

Perspectives of Paternalism in a Democratic Society

22–23 October 2021, Palacký University Olomouc, Czechia

online: <https://cesnet.zoom.us/j/97766045287>

On the Behavioral Political Economy of Paternalism

Christian Schubert (The German University in Cairo)

dr.c.schubert@googlemail.com

Most accounts of paternalism – especially those defending ‘libertarian paternalism’ – seem to treat the democratic political system itself as a black box: The government is typically understood as both benevolent and rational. This research strategy is of course understandable, but it limits the relevance of the insights gained thereby. I suggest to open up this black box, in two steps: First, those who make policy are self-interested; second, their voters are subject to all the kinds of biases that behavioral economics has uncovered so far. While the first step leads to a conventional Public Choice critique of paternalistic policies, the second, more interesting one, leads to a behavioral political economy view of paternalism, following Schnellenbach and Schubert (2015).

What does such a view imply? When we combine self-interested policy-makers and biased voters, we may find that choosing, e.g., ‘green nudges’ over more effective environmental policies might be attractive in simulating activity and engaging in symbolic policies that feel good, despite being largely ineffective. More generally, nudging may provide an attractive way to introduce all kinds of manipulative techniques into the political arena, thereby facilitating rent-seeking (Schubert 2017).

When we focus on biased voters, we may enter more controversial territory, discussing the question to what degree democracies should be organized as representative (as opposed to direct) democracies. Put differently, we have to talk about the rules of the (political) game, not the game itself. Since the very first implementations of modern democracies in the late 18th century, there was always the worry that “the people” (i.e. the entity seen as the new sovereign) might encompass individuals lacking the minimal requirements to vote ‘responsibly’. Constitutional checks and ‘independent’ entities such as central banks are supposed to maintain a minimum degree of rationality in collective decision-making. In other words, democratic theory has always been paternalistic at heart – in the peculiar sense of worrying not so much about individuals’ errors and biases but about collective follies. Recently, these worries have resurfaced with democratic choices that resulted in the Brexit or the election of Donald Trump, choices that arguably did not improve the welfare of the average UK or U.S. citizen, respectively. Instead, they arguably allowed powerful minorities to indulge in pleasurable ‘belief consumption’ (e.g. Bénabou/Tirole 2016). In the wake of this, polarization and epistemic issues (the spread of fake news, growing distrust in experts,...) have soared. If one argues that the political arena should not be used for purposes of belief consumption or entertainment (Maçães 2020), then a case may be made in favor of restricting the scope of direct democracy, and to strengthen the representative elements of democracy (Jones 2020).

Liberal Parentalism

Aviad Heifetz (Open University of Israel)

Enrico Minelli (University of Brescia)

Herakles Polemarchakis (University of Warwick)

aviad.heifetz@gmail.com

When a person makes sequential choices x_1, \dots, x_n in periods $i = 1, \dots, n$ towards a pre-scheduled target, as, for example, when saving for retirement, in each period i the person controls only the choice x_i at that period: even if the choice x_i induces strong incentives for future choices x_j in periods $j > i$, as is the case when investing in a pension fund with hefty fines for early withdrawal, still in period i the person cannot technically *commit* for the future choice x_j .

Yet, in period i the person may rejoice or worry towards his *anticipated* future choices x_j in periods $j > i$, and also like or regret choices x_k he has already made in the previous periods $k < i$. It is therefore natural to assume that even though in period i the person controls only the contemporaneous choice x_i , he has a utility function u_i over *entire* choice sequences (x_1, \dots, x_n) such that u_i represents his preferences, i.e. $u_i(x_1, \dots, x_n) \geq u_i(x'_1, \dots, x'_n)$ when and only when in period i the person weakly prefers the sequence (x_1, \dots, x_n) over (x'_1, \dots, x'_n) .

The person can thus anticipate the optimal choice $x_n^*(x_1, \dots, x_{n-1})$ he will be making in period n with u_n given any possible history (x_1, \dots, x_{n-1}) ; then deduce the optimal choice $x_{n-1}^*(x_1, \dots, x_{n-2})$ he will be making in period $n - 1$ with u_{n-1} given any possible history (x_1, \dots, x_{n-2}) and the anticipation of the reaction $x_n^*(x_1, \dots, x_{n-2}, x_{n-1}^*(x_1, \dots, x_{n-2}))$ to it in period n ; etcetera, by Backward Induction (BI). The resulting Backward Induction path is then

$$\hat{x} = (\hat{x}_1, \hat{x}_2, \hat{x}_3, \dots) = (x_1^*, x_2^*(x_1^*), x_3^*(x_1^*, x_2^*(x_1^*)) \dots)$$

In case the person's preferences do not change along time, in which case $u_i(x_1, \dots, x_n) \geq u_i(x'_1, \dots, x'_n)$ if and only if $u_j(x_1, \dots, x_n) \geq u_j(x'_1, \dots, x'_n)$ also in every other period $j \neq i$, then it is well known that by the *Principle of Optimality* in dynamic programming, even if the person were to have commitment power to choose the *entire* sequence (x_1, \dots, x_n) right

from the very start, in period 1, he would still be choosing the BI path \hat{x} that emerges without such commitment power.

But what if the person's preferences do change along time? Can a benevolent and liberal parental figure or entity, who has the capability to *commit in advance* to an entire sequence $\check{x} = (\check{x}_1, \dots, \check{x}_n)$, do so in such a way that the person will weakly prefer \check{x} over \hat{x} in each and every period, and strictly prefer \check{x} over \hat{x} in at least some period? Using a transversality argument, we prove that, somewhat surprisingly, the answer is 'yes' for almost every profile of utility functions u_1, \dots, u_n .

This optimistic view on the possibility of a liberal parental intervention changes dramatically when the person might misperceive his own future preferences, or be uncertain about them. His state of mind τ_i in period i then consists not only of a utility function, u_{τ_i} , but also of a *belief* β_{τ_i} about his future states of mind $(\tilde{\tau}_{i+1}, \dots, \tilde{\tau}_n)$, each with its own utility function $u_{\tilde{\tau}_j}$ and belief $\beta_{\tilde{\tau}_j}$ about *its* future states of mind $(\tilde{\tau}_{j+1}, \dots, \tilde{\tau}_n)$. We show that there are states of mind (τ_1, \dots, τ_n) for which no committed parental intervention can improve wellbeing in some periods without jeopardizing wellbeing in others, and that even informational *nudges*, that only influence the *belief* in some period, might be painfully sobering. For such situations, we propose a novel, minimally illiberal normative criterion, by which the parent, when she aims at improving average/aggregate wellbeing along time, should only intervene in ways that *minimize the maximal disappointment* across the time periods.

Epistemic Paternalism in Times of Crises

Ivana Janković & Miljan Vasić (University of Belgrade)

miljan.vasic@f.bg.ac.rs

Epistemic paternalism (EP) is the practice of interfering in the process of inquiry of another, without prior consultation and for the sake of her epistemic good (Ahlstrom-Vij, 2013). We will examine two main types of EP: eudaimonic and strict (Bullock, 2018). In the case of eudaimonic EP, epistemic improvement is used only as a means to achieve non-epistemic benefits. When it comes to strict EP, epistemic improvement is valued *per se*. Emma Bullock criticizes both of these types. She claims that, on the one hand, eudaimonic EP collapses into general paternalism (thereby making the term “epistemic” redundant), while, on the other, strict EP is unjustified since epistemic values fail to outweigh violations of personal autonomy.

In this talk, we will show that both of Bullock’s conclusions can be refuted. Firstly, we will show that, regardless of whether our interference aims to produce epistemic or non-epistemic benefits, the condition of epistemic improvement itself provides a clear distinction between EP and general paternalism. Namely, in some cases of EP, the epistemic improvement in question comes down to mere concealment of one’s epistemic shortcomings, but it is an epistemic improvement, nevertheless. Moreover, if the motivation behind interference was non-epistemic, this is an instance of eudaimonic EP and thereby distinctive from cases of strict EP. Secondly, we will argue that epistemic improvements can advance one’s personal autonomy (rather than violating it) by improving specific epistemic virtues, such as reliability, good judgment, responsibility, and the ability to make rational decisions (cf. Croce, 2020).

The question of justification of EP becomes even more pronounced in times of crisis, such as the current pandemic. People are more exposed to pseudo-scientific information, fake news, and unverified content; as a consequence, public health is being compromised (Lechanoine & Gangi, 2020). In situations like this, external epistemic intervention (private or public) might be necessary to foster better decision-making. Given all this, we will defend a moderate form of strict EP which is motivated by both epistemic and non-epistemic benefits. Unlike the cases of general paternalism (e.g. mandatory vaccination), EP is a more plausible position since it is less likely to provoke public outrage. On the other hand, this moderate position is also superior to more radical forms of strict EP that completely neglect non-epistemic benefits that should be our primary concern during a crisis.

Bibliography:

Ahlstrom-Vij, K. (2013). *Epistemic Paternalism: A Defence*. Basingstoke: Palgrave Macmillan.

Bullock, E. (2018). “Knowing and Not-Knowing for Your Own Good: The Limits of Epistemic Paternalism”. *Journal of Applied Philosophy*. 35 (2). 433-447.

Croce, M. (2020). “Epistemic Paternalism, Personal Sovereignty, and One’s Own Good”. *Epistemic Paternalism: Conceptions, Justifications and Implications* (Axtel, G. & Bernal A. eds.) London, New York: Rowman & Littlefield International. 199-217.

Lechanoine, F & Gangi, K. (2020). “COVID-19: Pandemic of Cognitive Biases Impacting Human Behaviors and Decision-Making of Public Health Policies”. *Frontiers in Public Health* (24).

<https://doi.org/10.3389/fpubh.2020.613290>

The Conceit of Behaviorally Informed Paternalism and the Anti-Psychological State

Petr Špecián (Prague University of Economics and Business; Charles University; IREF)

petr.specian@vse.cz

If people end up harming themselves in predictable ways once left to their own devices—even worse: if the imperfections of their rationality are being commercially exploited— how should the government respond? Behaviorally informed paternalists offer an increasingly popular answer. Namely, people need to be nudged (or shoved, if necessary) towards better choices. To arm the government for a battle with “internalities,” the paternalists have gotten busy designing interventions to alleviate the burden of rationality’s bounds. Among them, those of the “libertarian” streak achieved the greatest notoriety for offering cheap, discrete, and non-coercive fixes to irrationality, which ostensibly make people “better off as judged by themselves.”

But how about the compatibility of paternalism with democracy? Although the paternalists’ commitment to value non-imposition via subscription to the “as judged by themselves” (AJBT) standard of welfare improvement is laudable, its persuasive implementation is lacking. Paternalists’ efforts to address irrationality in consumption overload the political channel through which people’s preferences are transmitted. With no transparent way of how the AJBT standard could be upheld and people’s “true” preferences identified, paternalists end up relying on discretion, not rules. Therefore, implementation of paternalist policies—which tend to take the shape of precisely targeted micro-interventions and meticulously crafted choice architectures—means either that democratic citizens must be extremely watchful to detect potential misuse or that the policymakers’ benevolence must be strongly presumed. Given what we know about human political behavior, the former road appears impassable and the latter perilous, especially in the stormy weather democracy currently experiences.

Inadequacy of paternalist solutions to irrationality does not imply that its threat is not a serious one, however. Indeed, liberal democracy’s ability to keep political manifestations of irrationality within manageable proportions may be faltering. Against this background, I argue that democracy’s resilience could be reinforced by implementing what I call an Anti-Psychological State: a general label for an institutional design crafted with the specific intent to discourage political irrationality while not sacrificing democratic responsiveness.

The concept of the Anti-Psychological State draws on several sources of inspiration. The first is mechanism design: proper incentives must be created to motivate the people to use their best abilities in service of democratic collective choice and to feed the aggregation mechanism as much reliable information—inevitably scattered across many minds—as possible. The second is behavioral science: the institutional design must employ experimental methods, and it must be undertaken in acute awareness of the current state of knowledge regarding the quirks and limitations of human reason. The third source of inspiration is democratic theory since it delivers original and audacious visions of possible reform.

Accordingly, my example of how an Anti-Psychological State might be set up combines the radical proposal for “open democracy” with the mechanism design insights of the quadratic voting and the behaviorally informed notion of boosts to decision-makers’ competencies. I also demonstrate that these proposals show remarkable complementarity: the shortcomings of each building block are neatly compensated by the other blocks’ strengths.

What to Do, if Anything, About Conspiracy Theories?

M R. X. Dentith (Beijing Normal University)

m.dentith@mrxdentith.com

Interest in conspiracy theories and the problem of belief in unjustified conspiracy theories has grown in recent years. This has led to a plethora of diagnoses and potential cures or solutions to the problem of conspiracist ideation in contemporary public discourse. Yet there are two lingering worries about the supposed problem of conspiracy theories: just how big the problem of belief in conspiracy theories is (i.e. are they really that popular?) and are we capturing how most of the people who believe in conspiracy theories actually think and talk about them?

I will argue that conspiracy theories are not the problem we either believe them to be or are told they are. Rather, we should frame the debate around conspiracy theories in terms of both the evidence, and how we should define both what counts as a 'conspiracy theory' and who the 'conspiracy theorists' are. When this is taken into account, then many of the diagnoses and potential cures and solutions to the problem of both conspiracy theories and conspiracy theorists in public discourse seem mistaken— and possibly even sinister—as a consequence.

Consequently, we ought to consider conspiracy theories with a different frame: I will argue we should start our analysis with the proviso that we treat conspiracy theories seriously, and investigate to see if they are warranted. However, I will also argue that this does not mean we have to treat each and every conspiracy theory we encounter as *prima facie* plausible. Rather, there are certain heuristics or features we find amongst certain conspiracy theories which warrant a suspicion of such theories, which in turn allows us to prioritise which conspiracy theories we should investigate now versus those we can analyse later.

Regulation of Digital Media and Cyber Communications in India: An Archaic Move

Bhumika Sharma (Himachal Pradesh University, Shimla)

s24bhumika@gmail.com

Until 2000, only indirect legislations governed the cyberspace in India. The cyber law framework in India has emerged over the last 20 years. Section 87 (2) of the Information Technology Act, 2000 empowers the Government to enact rules under the Act. Since 2000, a series of Rules have been made by the Government from time to time in accordance with the social and global developments. The Act of 2000 read with its Allied Rules authorise the government to regulate, monitor and intercept the communications and messages. The authorities are also permitted to give directions to the intermediaries in this regard. In 2011, the Information Technology (Intermediaries Guidelines) Rules, 2011 were made to regulate intermediaries in India. The Information Technology (Intermediary Guidelines and Digital Media Ethics Code) Rules, 2021 have enhanced the entities covered earlier under the Rules, 2011. Part - II of these Rules are administered by Ministry of Electronics and IT. Part - III relating to Code of Ethics and procedure and safeguards in relation to digital media is administered by the Ministry of Information and Broadcasting. Part II brought the online curated content platforms within the purview of the regulation. Similarly, the publishers of digital news portals also fall within these new Rules by Part III. They further focus more on in house and self-regulation mechanism whereby a robust grievance redressal mechanism. The OTT platforms, called as the publishers of online curated content in the rules, would self-classify the content into five age-based categories - U (Universal), U/A 7+, U/A 13+, U/A 16+, and A (Adult). Platforms would be required to implement parental locks for content classified as U/A 13+ or higher, and reliable age verification mechanisms for content classified as "A". The publisher of online curated content shall prominently display the classification rating specific to each content or programme together with a content descriptor informing the user about the nature of the content, and advising on viewer description (if applicable) at the beginning of every programme enabling the user to make an informed decision, prior to watching the programme. These new Rules show how freedom of speech of the media in the digital age has been damaged by the Government. Content Regulation has become stricter now. It is thus the aim of this paper to evaluate how the evolving legal framework for the cyber and digital space in India is an example of paternalism. It will attempt to address that how far the government is justified in imposing restrictions on the cyberspace on the ground of regulation.

A Hobbesian Remedy for the Polarizing Effects of Availability Cascades on Social Media

Tommaso Ostillio (Kozminski University; University of Warsaw)

Giulio Sciacca (University of Genoa; FINO Consortium)

tostillio@kozminski.edu.pl

Several studies have lately revealed that social media conceal at least three pitfalls. Firstly, social media can negatively impact sociopolitical processes by becoming vehicles of incorrect information that augments political polarization (Lee et al. 2017; Ostillio 2018). Secondly, social media can quickly become a source of incorrect beliefs for those individuals featuring low digital literacy (Guess et al. 2019). Thirdly, intentional and deceitful manipulations of social media content can generate information cascades that increase group polarization across defined groups of social media users (Jang and Kim 2018, Wang et al. 2018, Colliander 2019). Importantly, all these findings corroborate Duffy's (2018) and Sunstein's (2018) claim that the circulation of biased content on social media may become a potent driver of sociopolitical instability in Western democracies.

Nonetheless, if Duffy's (2018) and Sunstein's (2018) claim is – as it seems - correct, then a thorny dilemma arises: should Western democratic regimes introduce severe penalties for those who spread fake news and fabrication on social media? Or should they let the market autonomously provide efficient technology that allows people to identify biased social media content? More concisely, is a libertarian solution to the plight of fake news and fabrication on social media preferable to adequate legislation that punishes those who spread such content?

Our paper answers this question negatively and contends that governments should instead engage in a merciless battle against biased social media content.

To prove this contention, we resort to Hobbes's (1651, part II, §18 and §29) claim that humans can avoid falling back into the Hobbesian state of war only if they stipulate a covenant that, among other things, grants the sovereign the right to ban those seditious doctrines that threaten peace and sociopolitical stability in the Hobbesian commonwealth. Specifically, we show that the Hobbesian covenant resolves conflict in the same way as correlated equilibria can resolve conflict in non-cooperative games. Namely, we show that citizens, like game theory players, enjoy greater welfare whenever they observe rules that mitigate the risk of social conflict effectively.

Accordingly, Western citizens would better off if Western governments introduced severe penalties for those who intentionally spread fake news and biased content that goes viral. That is because efficacious regulation would create disincentives for spreading biased content on social media. Besides, the latter would assuage the polarizing effects of availability cascades on social networks. Notably, these claims find strong support in the fact that the American government failed to provide adequate policy response to the two first two waves of the Covid19 pandemic because it failed to contain the polarizing effects of the large-scale infodemic (i.e., the epidemic of false information) about covid19 on social media (Evanega et al. 2020, Milosh et al. 2020). On these grounds, we conclude that Hobbes's (1651) teaching about seditious doctrines is valuable because it reminds us that only effective monitoring and regulation of social media content can obviate the adverse effects of biased content on Western societies and Western public institutions.

Bibliography:

Colliander, Jonas (2019): "‘This is fake news’: Investigating the role of conformity to other users' views when commenting on and spreading disinformation in social media". *Computers in Human Behavior* 97, p. 202-215.

Duffy, Bobby. *The perils of perception: Why we're wrong about nearly everything*. Atlantic Books, 2018.

Evanega, Sarah, Mark Lynas, Jordan Adams, Karinne Smolenyak, and Cision Global Insights. 2020. "Coronavirus misinformation: quantifying sources and themes in the COVID-19 'infodemic'." *New York Times*. Accessed December 22, 2020.

Guess, Andrew, Nagler, Jonathan, Tucker, Joshua (2019): "Less than you think: Prevalence and predictors of fake news dissemination on Facebook". In: *Science advances* (02 January 2019) 5, No. 1. Washington: AAAS. DOI: 10.1126/sciadv.aau4586, visited on 10 November 2020.

Hobbes, Thomas. (1651) 2017. *Leviathan*. Early Modern Texts. (Source: <http://www.earlymoderntexts.com/authors/hobbes>)

Jang, S. Mo and Kim, Joon Kyoung (2018): "Third person effects of fake news: Fake news regulation and media literacy interventions". *Computers in Human Behavior* 80, p. 295- 302.

Lee, Sun Kyong, Lindsey, Nathan J., Kim, Kyun Soo (2017): "The effects of news consumption via social media and news information overload on perceptions of journalistic norms and practices". *Computers in Human Behavior* 75, p. 254-263.

Milosh, Maria, Marcus Painter, Konstantin Sonin, David Van Dijcke, and Austin L. Wright. 2020. "Political polarisation impedes the public policy response to COVID-19." *VoxEU*. Accessed December 27, 2020. <https://voxeu.org/article/political-polarisation-impedes-public-policy-response-covid-19>

Ostillio, Tommaso (2018): "The Trigger Effect: Cognitive Biases and Fake News". *Internetowy Magazyn Filozoficzny Hybris* 44, No. 1, p. 86-104.

Sunstein, Cass R. # republic. Princeton University Press, 2018.

Wang, Quansheng/Yang, Xue/Xi, Wanyu (2018): "Effects of group arguments on rumor belief and transmission in online communities: An information cascade and group polarization perspective". *Information & Management* 55, No. 4, p. 441-449.

Institutionalizing Micro-Targeting? Using Big Data to Strengthen Democracy

Catalina Carpan (University of Warwick)

c.s.carpan@warwick.ac.uk

Currently, the academic literature on political micro-targeting has focused primarily on conceptualizing it as a threat to privacy and democracy, and on making recommendations for how it should be regulated. While regulatory solutions are necessary, they are not sufficient, because of technical, legislative and social challenges such as a lack of transparency, limited expertise on rapidly evolving digital behavioural strategies, and poor digital competency levels among citizens.

Drawing on normative theory, my paper contributes to this conversation by framing micro-targeting as a potential tool for strengthening democracy, if its power is harnessed in the right kind of way. Compared to other forms of campaigning and traditional nudge interventions, micro-targeting is remarkably effective due to its reliance on Big Data analysis to determine voters' most susceptible triggers, personalized interventions and online choice environments that facilitate hyper-nudging. To bring into sharp relief more positive or democratic uses of micro-targeting, it is helpful to think in terms of specific institutional innovations and their possible functions. Embracing the challenges of institutional design in this context, I propose the establishment of a Commission for Micro-Targeting that would use the practice in order to (i) build citizen resilience against online nefarious content and manipulation, and (ii) where possible, address long-standing democratic challenges such as rational ignorance and the uneven distribution of political knowledge, interest and competency. I also discuss the possibility that these interventions may be used to steer voters away from voting based on racist beliefs through personalized de-biasing techniques. Already, there is an emerging literature on how traditional nudging may be used to stop the spread of disinformation, encourage people to inform themselves or to vote. Against this background, my proposal for an independent body that would conduct micro-targeting in order to achieve these good outcomes, appears to be the next logical step. The Commission would first set clear competency-based, behavioural and attitudinal goals to work towards, such as building resilience against manipulation, dis/misinformation and other nefarious content, on the one hand, an encouraging citizens to inform themselves, consider opposing points of view and reason in the right kind of way. Second, the Commission would use Big Data combinatorial analysis to sift through legitimately acquired data points in order to identify the most susceptible triggers for each citizens. Third, behavioural experts would help design interventions for each cluster of citizens towards the desired goals; these interventions would range from rational persuasion strategies, to informational nudges – those that either provide more information or sway people to become better informed – and to non-informational nudges, which aim to sway behaviour in ways that do not depend on informing the nudgee. Finally, each citizen will be targeted with the optimum intervention in a way that harnesses the algorithms of social networks, which are designed to maximize effectiveness. In the final section of the paper, I spell out criteria that should be met by institutionalized micro-targeting interventions: that they should be non-manipulative, resistible, and respect the transparency and publicity requirements.

The Ethics of Boost

Till Grüne-Yanoff (Royal Institute of Technology, Stockholm)

gryne@kth.se

In recent publications, we have characterized boosting as a type of behavioral intervention that differs in important ways from its famous cousin, nudging. In particular, we have defined Boosts as policymakers' effortful interventions on groups of people with the aim of improving people's competences and through that improve their decisions according to their own standards (Hertwig & Grüne-Yanoff 2017). In this presentation, I will explore the characteristics of boosts along a number of normative dimensions, occasionally contrasting them to the normative properties of nudges. Specifically, I want to answer real or potential critics who criticize behavioral interventions and boosts along the following normative dimensions: for limiting people's freedom; for undermining people's autonomy; for being paternalistic; and for occasionally increasing inequalities between people. I will argue (i) that the criticism of boosts in any of dimensions is rarely justified, and (ii) where it is justified, boosts do better than nudges.

Hard Paternalism and Familism

Andrew T. W. Hung (The Hong Kong Polytechnic University)

andrew.hung@cpce-polyu.edu.hk

The regulations of use of drugs, pornography, prostitution and gambling, which are considered as self-regarding activities, have been controversial for a long time in liberal society. On the one hand, even if we assume that these activities are harmful, liberals argue that to prohibit these activities is violating Mill's harm principle; it means that they are harmful to the self only, they are not activities that harm others. On the other hand, the paternalists argue that the responsibility of the state is not only to protect individual liberty, but also to promote human well-being; at least, the state should protect people from harming themselves. In order to integrate liberalism and paternalism, Joel Feinberg argues for soft paternalism which argues that it is permissible to intervene one's action if one's action is not fully voluntary. However, the criterion of voluntary is also very controversial, as how can the others make a judgement that one's action is involuntary if one acts consciously and deliberately. This paper attempts to defend hard paternalism and argue that it is legitimate, at least sometimes when the harm is serious and irreversible, to prohibit certain activities for protecting the person no matter if one is voluntary or not. This paper also argues that in addition to hard paternalism, regulating these activities is also based on familism, because these activities are not really self-regarding, in many cases they are harming their families and finally the society.

The Ethics of Market Creation: Changing Behaviour by Adding an Option

Lukas Fuchs (University College London)

lukas.fuchs.17@ucl.ac.uk

The creation of new markets can introduce new options for consumers. Just in the last century, cars, vitamin pills and smartphones were introduced to the market. Introducing new green products into the market, such as green energy or green cars, may induce consumers to change their lifestyle. The creation of markets is not reserved for private entrepreneurs; states, as well as other actors in civil society, can shape and create markets. A new option, which is superior to the existing ones, may change consumers' consumption choices, without the need for coercion, incentive or nudge.

Should policy aim to change citizens' behaviour by creating markets in which new, say healthy or green, products can be purchased? The ethics of changing other people's behaviour is subject to intense debate, particularly when brought about through public policy. The ethics of coercion, incentive and nudge has been discussed extensively in the literature. This paper, by contrast, will analyse the ethics of adding an option as a mechanism for behavioural change. May states change individual behaviour by adding options in the market? What are the constraints of such policy? How does the ethical status of this mechanism compare to coercion, incentive and nudge?

This paper addresses worries that could be raised against market creation. After introducing adding an option and market creation as mechanisms of behavioural change, the paper compares adding an option with the other forms of behavioural change on different evaluative criteria, namely freedom, autonomy and paternalism. The overall argument is that adding an option is comparably soft on these parameters.

“Don’t Tell Me What to Do”: Nudges, Autonomy, and The Impossibility of Libertarian Paternalism

Ezechiel Thibaud (Lingnan University)

ezechieltthibaud@ln.hk

The question of the potentially disruptive nature of behavior change technologies – or nudges – and their impact on our moral abilities is crucial in a world where moral assistance seems to become more and more standard. Although often presented as freedom-preserving tools, nudges raise the legitimate question of the tension between external influence and autonomy. Nudges are described by their advocates (Thaler & Sunstein 2008) as an instance of libertarian paternalism. This paper looks at the libertarian paternalist claim for nudges and argues that the argument is self-defeating. It also explores the possibility that conceiving nudges as libertarian- paternalistic puts us at risk of excessive individualism: nudges become the superficial plaster aiming at fixing the deep social and political problems of collectively informed decisions, public involvement, and education.

Although these may appear as conflicting concepts, proponents of nudges have managed to conflate them, presenting these tools as simple choice architecture processes, aiming at gently pushing agents in the direction of what they may fail to recognize as their own good, and this without coercion or control. The claim for paternalism relies on the idea that agents often lack the rational capacities to identify what the best choice could be, and therefore need assistance. But for a nudge to remain “libertarian”, the condition of easy resistibility must be met (Saghai 2013). This condition implies that agents must be able to recognize a nudge and resist the temptation to follow the nudge’s lead, based on an informed desire evaluation. Nudges also aim at triggering what is referred to as System 1 mechanisms: cognitive activities that are faster, more automatic, and less conscious than System 2 mechanisms – based on slow, rational, and meticulous evaluations.

In this paper, I argue that the libertarian-paternalist aspect of nudges fails on two accounts: First, given the conditions for nudges presented by their proponents, the concept of libertarian paternalism becomes self-defeating: if the need for nudges is justified by agents’ lack of rational capacities to identify their best interest, it becomes unclear how they will be able to display the rational capacities required to resist the influence of the nudge – condition for its “libertarianism”.

Second, I argue that the libertarian aspect of nudges emphasizes the wrong problem. By focusing on rectifying individual choice architecture, nudges prevent us from looking at the deeper question of how agents can come to make informed choices in the first place. This, I argue, requires that we look at the collective and political aspects of choice architecture and place the emphasis on the promotion of autonomy.

Do We Live in a 'Post-Truth' Era?

Kristoffer Ahlstrom-Vij (Birkbeck College, University of London)

k.ahlstrom-vij@bbk.ac.uk

Have we entered a 'post-truth' era? This article is an attempt to answer this question by (a) offering an explication of the notion of 'post-truth' from recent discussions, (b) deriving a testable implication from that explication, to the effect that we should expect to see decreasing information effects – that is, differences between actual preferences and estimated, fully informed preferences – on central political issues over time and then (c) putting the relevant narrative to the test by way of counterfactual modelling, using election year data for the period of 2004–2016 from the American National Election Studies' Times Series Study. The implication in question turns out to be consistent with the data: at least in a US context, we do see evidence of a decrease in information effects on key, political issues – immigration, same-sex adoption and gun laws, in particular – in the period 2004–2016. This offers some novel, empirical evidence for the 'post-truth' narrative.

Science, Motivation, and Shame: Two Arguments Against Emotional Paternalism

Sarah Malanowski, Ashley Kennedy & Nicholas Baima (Honors College, Florida Atlantic University)
nbaima@fau.edu

In a liberal democracy, scientific beliefs matter, because they influence the way that people act. If someone doesn't believe that a face mask, for example, is an effective tool in preventing the spread of infectious disease, then they will likely balk at a rule that requires them to wear one. Because of this problem of motivating public action, some have recently suggested that epistemic paternalism needs to be implemented (John 2019). However, aside from being potentially morally problematic, it isn't clear that information alone, even of the paternalistic sort, is enough to motivate people to act. In other words, it is rarely enough that someone simply trusts the information that they are given; they must also accept the recommendation to act (Bennett 2020). So, what is needed is something that will motivate the public to act in accordance with policy recommendations, and certain emotions, such as shame, have been leveraged to play this role; we call this emotional paternalism.

Implicitly and explicitly, shame has been utilized by authority figures and institutions to alter the public's behavior. By exposing an individual's or group's undesirable behavior, their reputation is damaged, and thus the individuals or groups in question are (presumably) motivated to abide by social norms. However, while shame can be an effective motivator and thus an effective means to alter behavior, we argue that there are significant practical and moral costs that come with the use of shame as a means to change public behavior.

First, while shame may help to promote behaviors in some, it does not work when the individuals or groups being shamed do not trust or respect the authorities who are doing the shaming. And often, the people who need the most convincing are those who are least likely to have this level of trust and respect. Thus, shaming in these instances is likely only to create further resentment for and backlash to the message being conveyed. We support this claim by drawing upon real-life examples of effective and ineffective uses of shame in public policy. From these examples, we develop two models of public policy shaming: community-led shame and authoritarian shame, and we argue that the latter is a more egregious form of emotional paternalism.

Second, we argue that the use of authoritarian emotional paternalism, in the form of shaming, is morally problematic because it violates the values trust and openness. As such, it violates the aims of a liberal democracy (Nussbaum 2004). We, however, do not think that community-led shame necessarily violates these liberal values. Nevertheless, because community-led shame is small-community-focused, it cannot be used for large-scale public policy, and though it is less authoritarian, it may lead to other moral problems. In sum, we argue that public shaming is not the best way to motivate public behavior, and conclude the paper with some suggestions of alternative ways of promoting science-based behavioral changes that are both efficacious and moral.

Bibliography:

- Aaltola, E. 2021. "Defensive over Climate Change? Climate Shame as a Method of Moral Cultivation." *Journal of Agricultural and Environmental Ethics* 34. <https://doi.org/10.1007/s10806-021-09844>.
- Abaluck, J. et al. Forthcoming. "Normalizing Community Mask-Wearing: A Cluster Randomized Trial in Bangladesh." NBER Working Paper No. 28734. April 2021 JEL No. I12, I14, I15
- Anderson, E. 2011. "Democracy, Public Policy, and Lay Assessments of Scientific Testimony." *Episteme* 8 (2): 144–64.
- Bennett, M. 2020. "Should I do as I'm Told? Trust, Experts, and Covid 19. *Kennedy Institute of Ethics Journal*. Special Issue: <https://kiej.georgetown.edu/trust-experts-and-covid-19-special-issue/>
- Drummond, C. and B. Fischhoff. 2017. "Individuals with Greater Science Literacy and Education have more Polarized Beliefs on Controversial Science Topics. *Proceedings of the National Academy of Sciences of the United States of America* 36: 9587–9592.

- Goldman, A. 2001. "Experts: Which Ones Should You Trust?" *Philosophy and Phenomenological Research* 63: 18–212.
- Jacquet, J. 2015. *Is Shame Necessary: New Uses for an Old Tool*. New York: Vintage.
- . 2017a. "Guilt and Shame in U.S. Climate Change Communication." *Oxford Research Encyclopedia of Climate Science*. DOI: 10.1093/acrefore/9780190228620.013.575
- . 2017b. "Shame." In Szeman, I, J. Wenzel, and P. Yaeger (eds.). *Fueling Culture*. New York: Fordham University Press.
- John, S. 2019. Expertise in Climate Science. In Coady, D and J. Chase. *The Routledge Handbook of Applied Epistemology*: 131–141.
- . 2018. "Epistemic Trust and the Ethics of Science Communication: Against Transparency, Openness, Sincerity and Honesty." *Social Epistemology* 32: 75–87.
- Kahan, D., Braman, D., and Jenkins-Smith H. 2011. "Cultural Cognition of Scientific Consensus." *Journal of Risk Research* 14: 147–74.
- Kahan, D., et al. 2012. "The Polarizing Impact of Science Literacy and Numeracy on Perceived Climate Change Risks." *Nature Climate Change* 2: 732–35
- Mehta, L. and Movik, S. (eds.) 2010. *Shit Matters: The Potential of Community-Led Total Sanitation*. Practical Action Publishing.
- Perez-Truglia, R. and Troiana, U. 2018. Shaming tax delinquents. *Journal of Public Economics*, 167: 120–137.
- Williams, B. 1993. *Shame and Necessity*. University of California Press.

Wellbeing and Respect – Still the Central Values

Kalle Grill (Umeå University)

kalle.grill@umu.se

Traditionally, paternalism has been understood in terms of the value conflict between the promotion of some person's wellbeing and respect for her will or desire, particularly in matters where she has legitimate control or authority. In recent years, however, this traditional understanding has come under pressure both from analysis of group cases and from intense debate on more subtle forms of behavioral influence than the traditional dichotomy between coercion and persuasion.

The conceptual debate on traditional paternalism has struggled to account for cases that involve more than one person in either the role of paternalist or that of paternalized. Paternalism by groups upsets the traditional focus on the promotion of wellbeing as a motivating reason, since groups consist of many people, who may have quite different motives. Paternalism towards groups upsets the traditional concern with promoting one person's good against her will, since the promotion of several people's wellbeing, against their will, need not imply that the interference with each person is what promotes her wellbeing. Talk of "impure" paternalism and "mixed" cases does not sufficiently account for these complexities.

Richard Thaler and Cass Sunstein's has quite successfully branded an allegedly nonintrusive strategy for promoting wellbeing "libertarian paternalism". Paternalism scholars first resisted this term, but many have come to accept it or even endorse it. One reason to do so is that the behavioral influences that Thaler and Sunstein propose, while not interfering in the traditional sense, may be manipulative, or encroaching on autonomy in some other non-coercive way.

I propose that a fruitful reaction to these developments is to stay close to the traditional understanding of paternalism as a value conflict between wellbeing and respect. We typically have reason to promote wellbeing and reason to respect will or desire, in particular concerning matters over which there is legitimate control or authority. While both group cases and new and sophisticated forms of behavioral influence warrant independent investigation, this does not necessarily change what values are fundamentally at stake.

Criticisms of nudging, such as its being manipulative, covert, infantilizing, non-rational or counter-productive, typically boil down to either negative wellbeing effects or else a failure to respect what people want for themselves. If some influence is desired by those effected as well as in their interest, it is rarely morally problematic on their account.

In group cases, if we consider what all the involved parties want for themselves as well as what is in their interest, we can typically arrive at a coherent evaluation of any interference with some or all of them. The gravest problems with group cases stem from a tendency in the debate to try to determine when exactly some action is paternalistic, though this need not concern us once we know how to morally evaluate that action.

In conclusion, I propose that wellbeing and respect are still the two values that make cases of paternalism, whether traditional or libertarian, so difficult, interesting, and important.

The Alleged Rationality of Conspiratorial Thinking

Filip Tvrđý (Palacký University, Olomouc)

filip.tvrdy@upol.cz

Discussions in contemporary epistemology are burdened with confusions about the terms “rational”, “rationality” and their antonyms. In economy, for an agent to be rational simply means to satisfy the Bayesian probability axioms, but the situation in philosophy is much more complicated. Two kinds of rationality are usually distinguished. Epistemic rationality is an ability to achieve justified and true beliefs, whereas instrumental rationality is a capacity to act in accordance with one’s own interests. This division cleared the way to contemplation about rational irrationality, which is the case when an acceptance of epistemically unwarranted beliefs may increase instrumental profit for an individual. In my presentation, I will criticize this approach. The proponents of rational irrationality 1) misinterpret the primacy of theoretical reason over practical one; 2) underestimate social nature of normative systems; and 3) misunderstand the evolution of human cognition. I will illustrate my conclusions with an example from applied epistemology – unwarranted conspiracy theories. It is misleading and disparaging to use the term “rational” in relation to sets of unjustified beliefs that are based on conspiratorial intuitions. If philosophers cannot make their ideas on rationality clear, it might be better to leave out rationality completely from the epistemological discourse.