong, recognising that we can't eliminate all the possibilities of the automation being wrong. It is exactly when the machine is silent that we don't know when it's time for us to intervene.

With new technologies, with the push for autonomous vehicles, and with the growing number of semi-autonomous features in cars, we see a growing research literature on the need for drivers to transition from being on the loop to being in the loop (see, e.g., Merat et al., 2014). The idea of truly autonomous vehicles is to keep the driver – sorry, passenger – completely out of the loop, but there is a growing recognition that we may not be as close to it as some people would like to believe. Meanwhile, car manufacturers

integrate various systems to alert the driver to stay closer to the loop, if not in the loop. Some cars have sensors in the steering wheel to make sure the driver's hands are on the wheel even if the lane-keeping and lane-centring features are engaged. Some cars have eye-tracking cameras that alert the driver if their eyes are off the road for more than a few seconds, or if their eyelids droop as a sign of possible sleepiness. But it's unlikely that ATC equipment manufacturers will integrate such features into their designs, and the increasing popularity of AI algorithms only raises more challenges (see, e.g., Mazzolin, 2020). And so the struggle to stay close to the loop while on the loop, and to know when to jump in and intervene, will remain with us for the foreseeable future. So what can we do?

On the design side, we need our systems not to be silent. We want clear indications of states, of impending transitions, and of any situations in which the conditions necessary for such transitions do not obtain. We'd like our systems to 'communicate their intentions' so we know what to expect and are not surprised.

On the controller's side – in the absence of better design – we must invest in good training and good support materials that enable controllers to understand their systems, including the logic and rationale of their design. This should help controllers to know which machine state is the right state for a given traffic and environmental situation, and to be prepared to jump in on short notice if anything looks out of place. To be able to do that, one must know what's the right place for every piece of the current ATC picture. And one must maintain a state of mind I call 'proper paranoia', recognising that despite the very high reliability of the machines and automated systems we work with, they may behave in an unexpected way any minute and with little, if any, warning.

By building systems that are sensitive to the challenges of holding operators on the loop, designers can support us better. By maintaining proper paranoia, controllers and other frontline operators are better able to hold that on the loop position. As the accident record shows, the space between in the loop and out of the loop is fertile ground for us all to better ourselves and our systems for the sake of safety.

REFERENCES

Degani, A. (2004). Taming HAL: Designing interfaces beyond 2001. Palgrave-McMillan.

Dekker, S. W. A., & Woods, D. D. (2024). Wrong, strong, and silent: What happens when automated systems with high autonomy and high authority misbehave? Journal of Cognitive Engineering and Decision Making, 0(0).

Mazzolin, R. (2020). Artificial intelligence and keeping humans "in the loop." Modern Conflicts and Artificial Intelligence. Center for International Governance Innovation. https://www.cigionline.org/articles/artificial-intelligence-and-keeping-humans-loop/

Merat, N. A., Jamson, H., Lai, F. C. H., Daly, M., & Oliver M. J. &, Carsten, O. M. J. (2014). Transition to manual: Driver behaviour when resuming control from a highly automated vehicle. Transportation Research Part F, 27, 274–282

Sarter, N. B., Woods, D. D. (1995). "How in the world did we get into that mode?" Mode error and awareness in supervisory control. Human Factors: The Journal of the Human Factors and Ergonomics Society, 37(1), 5–19.

Wiener, E. L. (1989). Human factors of advanced technology ("glass cockpit") transport aircraft. NASA Contractor Report, No. 177528. Moffett Field, CA: NASA Ames Research Center.

DR. IMMANUEL BARSHI is a Senior Principal Investigator in the Human Systems Integration Division at NASA Ames Research Center. He studies the skilled performance of astronauts and pilots, mission controllers and air traffic controllers, their ability to manage challenging situations, and their vulnerability to error. He holds Airline Transport Pilot certificate with A320, A330, B737, and CE500 Type Ratings, and is a certified flight instructor for airplanes and helicopters.