Breaking the Overton Window: on the need for adversarial co-production

Elliott Aidan Johnson, elliott.johnson@northumbria.ac.uk
Irene Hardill, irene.hardill@northumbria.ac.uk
Matthew T. Johnson, matthew7.johnson@northumbria.ac.uk
Northumbria University, UK

Daniel Nettle, daniel.nettle@ens.psl.eu
Northumbria University, UK
Ecole Normale Supérieure-PSL, CNRS, France

Co-production has emerged as one of the key concepts in understanding knowledge-policy interactions and is associated with involvement, for example, of users of public services in their design and delivery. At a time of permacrisis, the need for transformative evidence-based policymaking is urgent and great. This is particularly important in highly distressed 'left-behind' communities targeted by the UK Government for Levelling Up, which constitutes an attempt to improve the infrastructural, economic, social and health outcomes of less affluent parts of the UK. Often, policymakers regard the transformative policies capable of addressing these crises as beyond the 'Overton Window', which describes a range of policies in the political centre that are acceptable to the public (Lehman, 2010). This window of opportunity can shift to encompass different policies, but movement is slow and policymakers generally believe that significant change lies outside. In this article, we build on recent debates in Evidence & Policy on co-production by outlining an embryonic approach to overcoming this Overton Window-based roadblock in evidence-based policymaking: adversarial co-production, which involves working with opponents of evidence-based policy to develop means of persuading potential beneficiaries to support introduction. This emerging approach has been deployed in examination of public preferences with regard to welfare reform, but can be applied to a wide range of policy areas. We outline briefly the history of co-production, before setting out the process by which adversarial co-production was developed. We then describe the impact of adversarial co-production on public preferences on basic income (BI). This enables us to set out challenges and opportunities for those with an interest in addressing our crises, serving to stimulate genuine debate on longstanding assumptions about the limits of evidence-based policy and public opinion.

Key words adversarial co-production • policymaking • evidence-based policy • crisis • UK policy

To cite this article: Johnson, E.A., Hardill, I., Johnson, M.T. and Nettle, D. (2023) Breaking the Overton Window: on the need for adversarial co-production, Evidence & Policy, XX(XX): 1–13, DOI: 10.1332/17442648Y2023D000000005
Introduction

Co-production has been described as a ‘buzz-word’ (Adelle et al, 2019: 1). It has emerged as one of the key concepts in understanding knowledge-policy interactions (Bandola-Gill et al, 2023b: 276) and is associated with involvement, for example, of users of public services in their design and delivery (Boviard, 2007). The need for improved policymaking is more important than ever, both within the UK and beyond. We now appear to be in a period of permacrisis, with the longer-term impacts of the Global Financial Crisis, a decade of austerity, pandemic, geopolitical conflict and the cost-of-living crisis all compounding and being exacerbated by a climate emergency that may only get worse over time (see Turnbull, 2022). The Financial Times recently described the UK as increasingly a poor country with some very rich residents (Burn-Murdoch, 2022). With high levels of polarisation and alienation from politics (Uberoi and Johnston, 2022), policymakers need to persuade an alienated electorate to support evidence-based policies (EBP) to mitigate poverty, inequality and the climate emergency. This is particularly important in highly distressed ‘left-behind’ communities targeted for Levelling Up – a policy that constitutes an attempt to improve the infrastructural, economic, social and health outcomes of less affluent parts of the UK (Secretary of State for Levelling Up, 2022). Too often, policymakers believe that the policies that can transform society and address our crises sit outside of an Overton Window of public acceptability, particularly in left-behind communities (see Johnson et al, 2023).

There is good reason to understand the role that forms of co-production might play in changing this and how an innovative form can address two issues that represent an equipoise in public policy. On the one hand, Bandola-Gill and colleagues (2023a; 2023b) suggest that preference for technocratic legitimacy leads to presentation of quantitative evidence disconnected from the political legitimacy associated with people’s values (see also Hornikx, 2007). The rejection of apparent economic expertise by supporters of Brexit (that is, Britain’s leaving of the European Union), has been presented as evidence of this deficit. On the other hand, we have argued that the opposite has become true, with policymakers, in the wake of Brexit, believing that citizens of left-behind communities in particular hold fixed value-based positions that preclude the presentation and implementation of evidence-based transformative policy (Johnson et al, 2022). As a consequence, otherwise progressive politicians have often elected to support conservative social and economic positions in order to appear sensitive to existing value sets, rejecting the capacity for persuasion (Johnson et al, 2022). In this regard, policymakers take people’s positions, as represented in focus groups, as fixed and they alter their public positions accordingly. Thus the evidence base they end up incorporating into policy is that of voters’ expressed positions, rather than actual evidence of impact. The approach is often deleterious to citizens themselves, since it leads, necessarily, to evidence-based policy being rejected. This is true, for example, of UK Labour’s recent decisions to U-turn on commitments to wealth and higher rate income taxes and an adequate Green New Deal for which there are good evidence bases (Crerar, 2023).

In this debate piece, we wish to build on recent debates in Evidence & Policy on co-production, led by Bandola-Gill and colleagues, by outlining an embryonic approach to overcoming this roadblock in evidence-based policymaking through adversarial co-production. Adversarial co-production involves working with opponents...
of critical EBP to persuade citizens of EBP’s value and to shift or break open the Overton Window accordingly. This emerging approach, which seeks to break down echo chambers, has been deployed in examination of public preferences with regard to welfare reform (Johnson et al, 2023), but has the potential to be applied to other policy areas. It is less concerned directly, than in work on co-productive agility by Chambers and colleagues (2022), with addressing epistemic injustice, but has a similar commitment to challenging dominant assumptions among political elites about which EBPs can be pursued. In what follows, we argue that adversarial co-production in policymaking has the capacity to create opportunity for critical transformative change and can be applied to policies at various spatial scales. We briefly outline the history of co-production, before setting out the process by which adversarial co-production was adapted from behavioural science. We then describe a pilot study of adversarial co-production on public preferences on basic income (BI). This enables us to set out a series of challenges and opportunities for policymakers with an interest in addressing our crises.

The history of co-production

In order to understand the need for engaging adversaries, it is important to understand the basic promise of co-production in policymaking. The involvement of service users in the design of public services gained momentum during the 2008 economic crisis and period of austerity (Fotaki, 2015), aiming simultaneously to improve design and legitimacy of approaches. Co-production as a methodological approach has been adopted in public administration, science and technology studies and sustainability science with similar developments in the health sciences (Miller and Wyborn, 2020). For Swedlow (2012), this emergence is explicable in cultural theoretical terms, with the adoption of co-production representing a post-Global Financial Crisis change in social and political order mirroring with a shift in beliefs about human and physical nature, noting the ways in which scientists involved in struggles over land and wildlife management in North America championed cultural constructs and policies. These trends have been associated with attempts to address epistemic injustice and practical deficits by mobilising multiple knowledge, including \textit{techne} and \textit{phronesis}, of those subject and party to policy, including citizens and practitioners (Flyvbjerg, 2001).

The implications for policymakers have been significant. A number of epistemological traditions, including feminism and pluralism (see Mouffe, 1999), have come to foreground co-production through user engagement, presenting the method as means of social, economic and political transformation, hoping to change the world not merely to interpret it (to misquote Marx). Indeed, there is evidence to support this view: co-producing knowledge can deliver academic and wider impacts on economy and society (Hardill and Baines, 2009; Hardill and Mills, 2013; Adelle et al, 2019). The COVID-19 pandemic granted renewed focus on co-production in health (Redman et al, 2021) and public service settings (Bandola-Gill et al, 2023a). A recent evaluation of methods employed by research teams on UK investments supported by UK Research and Innovation (UKRI) funding noted a number of projects on the impact of the pandemic on economic and social life undertaken during the emergency that were underpinned by the principles of co-production (Hardill et al, 2022).
The pandemic transformed many aspects of everyday life across families, communities, institutions, and civil society (Kupferschmidt and Cohen, 2020; Calvert and Arbuthnott, 2021; Hennessey, 2022). While it may be tempting to frame COVID-19 as indiscriminate, its spread laid bare existing deep-rooted socioeconomic inequalities, across categories such as health, employment, and housing (Bentley, 2020; Peirson, 2021). This has brought into focus the vital importance of collaboration across disciplines, paradigms, and methodological traditions (Kara and Khoo, 2020; Tremblay et al, 2021). Hardill et al (2022) followed Wiles et al (2011) in identifying innovation at one of three levels: inception, adaptation, and adoption. Hardill et al (2022) emphasise the importance of interdisciplinary adaptation of existing methods. Chambers and colleagues (2022) have advanced co-productive agility as a normative means of advancing epistemic justice and fostering learning and mutual respect. This is grounded in ethical commitment to bottom-up, long-term community change. We share concern consistent with potential long-term outcomes, but have sought, in an era of crisis, to adapt co-production to shorter-term foci.

Methodological innovation and adaptation

One key potential source of adaptation for policy development from behavioural science is adversarial collaboration (Mellers et al, 2001; Clark et al, 2022). The approach, which deploys collaboration to denote work between academic colleagues, as opposed to co-production with non-academics, is intended to improve research via collaboration with those who hold competing hypotheses, and partly involves each side exploring what it would take to convince them of the position they currently oppose. Daniel Kahneman, the originator of Adversarial Collaboration, and subsequent proponents, most notably those involved in the Penn Arts and Science (2022) Adversarial Collaboration Project, conceive of the method as a means of improving the process underpinning protocol development, data collection and publication through the following principles:

(a) understand and articulate their opponents’ perspective so well that each side feels fairly characterized; (b) work together to design mutually agreed-upon studies that have potential to adjudicate competing hypotheses and that they agree, ex ante, could change their minds; and (c) jointly publish the results, regardless of the outcome. (Clark et al, 2022: 6)

The approach is informed by the notion that scientists ‘are constrained by the same cognitive biases, limitations, and tradeoff calculations as mere mortals’ (Clark et al, 2022: 3). Some of its proponents contend that behavioural science is particularly prone to the problems this perpetuates as a result of: 1) its having comparatively low consequences, mainly confined to embarrassment, for producing ‘inaccurate’ findings; 2) its focus on provocative, moral-political topics that may be subject to particular social desires of the researcher and/or broader society; and 3) the fact that it deals with very complex, abstract and interacting variables and social/psychological constructs that can be finessed to obtain different results (Clark et al, 2022: 4). This has produced a range of innovative pieces of research with underpinning design on which a larger proportion of colleagues agree. It is a transformative approach at a time of paradigmatic polarisation.
Policymaking is subject to a similar or greater degree of methodological contention. In particular, colleagues find themselves in disagreement over the feasibility of policy, especially where there is a critical need for intervention and a strong evidential basis for implementation (Johnson et al, 2023). Those in public policy and politics often hold that there is an ‘Overton Window’ of policies acceptable to the public that provides parameters for action (see Lehman, 2010). There is belief that the public, including those who are likely to benefit most from implementation, will reject policies for value-based or cognitive reasons that evolve slowly over time and not in response to presentation of the policy itself. There is widespread belief that transformative policy of the sort needed to address chronic health and economic inequalities or the climate emergency sits outside the window, shifting policymakers’ appraisal of ‘sensible’ policies toward approaches that are inadequate or actually exacerbate harm.

This is apparent in the case of policymaker adoption of ‘nudging’, which formed a central pillar of public health mitigation and messaging during the pandemic. Nudging revolutionised debate on implementation of social policy, with Thaler and Sunstein (2008) presenting policymakers with means of avoiding ‘big state’ intervention, non-coercively influencing behaviour toward evidence-based ends. Shove’s (2010) account of the ‘ABC model’ holds that policymakers change attitudes by informing individuals of the costs of behaviour and providing incentives for change. Their approach holds two critical deficits. First, as Dowding (2020) notes, this privatises responsibility for action and suggests that active state intervention is unjust or counter-productive, abrogating policymakers of responsibility for outcomes. Second, and more critically, there is an emerging body of evidence to suggest that nudging does not work for societally important outcomes. Maier et al (2022) have shown that evidence in support of nudging is subject to publishing bias. When that bias is controlled for, there is little or no evidence of efficacy across such key areas of study as health, finance and environment, and that is without even taking into account recent allegations of data falsification by leading figures in the field (O’Grady, 2023). This indicates that academics, policymakers and governments may have invested significant resources in an approach that is incapable of achieving the ends to which we have a shared interest. Indeed, given the emerging evidence on nudging, it is possible that belief in it may have reduced the efficacy of pandemic responses that often depended upon it. As a field, we all share a responsibility for this. Too often, belief in the Overton Window and belief that sensible can be equated with minimal has led us to endorse research, policy and practice that has had a distortionary effect on public resources and people’s lives. There are good reasons to believe that Labour’s recent positions on wealth tax and a Green New Deal (Crerar, 2023) are examples of this, since there is little evidence that doing what is currently done is sufficient.

It sounds banal to suggest that the crises our societies face cannot be addressed by inadequate policy, but too often this position has appeared radical in ways that are historically distinctive. Human history is one of persistent change and reform in response to crises, but also by virtue of policy entrepreneurs’ persuasion. The neoliberal reforms of the past four decades, many of which have contributed to our permacrisis, represented clear, radical shifts that contemporary assumptions would present as unfeasible. There is a need to remind ourselves that change is possible, and to explore means of using policy to that end.

The methodological deficits identified by Clark et al (2022) have the potential to be amplified throughout the stages of the policy development process. Policies are
often presented to the public on the basis of ‘objective’ benefits assumed by those who design them to be salient, but which may be regarded as irrelevant or less relevant by communities. The avoidance of macro-economic harm by remaining in the EU was clearly irrelevant to many voters in left-behind communities who had witnessed long-term decline in living standards and who simply could not imagine their lives declining further. Policymakers face difficulty in disentangling their material interests from the interests of others and in grasping the bases of differences of opinion on policies. This, combined with the assumption that perceptions are fixed and grounded in value commitments with no clear relationship to material interest, have contributed to inertia and error in policymaking and created path dependencies that compound each crisis in turn. We need to remind ourselves that necessary change is possible. Working with opponents of EBPs from the general public to adversarially co-produce persuasive narratives to promote those EBPs is a means of achieving this.

**Persuasion is possible**

One of the key deficits in co-production is that it has often implied that we ought to take citizens’ positions at face value and as fixed, and to adapt EBP to the Overton Window of citizens’ views. Counterintuitively, this may actually diminish public discourse. As in Chambers et al (2022), we believe there are good reasons to reappraise the capacity for co-production to persuade. For example, during the pandemic, evidence emerged within the UK from furlough, which granted state support for wages during the pandemic, and a range of other interventions that reducing conditionality and increasing universality within the welfare system would enable policymakers to mitigate the impacts of COVID-19. We were funded by the Wellcome Trust to examine the prospective impacts of one such policy, Basic Income, on mental health. Our findings indicated that BI was an economically affordable, multipurpose policy instrument capable of securing transformative impact in public health, preventing or delaying, for example, 200,000 to 550,000 cases of anxiety and depression between 2010 and 2030 within the 14–24-year-old cohort – enough to pay for 750 mental health hospital nurses annually. Even though the policy is an economic instrument, exploration of public opinion has generally been driven by consideration of fit with people’s values, rather than understanding of its material impacts (see Hamilton et al, 2021). This focus on presenting BI in terms of values, and the notion that people’s values are fixed parts of people’s identities, meant that there was widespread belief among policymakers concerned with political legitimacy that BI as a transformative policy sat outside the Overton Window (Cowan, 2020; Gopal and Issa, 2021) and was not feasible for implementation due to public opposition (see Berry, 2021).

We examined debate on persuading prospective beneficiaries of policies, focusing on Fisher’s (1987) contrasting of two paradigms: rational choice: actors develop preferences by evaluating arguments according to the quality of evidence presented; and narrative: human beings are storytellers with preexisting beliefs who identify what they see as good reasons for preferences from historically and culturally contingent stories. While Honikx (2007) and others have asserted the value of the former, Morgan and colleagues (2002) contend that narrative is more persuasive than statistics. We decided to explore the impact of narrative by adopting adversarial collaboration from behavioural science and adapting it as ‘adversarial co-production’ for policy development.
Our approach sought to build on and invert Greve’s (2020) evaluation of evidence of impact from narratives on perceptions of welfare policy via ‘myths’ that link people’s often genuine concerns for their material interests to often self-defeating policy appraisals. This phenomenon is apparent in the opposition of many of the poorest in society to taxing the richest (Shapiro, 2002) or with support for inequalities justified through skill and opportunity in the ‘American Dream’ being negatively correlated with socioeconomic status (SES) (Hochschild, 1995). We noted evidence of narratives being deployed successfully to highlight people’s genuine material interests, with impact on lower-SES groups increasing support for redistributive policies that enhance resources (Piff et al, 2020). Successful narratives ‘help each other see from different perspectives’ (Stone, 2011), invoking interests that align with policy content.

As such, we formed a prolific.co panel (n = 858) in which participants were asked to evaluate BI based on researcher-produced presentation of its evidenced role in pandemic public health and economic security. Overall levels of support on a 100-point scale were higher than expected (mean 71.99, s.d. 26.45), with older respondents providing higher levels of support for the health narrative and younger respondents the economic narrative. We then asked a subset (n = 20) of respondents who had stated strong (≤30 on a 100-scale) opposition to BI to generate narratives to persuade those like them to support the policy (July 2021). These narratives were synthesised into six distinct narratives and tested (August–September 2021) in a second prolific panel, which was formed by identifying 105 strong opponents of BI (≤30 on hundred point scale) from 677 different prolific.co members from ‘red wall’ constituencies in two of the four UK jurisdictions in Wales and the North and Midlands of England. Participants were asked to rate the adversarially co-produced narratives and to rate BI overall following presentation of all six narratives. Post-study support for BI was significantly higher than pre-study support (mean 46.99, s.d. 28.60, compared to pre-study mean 15.56, s.d. 10.25). This represents a mean increase of 31.43 points on the 100-point scale (t = 27.07, p < 0.001). Accordingly, participants reported that their views on BI had been substantially affected by the arguments they had read (mean 47.85, s.d. 30.43). Moreover, the more they felt their views had been affected, the greater their increase in support (r = 0.60, p < 0.001) (Johnson et al, 2023). The narratives were then deployed in a separate study on the relationship of material conditions, mental health and faith in government to perception of welfare reform (Johnson et al, 2022). Again, the narratives demonstrated significant impacts on perception of welfare reform and illustrated specific challenges for policymakers in persuading beneficiaries of the possibility of reform. This stands at odds with any working assumptions regarding fixed values and the Overton Window (see Lehman, 2010). If there is an Overton Window, it is one that can be opened, extended or broken by persuasion. Support for Brexit, which was once a fringe belief, is evidence of this. Persuasion, in this case, is co-produced by policy researchers presenting evidence and opponents presenting narratives, which are then synthesised. While we co-produced with opponents of a policy to which there is assumed widespread opposition, it is perfectly possible that proponents of a policy could be engaged by opponents to co-produce narratives of opposition.

Adversarial co-production can be conceived of as in-project innovation or innovation through adaptation that meets principles of co-production with regard to collaboration and knowledge co-production (Bannister and Hardill, 2015). There is significant need, first, for examination of the scope of adaptation to different
policymaking fields; and second, for methodological inquiry into the effects of adversarial co-production on individual and group perception of policy and behaviour in distinct fields.

**Why does it work? The social nature of reason**

One question that springs to mind is: why should adversarial co-production work? Surely, people can use their reason to figure out their position on a question or policy; having done so, it is hard to see why being exposed to the reasoning for a different conclusion should make any difference. If they thought that conclusion was sound or congenial, they would have arrived at it independently. This conclusion draws on an intuitive, but outdated conception of the human capacity to reason. Traditionally, scholars imagined that the function of reasoning was *intra*-personal: a capacity that allows individual minds to work through complex questions and reach true conclusions. A more modern view is that the function of reasoning is primarily *inter*-personal or argumentative (Mercier and Sperber, 2011): reasoning works dialogically, through persuasion, to allow multiple individuals, whose interests may differ, to coordinate on joint action.

In accordance with the interpersonal view, people tend to produce very weak arguments for beliefs they already hold, until they are challenged through dialogue by others who hold different positions (Trouche et al, 2016). When challenged, they either produce better arguments to persuade their interlocutors, or change their minds when it becomes clear that there are grounds for the network to settle on a different course. This research demonstrates the strengths and weaknesses of human rationality. When individuals have comfortable or intuitive positions to hand, their reasoning is often one-sided and superficial. Through argumentative processes, lay people can create jointly deep and nuanced rationales for courses of action (Trouche et al, 2014). Experts and non-experts alike grossly underestimate the extent to which discussion in groups can improve the quality of reasoning (Mercier et al, 2015). This psychological research provides the foundations for contemporary experiments with deliberative democracy (see Mouffe, 1999; Mercier and Landemore, 2012), and grants real scope for adversarial co-production of narratives on transformative EBP.

**Risks of co-producing with adversaries**

There is, though, real need to consider the risks of the approach. There is at least the possibility that those with extreme and outlying views who oppose EBP might produce distorting or deliberately damaging narratives capable of undermining support overall. Indeed, there are various extremist groups who have an interest in advancing policy that actively discriminates against others and exacerbates crisis. There is real reason to question whether climate change sceptics or eugenicists should have a voice in policy presentation. It is precisely this phenomenon that motivates our interest in adversarial co-production. There is a recognised need to overcome ‘echo chambers’ that provide potentially misleading understandings of people’s positions. It is not that adversarial co-production advances policies supported by climate change deniers or eugenicists, but that engagement with those who oppose EBP enables identification of salient elements of evidence and prospective impact to persuade others who hold similar positions. This is a crucial means of engaging citizens who
are currently alienated from the technocratic approach criticised by Bandola-Gill and colleagues (2023a) and who need persuading that the policies are not elite abstractions or impositions. This is crucially important for left-behind communities, whose voices are silent or go unacknowledged and unheard, and who are most in need of intervention, and demonstrates the genuine reflexivity advanced by Bandola-Gill and colleagues (2023a).

Indeed, the risks of adversaries distorting policies can be mitigated during the research process outlined above by virtue of qualitative feedback among the broader population, as well as among fellow opponents. The greater risk, at present, is that those with such positions continue to be alienated from policies that have the potential to benefit them as members of society. With the growth in echo chambers and polarisation, the need for controlled co-production of narratives offers a means of addressing much broader social pathologies. This may be consistent with addressing epistemic injustice (see Mouffe, 1999; Chambers et al, 2022), but it is driven primarily by instrumental concern for finding means of persuading on policies with transformative impact.

Conclusion

Adversarial co-production offers potential impact in a number of policy areas. Pilot work on BI ought to serve as encouragement for colleagues committed to EBPs currently regarded as lying outside the Overton Window, but for which there is a fundamental need. There is an urgent need within policymaking circles in particular to grasp the reality that transformative policy, though expensive, is considerably cheaper than the alternative: mass social disintegration and climate-based extinction. Given Labour's express belief that essential EBP, such as wealth tax and an adequate Green New Deal, are not feasible due to the need to appeal to older, wealthy voters, there is a genuine need to understand a truism and make the case for transformative policy through persuasion: as Labour's forebears in 1945 recognised, sometimes the most expensive policies, such as the creation of the NHS, are so essential that society cannot afford not to implement them; the cost of pursuing policies that seem sensible by virtue of their being minimal and uncontroversial, but which are ineffective, is far greater. Adversarial co-production gives policymakers the ability to engage with firm opponents to come up with means of making policy capable of addressing our permacrisis; taking people’s opinions at face value and treating the Overton Window as fixed does not.

Note

1 Some colleagues now deploy the notion of evidence-informed policymaking (for example, CPI, 2018), but evidence-based policy remains the term preferred by the UK Government and by many colleagues, including ourselves.

Funding

This work was supported by the Wellcome Trust under Grant 223553/Z/21/Z. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.

Research ethics statement

The empirical study of public preferences regarding basic income was approved by the Faculty of Medical Sciences ethics committee, Newcastle University.
Contributor statement
MJ and EJ wrote the first and subsequent drafts of the manuscript, with comments from IH and DN. DN, EJ, IH and MJ conceptualised and designed the study. DN and EJ conducted data analysis and interpretation, with contributions from MJ and EJ.

Conflict of interest
The authors declare that there is no conflict of interest.

References
Burn-Murdoch, J. (2022) Britain and the US are poor societies with some very rich people, *Financial Times*, 16 September. https://www.ft.com/content/ef265420-45e8-497b-b308-c951baa68945.


