HELEN MACNAMARA
ALEX CHISHOLM

From: DAVID HALPERN
cc NR

Date: 28 September 2020
July 2020 (orig. draft)

INSTITUTIONAL LESSONS FROM COVID

You asked for some thoughts on the lessons from HMG handling of COVID. This note is a personal view, albeit situated in previous work on institutional weaknesses and failures, including *Behavioural Government*[1], King and Crewe's *The Blunders of our Governments,* and Eggers' *If We Can Put a Man on the Moon….* The summary points are:

1. **The early misstep.** Overconfidence and anchoring in our expert medical community led to a presumption that covid would be a flu-like wave, blinding it to the pursuit of near-suppression as a viable option and an expanded tracing system in particular. Our decision-making process was vulnerable to systematic error.
2. **Don't put your eggs in one basket**. Where daring punts were made on single solutions, such as the first NHS app, they generally failed, whereas when multiple solutions were tried, eventual success was achieved.
3. **'How?'** People mistook academic enquiry for policy. We neglected the 'engineering' of effective delivery.
4. **'Authoritative' versus 'authoritarian'.** The policy and comms strategy sought to occupy a position of 'soft authoritarianism', based on simple, rigid rules yet with weak enforcement. Arguably, we would have done better going for an 'authoritative' approach: risk and principles-based guidance that could be flexed.
5. **Putting humpty together.** A strong bio-medical perspective was not effectively tempered with robust behavioural or economic analysis, and a weak central process failed to give adequately balanced advice to the Cabinet of PM.

There will be another covid. Alongside improvements to core decision making, we should use civil service reform and the SR to strengthen the empirical and methodological competence of the UK policy profession. I am also going to seek to build a UK 'social and behavioural MIT' - an world-class institution that can deliver timely advice and solutions to the standard we need.

The puzzle

The central puzzle is how the UK brought together a world class group of experts, alongside a world class civil service, that together led to such an un-world-class

[1] https://www.bi.team/publications/behavioural-government/

INQ000129093_0001

outcome. The UK remains in the 'top-10' worst national performers by per capita death rates, despite having had the benefit of a lagged start. Many other countries have achieved 10-fold, or even 100-fold, lower per capita death rates than the UK (614)[2] - such as Norway (48), Australia (35), S Korea (7), NZ (5), Singapore (5). There are many factors involved in these differences, but we should urgently learn lessons, not least in anticipation of responding to the secondary social and economic shocks to come.

The 'easy' response is to blame the politicians. But the failure was much more extensive, and more subtle. Weaknesses in the UK response fit with James Reason's 'swiss cheese' account of failure[3]: an alignment between multiple institutional and human weaknesses, as seen in historic failures from 'Three Mile Island' to 'Deepwater'. It is also important to examine the counterfactual - countries or parts of our own system that performed better. Reason et al called these 'highly reliable' organisational forms: highly complex organisations, such as aircraft carriers or air traffic control systems, that are subject to multiple and varied challenges yet consistently do not fail.

1. The early misstep: overconfidence and anchoring

Arguably the most fundamental misstep in the UK response was the presumption that covid would be an unstoppable flu-like wave. This presumption was built into the Contain-Delay-Mitigate-Research strategy published in early March. It also underpinned the early (Chris Whitty) position on test and trace, and the Vallance view on 'herd immunity' (later air-brushed).

It is important to see that this presumption was not based on ignorance, but on a century of prior knowledge and assumptions. SPI-M had sophisticated models drawing on data going back to the 1918 flu showing the dynamics of prior epidemics. Chris Whitty had prior field experience, and was extremely concerned to get the 'wave' landed before winter. Jonathan Van Tam wrote the textbook on flu, and had strong views on the likely modes of transmission and spread. Patrick Vallance knew that vaccines would not be available till the end of the year at the earliest.

All this massed expertise converged on the conclusion that, once early containment had failed, a flu-like wave was inevitable. As such, the best that could be done was 'to flatten the curve'. This would take the worst of the negative effects out by (a) preventing the NHS being overwhelmed, (b) stopping the epidemic over-swinging (ie achieving 'herd immunity and around 60% instead of 90%), and (c) protecting the most vulnerable until the wave had passed.

2

3 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1117770/

But there was an alternative. The world - and more specifically the South Asians - had found another strategy: to build a sophisticated test and trace system strong enough to substantially suppress the virus, at least for long enough to enable treatment and vaccines to be developed. This is not just a matter of hindsight bias. There were a number of us - generally outside the biomedical bubble - who highlighted this alternative strategy. BIT studied the responses and techniques employed by S Korea, Singapore and others from the outset, including the particular techniques they used to do the tracing, and which of these might be transferable to the UK context (eg phone tracing and notifications, using commercially available tracking). This led us to question the limited ambition of the central 'delay' phase, though we found ourselves quietly dismissed as not really understanding the science.

If there was a single top lesson to be taken from the 2018 Behavioural Government report, it was 'beware over-confidence'. Humans are fundamentally prone to overconfidence in their own beliefs, systems and groups. This overconfidence tends to become more serious the more senior the individual.

Experts and academics are prone to the same weakness. Despite a century of statistical methods to directly estimate confidence, the academic world remains characterised by overstatement and overclaiming. Indeed, whole careers are built on battles between rival academics overstating their cases in the pages of academic journals, with corresponding overestimates of effect sizes from publication bias.[4]

There were multiple examples of such 'anchoring' within SAGE, with early hypotheses and views stuck with, despite mounting evidence against. Early evidence of low german death rates was repeatedly dismissed as being the result of cases occurring in 'fit skiers', and that their death rates would shortly converge on the UK numbers. The role of aerosol transmission and corresponding importance of ventilation was underplayed (relative to ballistic). Asymptomatic cases, and their policy implications, were understated. The growing evidence on masks was very slow to be taken on board.

At the same time, a proposal (supported by Matt H) to create an open wiki-style list of questions, or known unknowns, was quashed by PV. Instead PV asked Mark Walport at UKRI was asked to organise a website publishing existing research: a (delayed) shop-window of what we already knew, rather than a pinpointing of what we didn't. Ironically, the pride in our science and our capabilities, slowed our ability to learn lessons from other countries. Under cover of variations of 'it is very different there', there was an arrogance that we knew better, and would do better.

If I had to choose a single issue - and moment - that embodied this failure, I'd point to the unshakable conviction of the SPI-M modellers that suppression - the sustained holding down of covid prevalence - was not a viable strategy. For example, Graham

[4] https://eml.berkeley.edu/~sdellavi/wp/NudgeToScale2020-05-09.pdf

Medley, chair of SPI-M, was asked in SAGE in the second week of March, along with the other modellers: 'how certain they were that major second waves would arise in China and other Asian nations?' (ie that suppression was not a viable strategy). Medley answered immediately, and with total conviction: 'as close to 100% as possible...'yes, 100%'. His colleagues echoed the conclusion 'yes, 100%'. They were totally convinced that as soon as the harsh lock-downs in Wuhan, South Korea and elsewhere were lifted, cases would immediately surge again.

Nothing in science, and certainly in statistical modelling, is 100%. Let alone in the face of the data already emerging from the Asian experience. This was doctrine, not science.

Don't: get carried away with your own models (and ego)

Don't: shut down counter or heretical views, or seek an overly early 'single view of the truth'. Avoid having key meetings chaired by people with strong prior expertise in a particular area or specialism.

Do: use scenarios that ensure your policy is robust to a range of possibilities, and/or supplement with 'red-teaming' to challenge strong priors

Do: have key meetings chaired by intelligent generalists; seek to develop alternative viewpoints, and use mechanisms to keep questions on the table.

2. Don't put your eggs in one basket: diversity and innovability

Anchoring and rigidity was not limited to our core epidemiological assumptions. It also dogged our delivery. We generally succeeded when we set higher level functional goals and challenged innovators to deliver. We generally failed when we over-specified the objective and means to deliver it, and then bet the bank on a single solution.

The decision by NHSx to put all its eggs in one NHS App basket was a particularly grim example. When the idea was proposed for an app (March 6th) we, along with Ben Warner from No10, strongly supported it being taken forward, but also strongly recommended that this be done as part of a portfolio. We argued that more than one approach to the app should be developed (such as setting up 2 or three parallel teams, at least up to an early gateway); a re-skin of an existing app be attempted (eg Singapore, Australia); and non-app based approaches should also be pursued (eg SQREAM). NHSx not only pursued only one (internal) option, they persisted with this as deadline after deadline passed. This was a serious policy failure in its own right. A version of a tech-enabled app should, and could, have been delivered to be ready for the end of the first lockdown (when its value really would have kicked in). The app, delivered by a separate team after the first was shut, was delivered six months after the original target date.

In contrast, a relatively early and successful move - though highly contentious at the time - was to break the internal monopoly of PHE on the delivery of testing. This was kicked off with the No10 summit where we brought in multiple industry players and challenged them to deliver volume covid testing. While some of the tests failed to reach an adequate standard (notably the first batch of pinprick antigen tests), in general volume and quality was delivered across multiple platforms, and it has also led to the development of faster, more user-friendly, and low cost testing techniques.

If you are serious, the real test of this approach is whether you build it into your day-to-day practice. This means building in experimentation and deliberate variation (inc A/B formatting) into delivery. Even with the greatest minds and good intent, 'world class' isn't launched on day one, but is evolved and shaped through a relentless search for improvement.

A good example is how we tested variations in the early public health messages on samples of several thousand people, systematically enhancing levels of comprehension and intent to comply across interactions (as opposed to plumping for what a couple of designers thought looked cool). A recent, delivery based example is how we tested variations in approach within the test and trace system to follow-up with contacts to increase compliance. Alongside business as usual, we tested with several thousand cases the efficacy of (a) follow-up texts, (b) follow-up calls in person, and (c) a combination of texts and calls. The results were clear and had clear operational implications for the service. Calls boosted compliance. Texts, if anything, reduced compliance. Calls and texts were less effective than calls alone.

Unfortunately, such experimentation remains the exception, not the rule. If you are running at 200mph to deliver, it can seem a luxury to test variations. But if we want world class, or even top quartile, we have to build systems and practices that are designed to experiment, learn and get better.

Don't: put all your eggs in one basket, or prematurely shut down delivery options or questions

Do: promote and test multiple alternatives, and design systems with constituent elements as modular and interchangeable as possible so that as better alternatives are identified they can be swapped in with minimal cost and friction

3. 'How?': science versus policy and delivery

One of the very striking aspects of the early key phases of the covid crisis was the extent to which the traditional policy community became passive. Partly this was driven by the political mantra to 'follow the science'. It led to an extraordinary level of paralysis across operational areas and 'normal' civil service policymaking. Every question or

decision had to be passed up to SAGE and/or for clearance by CMO, literally down to the wording of posters.

As Chris Wormald remarked (in early March), but didn't seem to act on, SAGE was focused on 'what' not 'how'. We (BIT) saw this in our own line-by-line analysis of SPI-M models. These models were full of assumptions about what level of reduced contact different measures might lead to, and presumed levels of compliance. But the models themselves did not tell you how these levels of compliance or outcomes were going to be achieved: the guidance, operational changes, legal requirements, system capabilities, etc. It was 'abracadabra' delivery.

Testing again illustrates the situation. While the UK split hairs over the pros and cons of mathematical models, from how many lines of code the Imperial model had versus the London School or Warwick models, or the exact level of lagging in the ICT data, the Germans sharpened their already impressive testing systems. We did high science, they did engineering.

'Analysis paralysis' persisted deep into our operational practice. Important questions in the design of the T&T systems, and later JBC, were passed to the high priests of SAGE and SPI-M that they had no way of answering. These questions should have been answered within the T&T system. To give a concrete example, there were always key questions about where the T&T system should deploy its spare marginal capacity, particularly once the initial surge past 100k tests a day was achieved. For example, T&T's director of policy believed that it was worth testing more asymptomatic cases (that were judged of high risk), and other senior figures in the program believed that there was a strong case for exploring test and release strategies for people late in isolation. But there was a lack of clarity about who could sign off such testing or experimentation, so policy was instead referred for decisions by SAGE and CMO office. This created Kafka-esque loops. For example, on asymptomatic testing, the position was received back from SAGE and CMO office that such testing should not be done unless there was clear evidence that the incidence or prevalence was higher than in the general population. But the only, or at least fastest, way this could be determined was to conduct testing in that population…

Delivery was also dogged by HR issues - a carousel of people at both senior and junior level. We don't swap out our CSA every month or two, but we did exactly that across swathes of our delivery apparatus.

Don't: mistake discussions over aetiology for being policy or delivery - the 'how' matters greatly

Do: bring operational, economic and other policy-relevant issues into the room, and build and keep a talented delivery team who you are prepared to listen to and respect

4. 'Authoritative' versus 'authoritarian' (& local versus central)

One of the deep issues that the UK and other countries have had to wrestle with is the extent to which our covid policy, and communication, needs to be kept simple and uniform, versus more complex and varied - across geographical areas, time, and risk segments.

This was not an easy call. In general, the UK and the current No10 leaned towards keeping it simple, at least in the early phase. 'Stay Home' and the initial lock-down was simple and uniform. This broadly lined up with the evidence that the UK had been exposed to multiple seeding events and community transmission was relatively widespread. There was discussion about a more differentiated approach, such as London entering the lockdown early, but confusion of message and the Italian experience of displacement across boundaries were real and appropriate concerns.

There is a case, however, that the UK became stuck on a UK-style version of 'authoritarian' (in the sense of rigid, simple rules, with the implication of punishment for violations of those rules). In contrast, 'authoritative' means that there are a relatively small set of principle-based guidelines, around which there are explanations of why. This implies that instead of the person (or child) following the rule in a rigid way, they understand the underlying principles and can flex them accordingly. Countries following what we might call an 'authoritative' approach include Canada and Japan.

To illustrate, 'stay at least 2metres apart' (or else) is a clear simple rule. A more complex approach is the Japanese '3 Cs': avoid closed spaces, crowded places, and close-contact settings. The Japanese never set a specific distance to stay apart, but instead sought to help people understand the underlying risk factors. The problem with the 2m rule is that it was almost certainly too harsh in outside environments, and not effective or harsh enough within closed spaces.

Arguably, the UK tried to have its cake and eat it, and sometimes ended up with neither. 'Stay Alert' would be a good (ie bad) example. It was simple, but not directive enough to give a clear steer as to what to do nor what the underlying risk factors or principles were. Similarly, a number of UK policy and rule changes seeked to maintain simple messages (eat out to help out), but didn't communicate to the public a coherent underlying principle of safe behaviour. [The Cummings incident was also extremely unfortunate: Dom's response would have made sense if the UK had pursued a principle or authoritative strategy, but blew a huge hole in the rule-based approach.]

Our ambivalence towards giving the public more precise risk estimates, and enabling them to make informed judgements, indicates that the UK continues to be caught in this strategic no-man's land. The NHS app, for example, gives the user a risk score for the area they live (or wherever they entered their postcode as being initially), but it doesn't allow the user to enter: the postcode of where they work; the travel route and mode they take (such as car or tube); any personal or health factors that might affect their vulnerability (such as age or obesity); their risk of spreading covid to others (such

as living with an elderly relative); or the behaviours they engage in (eg always washing hands or wearing a mask). Furthermore, as of the time of writing, the risk level on the app doesn't line up with the partial lock-down areas, nor provides users with a list of what the rules are in that area…(see also 'how'? above)

A related issue we continue to wrestle with is the extent to which covid policies should be overtly varied by geography or better still, risk segment. Though local lockdowns have pulled us towards variations, this tends to be a local or regional variation of the national 'authoritarian' position. It is not, for example, backed by S Korean style hyper-local information about a case having occurred in your street, neighbourhood, or local shop, and linked action you should take. Similarly, we have not pursued getting messages (or anything) to second degree contacts of index cases - who we could, and in my view should, have been sending more personalised warnings and requests to curtail their non-household activity.[5]

Finally, the presumption of a simple rules-based (versus principle-based) approach is that there are also clear sanctions. We've generally been slow to get these in place, and certainly the police have been hesitant on enforcement.

This must be acknowledged as one of the more speculative conclusions on lessons learned. Nonetheless, there's a strong case to be made that the UK has tried to pursue a 'soft authoritarianism' - a space that is particularly hard to occupy. It is likely that an 'authoritative' approach would generally have been more effective - and that the British public would have been able to handle it. Of course, where possible messages should be kept simple, and ideally be clear calls to action (cf 'hands, face, space'). But in general - particularly when rules are likely to have to flex and change - we are better off making clear what the underlying principles and vectors are; giving the public, businesses and local areas as much personalised risk information as we can, along with support to comply; and then reserving the application of sanctions to especially high visibility and egregious examples of problematic behaviour.

Finally, in such a model variations in practice, within reason, can be a major asset in as far as they allow learning about more effective practice that can be scaled across a system (see 'eggs in one basket').

Don't: go for simple, rigid rules if you think you are likely to have to vary them, or have limited ability or intention to enforce them

Do: seek to be 'authoritative' in your guidance and policy, laying out clear principles, 'calls to action', and risk data to inform individuals and businesses, that have room for appropriate interpretation and agency at local and individual level

[5] A prototype trial of this has just been agreed this week, to be done in the new T&T lighthouse.

5. Putting humpty together: grip, the cuckoo and the aristocrat

It's hard at the best of times to pull together the Weltanschauung (worldviews) of different academic disciplines and government Departments. It was that much harder when racing against the exponential of an epidemic - and without a Jeremy Heywood at the centre to pull it altogether.

In late February, Chris Wormald did a beautiful impromptu masterclass for Matt Hancock about how the decision-making would work. SAGE would meet, and Patrick would report its conclusions. Chris Whitty would advise Matt as CMO. Mark Sedwill would look across range of the advice, plug gaps, and advise the PM. These views would be brought together, along with economic and other analysis at COBRA. The PM and Cabinet would then make final decisions.

The clunking reality was very far from this beautiful model. SAGE meetings were dominated by a biomedical perspective (eg: no economists were present). Meetings were long and overrunning. Patrick and Chris would literally run from Victoria Street over to Whitehall for COBRA, with no time for even the most rudimentary economic analysis or consideration of delivery challenge. (There was 'SPI-B' supposed to be inputting into SAGE, but that's a whole other story). If Mark was holding other meetings, I never saw them, nor evidence of their impact. It was only with the arrival of Tom Shinner, and later Simon Case, that anything approaching a broader analysis was attempted.

When SAGE got caught in vacillating around its models in early March, or when Chris and Patrick became uneasy about going back to the PM with changed advice, the only way out seemed to be to invade the pitch - as Dom C did by bringing in non-medical modellors to inject an alternative perspective and shake it out of its stupor. [Later criticised, ironically, for inappropriate political interference…] My own increasingly desperate and specific emails in March to Chris, Patrick, and Mark - in particular to move forward urgently and early with Nordic-style social distancing measures (as we would now call them) based on the evidence we had collated from elsewhere seemed to have little impact, with no clear place to land.

Countries that did better moved fast. One of the factors seems to be that such countries had smaller, but more balanced decision-making circles. Jacinda, for example, did rely heavily on a medical epidemiologist that she had known for years, alongside a small circle that she gathered around her. She didn't have twenty medics, and everyone else in the waiting room.

Covid exaggerated a problem that has become quite deeply embedded in our research landscape. The UK elite is very proud of our science base. However, its narrowness and academic character were brought out in the covid crisis. In particular, our science base is extremely skewed to the bio-medial. This is reflected in the profile of UKRI spending, within HMG expenditure, and even the character of 'science' with the CSA network. It has become the 'cuckoo' in the R&D nest. The Councils that dominate the

UKRI budget (£7.5bn) and voice are biomedical. Across Whitehall, the only departments left with significant research budgets, are the MoD (£1.6bn), health (£1.1bn), and dfid/fcdo (£0.3bn). Equally importantly, our key R&D positions are also dominated the biomedical community.

Our CSA is bio-medical (Vallance). The Head of UKRI during the crisis (and previous CSA) was bio-medical (Wallport; though new Head is at least bio). The architect of UKRI was bio-medical (Nurse). The deputy CSA - though CSA at MoD with its sizable budget - is biomedical (Maclean). The DfID has a CSA from London School of Tropical Medicine (Watts, a mathematical epidemiologist), and its predecessor was of course the current CMO (Whitty). Our largest research foundations, such as Wellcome, are bio-medical.

Even when policy was (and is) discussing overtly non-biomedical policy measures to contain the virus, these are referred to as NPIs - 'Non-Pharmaceutical Interventions'... It's like the DfT referring to cycling as 'non-train based-travel'.

This over dominance is self-reinforcing, as these huge budgets, networks and institutions supply people to fill key posts, and steer committees. And all this is against a background where the returns to this research have been halving every 7years for the last half century. We all still hope that our bio-medical investments will deliver us some form of vaccine in the year to come, but it doesn't appear that this body of knowledge gave us much of an edge in the current crisis.

Against this stands the weakness of other disciplines. Alongside all the key medical judgements were behavioural and economic judgements. The 'first lines of defence' against covid rested heavily on shifting behaviour: getting people to wash their hands, maintain (unnatural) social distance, wear masks and so on. Behavioural factors are also key at the 'second-line of defence': getting the right people to come forward to get tested, convincing them to share and prompting their memories of contacts, and then getting people to comply with self-isolation. Even the success of the ultimate and hard medical 'third line of defence' - vaccines and treatments - will rest heavily on behavioural factors, such as whether people will be put off vaccination by social media scare stories, or whether they will comply with treatment and vaccine schedules.

Has our behavioural and economic analysis been up to the job, or even in the room for the key decisions? The Treasury engagement has been erratic - sometimes barely involved at all (such as in the SAGE dominated early and middle phase), and at other times a hard block (such as around providing financial support for self-isolation). Perhaps reinforced by the lack of a strong central process, HMT reverted to 'aristocrat' mode. Requests could be sent into the black box of the HMT palace. They might get listened to, or they might not. All very much a mystery. A fortune might be spent on an inflexible furlough scheme, that seemed to learn few lessons from more effective models elsewhere, but a fraction of that to reinforce the key behaviour needed on self-isolation was a hard no.

I hope that one of the abiding lessons of covid, not least in the face of the social and economic aftershocks that are just beginning, is that we recognise the lacunae in our expertise and our failure to integrate the knowledge we did have. There is an extraordinary skew in our research base to bio-medical, in our key R&D institutions, and in our thinking. We shouldn't respond by shutting down our medical labs, but by applying the same methodological rigour and at least a little resource to raising the game of our behavioural and economic applied sciences.

Either way, we can't have a central policymaking machinery that is caught napping, or sitting on its hands, waiting for someone else to 'put humpty together' and develop coherent policy that balances a full range of considerations.

Don't: let a single perspective or discipline dominate policy - or indeed R&D or the way we think

Do: ensure a range of perspectives, with robust underpinning, are brought together to provide Cabinet and the PM a balanced policy perspective.

Do: in the wake of covid, re-balance the UK R&D base to systematically strengthen our behavioural, economic and operational science (and don't presume the existing institutions are up to the job)

Do: dust off the 'behavioural government' report, and get use the forthcoming civil service reform paper to drive a widespread improvement in the methodological skills of civil servants, including the ability to interrogate the empirical robustness of expert claims.

Conclusion

The UK will emerge from covid bloodied, bruised and poorer. The good news is that it is likely that around 200,000 premature deaths may have been prevented relative to a plausible worst case scenario, albeit at immense cost.

Future enquiries will likely pick up on the alternative paths that could have been pursued to achieve better outcomes. It seems unlikely to be a story of simple villains. The vast majority of the key actors were good people, trying to do the right thing. I also don't buy simple blame games that put it all on the politicians. Indeed, I'd like to put on the record that what I saw of Matt in particular made me feel that he mainly made the right calls, and did an extraordinary job in trying to keep his Department motivated and focused.

At the core will be a 'swiss cheese' failure: an alignment of multiple circumstances and institutional weaknesses that together led to a bad outcome. If the UK had had experience of SARS, we would have been better prepared. If our experts weren't so fixated on flu, they might have grasped the opportunity to pursue a near suppression

strategy and mass tracing (even before testing was available). If the process wasn't so clunky, lock-down could have been done earlier and arguably lighter. If CMO wasn't so fixated on the risks of 'behavioural fatigue', a different strategy might have been pursued. If Jeremy Heywood hadn't died, there would have been a stronger grip in the centre. If brexit struggles and the need to recover from punishing elections and parliamentary battles hadn't been running, Ministers might have moved earlier. If we'd been more focused on engineering and behaviour rather than bio-medical science, we would have done a better job on delivery. And so on.

For all of us, and particularly those at the edge of the infamous policy vortex, we are left wondering what more could we have done. I have certainly had many sleepless nights of frustration and anger. It wasn't just a matter of hindsight. Some good advice was given. It just didn't seem to win through often enough, or at least too slowly (such as the case for earlier but less severe social distancing; the rapid scaling of tracing [and testing]; the earlier adoption of masks; weak support and enforcement; the use of more targeted messaging; headline messages with clearer calls to action; the use of trials and experimentation to optimise comms and delivery; the tracing of second degree contacts; the development of more than one alternative to the app; the use of prediction and delphi tools to build of a better risk model; and so on).

Of course, we can all tell ourselves a story of 'if only'. For all I know, our advice, even if adopted in full, would have made no difference. But I do think that there are important 'no regret', and quite deep, institutional lessons that we should learn.

Many of these are outside of my direct control. But there is one central lesson I do intend to work on, and would greatly appreciate the backing of others to make happen. The empiricism and quality of the behavioural and economic analysis that was incorporated into our policymaking was not good enough. I believe that there are two linked actions that we should pursue.

First, we must seize the opportunity to make our civil service more methodologically sophisticated. We should pursue this through the Civil Service reform paper, through embedding evaluation and experimentation via the SR, and through delivering on the political interest expressed through Michael Gove's Ditchley speech and Dom Cummings drive to reshape the empirical sophistication of the centre. I believe that if we had this in place, some of the key early policy missteps might have been avoided, and in particular a more methodologically confident policy profession would not have ceded decision-making so fully. Almost for sure, more strength in this area would have avoided the 'all eggs in one basket' around the delivery and operationalisation of policy, and would have helped us deliver - rather than merely aspire to - 'world class'.

Second, I am going to try and build a UK powerhouse of behavioural science that can serve our government and people. We need something like a social and economic MIT (perhaps 'a second foundation', in the words of Asimov). It needs to avoid the trap, and the weaknesses, of our current overly academic institutions. Rather it needs

to bridge between the worlds of academia, policy and entrepreneurship. Specifically, I am going to work with NR at Nesta, and NR at UKRI, to see if we plug this gap. We need an ecosystem able to generate, translate, and nurture the adoption of more effective practice. At at its heart we need an institution that can generate cutting edge, practical behavioural science at least of the methodological quality that we would expect of any 'hard science'.

There will be another covid. It's on us to make sure we are ready.

DAVID HALPERN