

RESPONSE

Learning lessons about lesson-learning: Covid complexity

Calum Paton 

Keele University, Keele, UK

Corresponding author. Email: professorpolitic@gmail.com

(Received 6 May 2022; accepted 18 May 2022; first published online 14 June 2022)

Abstract

In the article, ‘Learning Lessons from the Covid-19 Pandemic’, Powell (2022) rightly implies that there is a profusion of confusion in the ‘industry’ which has grown up around lesson-learning from the pandemic. His contribution sets out a helpful framework for classifying or making attempts at lesson-learning. He combines the tripartite classification of inadequate approaches to policy-learning and policy transfer developed 30 years ago by Dolowitz and Marsh (‘uninformed-incomplete-inappropriate’), which he inverts to produce a classification of approaches which are informed, complete and appropriate, with the framework of ‘outcome-mechanism-context’ from realistic evaluation. (I use the term realistic rather than realist, as the latter implies an epistemological stance as opposed to what was intended, which is that evaluation takes account of complexity in a realistic manner.) This produces a classification, and possibly an ‘ideal type’, of informed outcomes, complete mechanisms and appropriate context. Powell rightly implies that no overall conclusion is available from the literature reviewed. He does however imply that different approaches may work in different settings. This is true in one sense but misleading in another. This commentary argues that such ‘relativism’ is not only dangerous in practice but mistaken in theory.

Key words: Covid-19; successful policy

1. Powell’s conclusion

Powell rightly implies that none of the literature which he reviews and categorises (as far as each article’s content allows) within this framework is adequate for the purposes of lesson-learning (Dolowitz and Marsh, 1996). This is because the articles reviewed range from (in my words) arbitrarily chosen bipartite comparisons, through larger scale general essays (again, my term), to broad prescriptions, none of which satisfactorily correlate outcomes with causal or contributory factors. Powell does not provide his own empirical analysis of Covid data, which is of course not the point of the article. His only ‘conclusion’ from the literature reviewed is that there is no possible conclusion (from this literature, I would emphasise) as to which countries have achieved better outcomes in the pandemic through one unequivocally effective policy. He does not do a ‘meta review’, nor does he try to, as even the most (hypothetically) relevant studies which emerge from his literature search do not allow such. His ‘non-conclusion conclusion’ does however suggest that different approaches may work in different settings and that there is no single solution applicable in all cases.

This is true in one sense but misleading in another. Moreover, in the ‘dirty hands’ of politicians and various ideologists, such a conclusion may be used to excuse failure, or failure to consider adequately tough responses to the pandemic. Let me explain. Powell rightly draws the observation, from his chosen literature, that the relatively successful countries in South East Asia did not all *emphasise* (my emphasis!) the same approach. This is true, at the level of which particular ‘technical’ mix of ‘non-pharmaceutical interventions’ (i.e. governmental and social actions) they prioritised. Yet it may be misleading in that all successful countries, globally,

© The Author(s), 2022. Published by Cambridge University Press

were strong, tough and steadfast in tackling Covid and used whatever policies *were necessary in their social context* (again, my emphasis).

Successful countries did not eschew – or limit – lockdowns and/or other social restrictions through ‘political’ choice (of the sort which influenced the Johnson government in the UK, e.g.) as opposed to more legitimate reasons, e.g. policy reasons such as timely and comprehensive testing and tracing or social reasons such as the fact that compliant and/or unselfish societal behaviour rendered them less necessary. This is not the same as choosing a response to Covid by broad policy preference as opposed to what the local situation requires as measured empirically and analytically.

An example can clarify my point. Anti-lockdowners in the UK point to Sweden as producing a ‘good’ result without lockdown. Yet most of this claim is false, and the true part of the claim is misleading. Sweden produced a worse outcome than its Scandinavian neighbours which had tougher restrictions. The reason it did better than expected *given* the absence of lockdowns is not because it did *not* have lockdowns: it is because Swedish culture was more consensual in promoting appropriate social behaviour in the absence of statutory restrictions. Moreover to compare Sweden with the UK, for example, and conclude as do libertarian-leaning commentators that the latter did worse, or no better, *with* lockdowns than Sweden did *without*, is nonsensical. (As an aside, in any case, Sweden did virtually as bad as the UK in the first wave of Covid.) The UK’s lockdowns were too-late, too-loose and too-short. The UK witnessed more polarised social behaviour, with non-compliance by a large minority legitimised by a *cod* libertarianism which ignored the fact that ‘freedom’ to hurt others is the ‘freedom’ of Hobbes’ state-of-nature (nasty, brutish and short). Worse, such libertarianism was persistently tipped the wink by the UK government. The UK without lockdowns at all would have made death rates, which were already egregiously high, significantly worse.

This is not a criticism of Powell (2022), of course. But it is a health warning as to the (mis)use of a conclusion that it is ‘horses for courses’ when it comes to managing Covid. To be successful in combating Covid requires ruling nothing out and indeed having a hierarchy of required policies to be activated as necessary. While one single intervention is indeed not the universal explanation for success in a mechanistic and simplistic manner, different interventions – from closing borders; through effective testing, contact-tracing and isolation; to lockdowns – are related in a hierarchy where mobilising one depends on what has happened with others. In other words, effective Covid policy is not something which can be chosen from a menu of options; it requires a universally-understood approach as to when and how different interventions should be mobilised. One size does not fit all in the simplistic sense of mechanistic adoption of the same intervention on the same scale in all countries despite countries experiencing different stages, scale and scope of the pandemic. But one size does fit all in the sense of accepting that, if the policy aim is clear (for example, minimising Covid), then for countries in the same particular circumstance, there is likely to be one best answer. The World Health Organization, for example, was rightly criticised for certain faults, including (initial) softness with China. But its universal prescriptions (such as ‘test, test, test’) were more appropriate than many sceptics, whether academic or political, acknowledged.

2. A multi-country comparison

It might be useful to consider a contribution which *does* seek to draw comparative conclusions, *inter alia* to see if the two articles considered together stimulate further thought as to lesson-learning. Greener (2021) sought to draw conclusions for (the first wave of) the pandemic by comparing 25 OECD countries; and, as one of the UK’s most accomplished social scientists, he has made an ambitious attempt. I am only considering this one additional article here, as this is not a literature review, but a discussion for heuristic purposes stimulated by the challenge laid out in Powell’s article. Noting that Greener’s research only covered the first wave of the pandemic, we might assume that future analyses should cover the calendar years 2020 and 2021, as the

rise of the Omicron variant at the end of 2021 in a (for some) post-vaccine context provides a sensible cut-off point: most countries, other than most notably China, at the time of writing, have faced large numbers of Omicron cases.

What both studies, considered jointly, demonstrate (inadvertently) is that neither an overt framework nor comparative empirical study *per se* is necessarily enough to conclude which countries had successful policy, and certainly not in order to provide lessons in policy or decision-making. To put it another way, Powell's framework would need to be extended to become both less general and less generalising (it applies Occam's razor too ruthlessly); and Greener's empirical study loses 'bite' as a result of using multi-country statistical relationships to identify factors which explain which countries did well and others did not (as gauged by outcome measures such as deaths and case levels). This is not to disparage the attempt: it is a matter of 'horses for courses'; we may note that Greener *et al.* (2021) are clear about the sort of factor which may be salient in explaining 'good' (e.g. Germany) and 'bad' (e.g. UK) reactions to the first wave of Covid. Greener indeed starts by stating that there is a place for more detailed country comparisons.

Greener uses 'qualitative comparative analysis' (QCA). This is not the place to provide a critique of CQA, which is not actually all that qualitative, maybe in general and certainly in this case. The only 'mechanism' as opposed to 'context' (in the terms of realistic evaluation) which Greener's study considers is (the amount of) 'testing' undertaken by a country, as measured by numbers of tests per numbers of Covid cases. Using CQA, this comes out as 'necessary' for success. Yet this is too broad to be meaningful as a conclusion. The way CQA works – identifying different possible routes to a successful outcome for each country, based on different combinations of variables – means that while 'testing' comes out as necessary for success, it is not necessarily sufficient (see below). The factor 'testing' is so generic that it begs many questions, although a country's ability and willingness to mobilise testing at short notice in the first Covid wave is an important measurement.

A time-honoured methodological issue also arises – the difference between association and causality. For example, Greener's most significant factor, Covid-19 tests per case, may well be an association rather than a causality, or indeed the causality may be the other way around in some cases: already-successful countries test more, and follow it up with effective tracing and effective isolation, to keep Covid rates down. Such factors, hypothetical as they admittedly are, need to be unpicked.

None of this not a criticism of Greener but a pointer to what he would no doubt acknowledge – the need for detailed qualitative work to explore the 'mechanisms' (policies, decisions, changed social behaviour, etc.) which represent what governments were able to use or mobilise to improve the situation, as opposed to the contextual factors which they could not change, at least in the short term. That is, if we are seeking lessons and not (only) academic explanations, we should be focussing upon not (only) contextual factors but also the sort of factors which are unlikely to be reducible to 'independent variables' in a comparative quantitative analysis or 'factors' in CQA. Of course to use work such as Greener's as the start-point for, or accompaniment to, studies on lesson-learning may be worthwhile.

Speaking as a political analyst, we need to put the politics back in. It tends to be lost in multi-country statistical comparisons, unless these are ingeniously designed. Furthermore, in the language of realistic evaluation, Greener's explanatory factors conflate contextual factors and (only one) mechanism (i.e. policy or decision) and the identification of the scope for agency is lost or buried.

3. Learning lessons for countries, not providing excuses

Capturing the key decisions which allowed a country to outperform or under-perform, given its 'contextually'-based expectations, is surely the challenge for lesson-learning. This requires both qualitative depth and judgement as to the applicability of conclusions. Conclusions from

statistical comparisons cannot provide the necessary insight. This involves structured narratives. Powell points out rightly, nevertheless, that we have to avoid narrative comparisons which are context-free or context-misleading. Thus, in my terms, we have to negotiate a subtle path between the Scylla of statistical generalisation and the Charybdis of context-blind conclusions and therefore misleading recommendations.

Powell also points out a different issue, concerning lessons and recommendations handed down as universal wisdom applicable to all. Concerning the WHO mantra of ‘test, test, test’, for example, they consider countries which do not have the resources to do so. One might paraphrase their argument: simplicity has advantages, but universal simple prescription has a downside. The problem with WHO’s generalised good practice is not, of course, that it is arbitrary or that it fails to be research-based: indeed, WHO, like some individual countries, was too slow in some respects during the pandemic, waiting for evidence (from China; about transmission) or updated research (e.g. about masks) when the luxury of waiting does not exist in a pandemic. It is about whether ideal solutions are affordable. Powell points to the Bhilwara model from India as an alternative for low-income countries, in this respect, and such is a good caveat.

We should be clear however that the need for relevance based on context should not lead to what I will call ‘over-respecting’ context. ‘Culture’ is not a legitimate reason for policy-relativism. It might be argued (and indeed the literature on policy transfer and lesson-learning which Powell reviews generally does argue) that particular policies seen to work in one type of context (culture, in particular) should not be recommended for another. This is one of the key lessons drawn in conventional consideration of the pitfalls of comparative policy transfer. But observing such niceties from comparative methodology in an exceptional situation such as a pandemic might be a mistake. Bucking the culture may be necessary, in an existential moment. Consider how ‘libertarians’ rejected lockdowns on the grounds that they were not appropriate for a ‘free people’. In the UK, however, ‘freedom’ was often a euphemism for a weak government indulging its core constituency and seeking short-term popularity (Paton, 2022).

Avoiding approaches which ‘alien’ cultures use was a *leitmotif* in some liberal political cultures at the beginning of the pandemic: it was assumed until it was too late, to different degrees in the UK, Europe and the USA, that rigorous lockdowns were suitable for, or workable in, authoritarian political cultures only. Italy has an excuse: it had a stringent lockdown, but the fact that it was too late was less culpable than, for example, in the UK, which was observing (or failing to observe) Italy’s problems unfolding in real-time (Italy was ‘hit’ 2 weeks ahead of the UK) and failing to act until, armed with alarming predictions, the UK Prime Minister’s chief assistant forced action upon an unwilling boss. In the UK, scientific leaders (the Chief Medical Adviser; the Scientific Advisory Group for Emergencies) had self-censored with the ‘liberal’ assumption that border closures were not on the agenda and therefore should not be considered as a response to Covid, which obviously suited the political leadership. This was of a piece with the assumption that the scientific method must include seeing what works, i.e. getting evidence before deciding. But waiting for evidence when *not acting* is itself a decision – what political scientists call a non-decision (Bachrach and Baratz, 1970) – can be disastrous, and arguably was for many countries in the pandemic.

Effective leadership might mean *going against* culture: commanding and controlling where necessary. One of Greener’s ‘independent variables’, or factors, was the ‘degree of openness of countries to international visitors’. Governments sometimes behaved as if business trips and tourism were facts of life, independent of government agency. It would have been interesting to model the effect of closing borders and/or restricting travel, to compare those countries – such as New Zealand, Australia and various South East Asian states – which did so with others which did not, in terms of Covid outcomes. Again, the question arises: is the aim to explain the past, or learn lessons for the future? If the former, then ‘degree of openness’ is of course a relevant measure.

A general issue arises for lesson-learning, which points to a psychological danger, rather than a logical corollary, of differentiating by context. There is a danger that *interpreting for policy* the relation of country typologies to Covid responses leads to an acceptance of the view that expected

response must take account of context. As a result, ‘expected’ becomes accepted. There is a difference between the constraints of ‘normal’ policy transfer based on policy learning and exceptional imperatives in a crisis. In a nutshell, it is important to avoid excuses for a lack of ambitiousness in dealing with Covid.

4. Welfare typologies of countries

This leads to the question of whether different ‘types of countries’ were associated with different Covid outcomes. Greener discusses various typologies of health and welfare systems (see, e.g. Esping-Andersen, 1990; Bambra, 2007), and plots Covid outcomes against countries in terms of these typologies. He finds that the typologies of welfare systems which most correlate with Covid outcomes are those which focus on the variables of total social expenditure, on the one hand, and societal redistribution, on the other. Since there are four possible combinations of these two variables, the picture is complex, but he finds that high social expenditure and high redistribution seems to plot quite well against Covid success – with no such countries classifiable as ‘liberal’ (as opposed to social-democratic, conservative or ‘Southern’). Greener finds that the countries with a combination of low redistribution and high expenditure tend to produce poor Covid outcomes, but the exceptions (Australia and Japan) are so striking, in the context of low total numbers, that the conclusion is possibly tenuous. Similarly with low expenditure and low redistribution: the exceptions (New Zealand and South Korea) to a tentative conclusion of poor Covid outcomes are truly ‘headline’ exceptions. These were two of the most successful countries, arguably *the* most successful, bar China, especially at this stage.

There is a wider problem. The typologies of countries and their welfare systems from political science are neither consistent nor – especially in 2022 – all that convincing. One simple example: typologies class high spending, highly redistributive Finland, and also France and Germany, as ‘conservative’. The latter two are so classed because *inter alia* they have Bismarckian rather than Beveridge health systems: but this is wholly irrelevant to Covid. Moreover France is Bismarckian in name only. Greener extracts, from three typologies (see Greener, 2021), his two measures (expenditure and redistribution), which is instructive; but then to go on to use broad terms such as liberal, conservative, etc., to consider which types of country dealt/deal with a pandemic better is fraught with difficulty. I do not think Greener would demur, by the way.

At the level of instinct, it *does* seem plausible that ‘liberal’ countries are likely to have greater difficulty in terms of political culture in dealing with Covid (although one might distinguish between liberal political institutions and liberal-libertarian political culture). Perhaps, however, common sense is a better guide than initial instinct on the one hand or theoretical typologies, on the other: New Zealand and Australia are liberal polities, whatever typologies from political science may say (New Zealand had a social-democratic government; Australia had a populist libertarian conservative government). They were both models for handling Covid, until the Omicron variant arose, which was moreover in a post-vaccination context: they therefore avoided the mass deaths which unvaccinated populations incurred in previous waves, based on previous variants.

Typologies of welfare-type in terms of social expenditure and distribution may moreover have little relevance to the specific characteristics of public health response, which are arguably more about state capacity and leadership styles. The question of association as opposed to causality then becomes important, not least in the need to beware of false inferences. This is not to deny the importance of potentially correlating contextual variables such as social and demographic factors with Covid outcomes: it is, however, to distinguish these factors from those relevant to learning lessons about how a government can act better in real-time when confronted by a pandemic. Moreover, even beyond the sphere of public health, typologies such as Esping-Andersen’s may have had their day. To describe France, for example, as a ‘conservative’ polity in Esping-Andersen’s sense seems eccentric, bizarre and outdated.

5. Learning lessons for research

Returning to Powell: his classification may at the very least allow one to distinguish better between, on the one hand, unalterable contextual factors which constrain attempts to achieve the outcome of (say) a low number of deaths from Covid and, on the other hand, those policies or decisions which make the best of one's contextual position, i.e. the mechanisms used, in realistic evaluation's argot. The challenge for policy-learning is to disentangle unalterable contextual factors from ones which may be altered, even if with difficulty. Powell's framework, however, must be used in a manner which does not conflate different types of mechanism or different types of context (see below.) When it comes to 'operationalising' Powell's framework, there is a good argument for grouping countries with similar 'contexts' in order *then* to compare mechanisms, and even then, it is unlikely that mechanisms, as opposed to contexts, will be susceptible to statistical analysis.

In Greener's work, there is not a meaningful distinction within the 'independent variables', i.e. factors, in CQA's language (the hypothetical predictors of a good or bad Covid outcome) between contextual and mechanistic factors. In plain language, there is not a distinction between factors which can be influenced by governments and policy-makers in resisting Covid and factors which cannot, especially in the short or even medium terms (and therefore of no use in a pandemic when the mantra has to be 'act now'). The 'alternative routes' to better Covid outcomes (lower deaths and/or cases of Covid) conflate types of factors. Moreover, having 'sufficient conditions' for a good outcome do not mean that good outcomes must contain these conditions: the salience of such conclusions depends on how much of the total sum of good outcomes these sufficient conditions contain.

It can be argued that the point of Greener's research and similar exercises is not to provide a playbook for policy but to explain the past. This is a fair point, but even so, other problems arise.

As pointed out above, the solitary factor in Greener's study which reflects real-time decisions (i.e. policy made by government) is 'testing' (adjusted for need in a rudimentary way), which he finds to be a necessary component of all the different routes to better outcomes, i.e. widespread testing in those countries which are 'successful' (relatively) in the study. But this 'policy' variable isolated is too broad to represent policy in any meaningful way. Next, Greener seeks out 'sufficient' solutions for success. But the factors which enable this are contextual – demographic and socio-economic, such as proportions of elderly in the population and GINI coefficients for countries. Thus, we get a hybrid of academic explanation and (possible) lessons.

This is not to deny that it is useful to know the hypothetical effect of demographic and social factors upon a country's prospects. Such can then be used to show how variation around predicted levels of death can be explained. Germany, for example, had a much lower death rate than the UK amongst the elderly during the 'unvaccinated', earlier waves of Covid as a result of its health services having more resources – in general, in intensive care and in staff. This points to an intermediate variable – between unalterable 'context' and agency: in this case, availability of, and/or mobilisation of, health services.

The conclusion might therefore be that quantitative studies should not only distinguish between different types of contextual factor, but should then form the backdrop for detailed qualitative, narrative studies. 'Qualitative comparative analysis' is arguably mostly quantitative, and the 'qualitative' variables, which are then quantified to become input as factors, are themselves forced into a quasi-quantitative framework. Key factors which actually explain different countries' different outcomes, or significant parts of such, are missed altogether, as they are not susceptible to such analysis.

To improve the situation, using Powell's framework with careful sub-categories in each of the three domains could then be used to develop an empirical analysis of more use to understanding either which countries made the best of their situation or how future challenges could be better tackled. Put another way, if we want to be ambitious and quantify, we might seek 'expected values' of deaths, once we have a data set from Covid for our countries to be studied, with the weighted

average providing the baseline. These values would reflect the contextual factors – which ought not only to include demographic measures and other quantitative data but also measures of, e.g. ‘social compliance’. The latter – to give an example – would lower Sweden’s expected deaths, and show that its failure, not only by comparison with its Scandinavian neighbours but more widely, was even more pronounced than noticed – and it would certainly give the lie to those who point to its relative absence of restrictions as a success.

6. Concluding remarks

Finally, and returning to Powell, I would reiterate that, when it comes to (what in realistic evaluation are) ‘mechanisms’, i.e. policy- or decision-relevant variables, even were it possible in theory to quantify these, then the complexity in practice would probably rule out such quantification. For example, for the potential variable of ‘lockdown’, is it ‘Yes/No’? How do we quantify timing and duration in a manner which can capture (enough of) the richness of reality to generate policy-relevant conclusions? For lockdowns and other restrictions, how do we quantify enforcement? And so on. Powell provides a useful start-point for improving future lesson-learning. Likewise with Greener’s research: my discussion here is testimony to its value in stimulating wider thought.

Of course, the thing about pandemics is that learning from the previous one, when confronted with the threat of a new one, may not be relevant. Different viruses, of differing severity and different types and ease of transmission, require different reactions. Not over-reacting to the novel coronavirus at the beginning of 2020 was a ‘lesson’ from previous threats allegedly drawn by some policy-makers in the UK. The importance of risk-aversion in individual countries and regional blocks would, however, be the appropriate lesson to be drawn from Covid-19.

References

- Bachrach P and Baratz M** (1970) *Power and Poverty: Theory and Practice*. New York: Oxford University Press.
- Bambra C** (2007) Going beyond the three worlds of welfare capitalism: regime theory and public health research. *Journal of Epidemiology and Community Health* **61**, 1098–1102.
- Dolowitz D and Marsh D** (1996) Who learns what from whom: a review of the policy transfer literature. *Political Studies* **44**, 343–357.
- Esping-Andersen G** (1990) *The Three Worlds of Welfare Capitalism*. New Jersey: Princeton University Press.
- Greener I** (2021) Comparing country risk and response to COVID-19 in the first 6 months across 25 organisation for economic co-operation and development countries using qualitative comparative analysis. *Journal of International and Comparative Social Policy*, 1–15. doi: 10.1017/ics.2021.6.
- Greener I, Powell M and King-Hill S** (2021) Intra-crisis lesson-drawing in real-time: the pandemic lessons available in the UK media during the first months of COVID-19. *Social Policy and Society*. <https://doi.org/10.1017/S1474746421000610> (online).
- Paton C** (2022) *NHS Reform and Health Politics in the UK: Revolution, Counter-Revolution and Covid Crisis*. London: Palgrave Macmillan.
- Powell M** (2022) Learning lessons from the Covid-19 pandemic. *Journal of Health Economics, Policy and Law*.