Technology in Society 71 (2022) 102117

Contents lists available at ScienceDirect

Technology in Society

journal homepage: www.elsevier.com/locate/techsoc

Finding the 'nudge' in hypernudge☆

Stuart Mills

Lecturer in Economics, Department of Economics, University of Leeds, United Kingdom

A B S T R A C T

'Hypernudge' describes a group of phenomena which occur at the intersection of behavioural science and computer science, and law. The term been increasingly used in the latter field, though sparingly in behavioural science. As such, 'hypernudge' remains largely absent from the behavioural science lexicon, inhibiting the field from participating within vital discussions surrounding the use of psychological insights with ubiquitous computing.

In this article, I search for the 'nudge' in hypernudge by critiquing the differences between the two concepts from a behavioural science perspective. I 'find the nudge' in hypernudge by conceptualising a hypernudge as a system of nudges which change over time and in response to feedback. In this sense, a hypernudge is not a *type* of nudge, but an *arrangement* of nudges. This article then engages in an extensive discussion of the implications on this concept for the hypernudging programme, and for nudging more broadly.

1. Introduction

The term "hypernudge" (p. 122) has been offered by legal scholar Karen Yeung [1] to describe the combination of behavioural science with computer science for algorithmic regulation. The term respects this combination, with the concept of, '*hyper*,' attributable to the vast expansion of computational resources and data within computer science, and the concept of, '*nudge*,' attributable to the programme of light-touch behavioural modification found in behavioural science [2,3].

Following its introduction, the term has seen a small but growing usage as scholars seek to understand how 'nudging' and similar concepts are being used throughout the digital economy [4–8]. Despite this, neither [1] nor those who have subsequently used the term have explicitly explored the relationship, *if any*, between the notion of a hypernudge, and that of a 'traditional' nudge.[1] This is unfortunate, for it undermines conversations between behavioural scientists and critics of hypernudging, as the former struggle to understand the subject which the latter critique. As behavioural science and nudging increasingly take on a technological dimension [9,10,11,89], owing substantially to the proliferation of surveillance and data collection technologies since the introduction of *nudging* [2]; also see Refs. [12,13], such conversations will likely prove significant in the coming years. This article contributes a conceptual account of hypernudging which clearly relates it to behavioural science.

In this article, I offer a novel definition of a hypernudge as a system of nudges which change over time and in response to feedback. In doing so, I draw a clear distinction between nudges and hypernudges. Rather than hypernudges being a *type* of nudge, I argue hypernudges result from arranging nudges in a technologically mediated fashion to continuously influence human behaviour, including through real-time (re)configuration and dynamic personalisation. The importance I place on hypernudges as systems or arrangements of nudges could be reached etymologically by focusing on the *hyper* of hypernudge (i.e., *hyper* indicating the connection of nudges, in the same way that hypertext is the connection of text). But my analysis relies more substantially on various literatures spanning behavioural and computational social science.

I do not argue in this article that 'hypernudge' is a replacement of or superior concept to digital nudging e.g., [2,88], which also informs my approach in this article, nor does this article intend to marginalise various literatures regarding nudging online behaviour, such as consumption behaviour (e.g., Ref. [4]; Mele et al., 2021). My efforts in examining the hypernudge concept are two-fold, and complementary to these literatures. *Firstly*, to interrogate a nudge concept developed outside of behavioural science for any useful insights or warnings it may provide the behavioural science literature. *Secondly*, to emphasise the notion of *arranging* nudges and designing *systems* or *choice environments*, rather than just specific choice architectures.

The structure of this article is as follows. I begin by contrasting given definitions of nudging and hypernudging, showing how the hypernudge literature has various recurrent themes, but limited overlap with the

☆ I am grateful to Henrik Skaug Sætra and Viktorija Morozovaite for their thoughtful comments and suggestions on earlier drafts of this manuscript. All errors are my own.

E-mail address: s.mills1@leeds.ac.uk.

[1] A notable exception may be [6]; who offers several affordances towards behavioural science and economics in their discussion of hypernudging.

https://doi.org/10.1016/j.techsoc.2022.102117
Received 23 April 2022; Received in revised form 29 August 2022; Accepted 3 September 2022
Available online 10 September 2022

behavioural science literature. I thus explore where overlap does exist, and where the hypernudging literature invites more novel considerations, through the recurrent themes of personalisation, real-time (re) configuration and prediction, and hiddenness. This leads me to my definition of a hypernudge. Following this, I highlight three 'burdens' which hypernudges may create. I do not offer these burdens as an exhaustive account, but rather as interesting challenges which appear important given the broad principles of nudge theory. These three burdens concern avoiding the nudge, understanding the nudge, and being experimented upon by nudgers (so-called *choice architects*). I then conclude.

1.1. Nudging, and hypernudging

How does a hypernudge differ from a nudge? Addressing this question is a central objective of this article. It is useful to begin by considering the definitions which have been given for both terms. A nudge is:

> "[A]ny aspect of choice architecture that alters people's behaviour in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a nudge, the intervention must be easy and cheap to avoid. Nudges are not mandates" [3]; p. 8 [2]; p. 20)

Various examples of nudges exist. For instance, people tend to choose whatever option is set as the default option [14], and so changing default options has been a popular and effective nudging strategy [15]. People also tend to be susceptible to social pressure and social norms [16], and highlighting the behaviours of others has been found to significantly change individual behaviours in areas such as energy usage [17]. These are but two of a plethora of nudges which have been developed in recent years (see Ref. [18]; for a recent, authoritative review).

While the above definition is not without criticism [19–21], it is a succinct and established definition of what nudges do: nudges do not change economic incentives, and what they do change – the environment in which a decision is made, also called *choice architecture* – must not constitute changes which cannot be easily avoided, or which restrict choice [2,3].

A clear definition or description of a hypernudge is, by comparison, less forthcoming, with most authors within this small literature choosing instead to describe hypernudging as a broad group of characteristics: Yeung (2017, p. 122) describes hypernudges as:

> "[N]imble, unobtrusive, and highly potent, providing the data subject with a highly personalised choice environment … Hypernudging relies on highlighting algorithmically determined correlations … dynamically configuring the user's informational choice context in ways intentionally designed to influence decisions."

Other scholars have offered similar definitions. Morozovaite (2021, p. 115) writes, "Hypernudging is built on the insights of linkages between the behavioural economics-grounded theory of the nudge and information systems (IS) literature," before elaborating that, "Hypernudging is one of the most sophisticated forms of digital nudging that allows for dynamically personalised user steering, where the aim is to reach the right user, with the right message, by the right means, at the right time, as many times as needed" (p. 117). [5] chooses to highlight the major features of hypernudging, rather than offer an explicit definition. These features are distillations of the ideas given by Ref. [1]; with [5] also highlighting the *dynamism* of hypernudges, their *predictive capacity*, and finally, the *hiddenness* of hypernudges.

Additional, less-sustained discussions of hypernudging can also be found. [7] considers hypernudging in their discussion of the possibilities of big data being used to produce highly accurate predictions of human behaviour. Yet, they are concerned with big data being used to make singular, highly accurate predictions, and avoid discussions of inter-temporality and dynamic feedback which are taken as constituent elements of *dynamism* [1,5,6]. In this sense, hypernudge is a useful concept to Ref. [7]; but it does not form a substantial component of their analysis. [4] draw on the concept of hypernudging in their discussion of technology and consumer behaviour, but as with [7], use the term as an umbrella term for digital nudging. Finally, Smith and Villiers-Botha [8] discuss the use of hypernudges in influencing children, arguing that children require protection from hypernudges as their preferences and cognitive processes are still developing. Yet, as with previous scholars, they once more use the term generally to describe ubiquitous computing which is levered to influence behaviour.

This brief analysis of the existing literature provides clues for further exploration of the hypernudge concept, with technology's capacity to predict behaviour, to follow decision makers, and to change and hide nudges, all consistent themes. Yet, these themes do not inherently advance the concept of hypernudge as a *behavioural* phenomenon and could quite easily be seen as discussions of the *medium* through which nudging occurs, rather than an *evolution* of the nudging mechanism in itself (see, for instance, Refs. [22–24]; Weinmann, Schneider and vom Brocke, 2016). For instance, Mele et al. (2021) develop the concept of 'smart nudging,' as choice architecture informed by various data streams to improve consumer experiences. This is like the notion of the 'choice engine,' initially proposed by Ref. [25]; and more recently championed by Ref. [23]; as well as [2] in their revised edition of *Nudge*. This is to say nothing of the recent boom in personality targeting and personalised persuasion research, which suggests nudging individual personality types using data gleamed through vectors such as social media may produce highly influential behavioural interventions (e.g., Refs. [10,26,27].

The hypernudging literature positions itself as a particular type of digital nudging. [6] describes hypernudging as, "one of the most sophisticated forms of digital nudging." Smith and Villiers-Botha [8] also suggest – as the nomenclature of the term implies – that hypernudges go *beyond* standard nudging, perhaps interfering with preference development. Finally, Yeung's (2017) use of the term "choice environment," rather than the typical term, "choice architecture," suggests an intention to think larger than individual nudges.

Nevertheless, this initial reading of the literature demonstrates the need to go beyond previous descriptions and interrogate the similarities and differences between nudges as they are understood, and hypernudges as they *seem to be*. In doing so, I will present a conceptualisation of hypernudging which clearly integrates nudging, and reveals *how* technology transforms the humble nudge into the techno-social entity which has prompted such discussion.[2] To accomplish this task, I will begin by interrogating the three features of hypernudging proposed by Ref. [5]; *dynamism, predictive capacity*, and *hiddenness*. Based on the arguments and evidence assembled, I have chosen to split *dynamism* into *personalisation* and *real-time (*re*)configuration* and have chosen to discuss the latter alongside *predictive capacity*.

1.2. Personalisation

While personalisation is a feature commonly associated with big data and technology [1,22,27–29], the impetus for *personalised nudging* in nudge theory comes from what Sunstein (2012, p. 6) calls, "the problem of heterogeneity" (also see Ref. [30]). The problem of heterogeneity occurs when different (i.e., *heterogeneous*) individuals are nudged in the same way. Assuming the nudge has been used because it is expected to confer some welfare benefit (however this is defined [31,32]) across the population being nudged (i.e., a *net* benefit), because individuals within the population are different, it would be expected that some individuals

[2] The term 'techno-social' is a homage to the work of [12]; who consider nudging a substantial component to what they identify as the contemporary 'techno-engineering' of society [12]. do not *explicitly* discuss hypernudging, though the substance of the idea seems highly complementary.

greatly benefited from the nudge, while others greatly suffered because of it [30,86]. Thus, even when the net welfare benefit of an impersonal nudge is expected to be positive, a theoretical argument could be made that the net benefit of nudging could be increased if nudges could respect heterogeneity within the population [27]; Page, Castleman and Meyer, 2020; [29,30,33].

[1] also follows the problem of heterogeneity in their discussion of hypernudge. Considering the example of a speed hump as a nudge. Yeung (2017, p. 122) writes:

> "Although vehicles proceed slowly in residential areas to ensure public safety, speed humps invariably slow down emergency vehicles responding to call-outs. In contrast, Big Data-driven nudges avoid the over-and under-inclusiveness of static forms of design-based regulation."

Here, the speed hump functions as an impersonal nudge as it encourages vehicles to slow down but does not force or strongly coerce the driver via, say, a large fine [3]. However, not all vehicles are the same, and some – such as emergency service vehicles – would benefit from not being nudged to slow down. If it were possible to design a speed hump that did not impact emergency service vehicles – in other words, *personalise the nudge* – a theoretical argument could be made that net welfare has been increased: the gain of welfare of public safety afforded by the hump remains, while the loss of welfare of emergency vehicle response incidentally produced by the hump disappears.[3]

The role of personalisation in nudge theory, *just as it is in hypernudging*, comes not from *technology*, but from a theoretical argument regarding optimisation. Technologies such as big data and machine learning *may* be used to design personalised nudges [27,28], just as they are essential technologies in several of the hypernudge examples given by, say, [6]. Yet, as [28] argue, and as several studies show [34,81,83], not all *personalised nudges* are necessarily *big data* nudges, or indeed, *hypernudges*.

For instance, Page, Castleman and Meyer (2020) offer an example of what Porat and Strahilevitz (2014, p. 17) label "crude" personalisation – personalisation that is neither computationally- nor data-intensive. They personalise reminder SMS texts used to encourage high school students to complete their FAFSA application (a government programme designed to support students into higher education). These messages are personalised based on the progression of the individual students' application: a student who has not begun an application is sent a message to begin; a student who has completed part of the application is sent a message to finish; and a student who has finished is sent a message to make sure they have all the necessary additional materials prepared. The technologies utilised in this intervention are an automated SMS system which school districts already have, and a short computer script which reads the status of *in progress* and *completed* applications.

Reflecting on personalisation alone, one arrives at an interesting question: is the above study an example of *hypernudging*? In the absence of a clear definition (at present), one may conclude *maybe*. But when contrasted with the examples of hypernudging given in the literature, such as Google's targeted advertising or various body-tracking technologies [5,6], or when contrasted with the "highly sophisticated" algorithms which Yeung (2017, p. 121) describes, one comes to a firm answer: *no*. The purpose of this exercise is to recognise that, from the perspective of nudge theory, personalisation alone is insufficient to warrant a *hypernudge* concept which is distinct from *nudge* – indeed , the central discussion of personalisation found in [1] reconciles with the problem of heterogeneity found in nudge theory.[4]

1.3. Real-time (Re)Configuration and predictive capacity

Dynamism can be understood as change and adaption [6]. In this sense, personalisation is a description of *what* a nudge is changed into. Yet, as above, this *what* does not appear to be a distinguishing feature between nudging and hypernudging. Instead, I will turn attention to *how* a nudge is changed. Within hypernudging, this process concerns *real-time (*re*)configuration* and *predictive capacity*.

As well as describing personalised nudges as free from, "over- and under-inclusiveness," [1] also asserts they differ from, "static forms of design-based regulation," by which it is meant nudging as traditionally understood. This may be a substantial distinction between nudges and the hypernudge concept, yet as above, I have argued nudges can also be personalised, and in this sense can be changed. The substantial distinction, therefore, is not simply a matter of *change*, but as acknowledged within the hypernudge literature [1,5], *real-time* change in response to feedback.

Consider the example of Google Maps GPS discussed as a hypernudge by Ref. [1]. Google Maps can be used for directions when driving or navigating on-foot. As the user changes their location, and as other changes occur, such as traffic build-up, Google Maps can change the instructions it gives the user. Here, the method of nudging does not change, but the instructions being given do change, and in this sense, so too does the nudge [27]. *In principle*, this example seems rather analogous to the above example of personalised SMS texts sent to students: in both instances, the method of nudging does not change, and in both instances, the outcome individuals are nudged towards changes *in response to* individual behaviour. Yet, a key distinction between these two personalised nudges is the speed at which the nudges change. Google Maps GPS changes in real-time, based on an individual's GPS location, and the nudge in the form of the instruction changes in real-time also. This is also preferable for someone navigating in real-time. By contrast, the SMS texts are sent periodically (e.g., weekly), and only change at the pre-specified period, assuming the student has made progress in their application. This is preferable within the context of an application with fixed requirements and extended time horizons (e.g., several months or years).

The rapidity of change, therefore, represents an important distinction between nudges and hypernudges. So too does the way in which hypernudges *can* change. This relates to predictive capacity. For instance, the SMS texts change only in response to student progress on their applications, but a sophisticated GPS system such as Google Maps can change the instructions it gives in response to a variety of factors *in addition to* the individual's behaviour. For instance, by being linked to the internet, the system can detect heavy traffic build-up and direct an individual to avoid it through an alternative route.[5] Another example given by Ref. [1], Facebook's News feed algorithm, is a more extreme example of this, with the items shown to individual users often being curated through a complex, algorithmic calculation factoring in the individual's behaviours, but also the behaviours of *others* [35].

The purpose of this variety of data streams is to bolster the predictive capacity of the intervention. Predictions regarding nudges, hyper or otherwise, necessarily require testing and feedback. When nudges fail

[3] It is debatable whether a speed hump is an appropriate example here, insofar as it may not be considered a nudge. For instance, the costs of ignoring a speed hump *may* be significant if it results in a collision, or even if it results in mere discomfort or slight vehicle damage. Furthermore, alternatives exist. For instance, would not a sign asking drivers to slow down be more conducive to nudging? Perhaps, but the purpose of this footnote is not to debate the behavioural legitimacy of speed humps and traffic signs. Rather, it is to demonstrate the fraught nature of the definition of *nudging*, which may create difficulties both for this discussion and within the wider literature.

[4] This is despite Yeung (2017, p. 121), perhaps implicitly, assuming as others have [28,50] that personalisation *necessarily requires* sophisticated technologies when she attributes the avoidance of "over- and under-inclusiveness" to "Big-Data driven nudges" exclusively.

[5] GPS is often cited as an example of a nudge (e.g. Ref. [78], and therefore this distinction is important.

[36], choice architects often reflect on why, and re-evaluate their approach [37]. Many interventions are also investigated in laboratory settings, or in other specialised environments, and as such the behavioural findings of these studies should only ever be seen as predictions when applied in real-world environments [38]. These are the systems of adaption which nudges and hypernudges share.

Yet, through consideration of ideas already given, a distinct character associated with hypernudging can be seen. When one considers the problem of heterogeneity, the means by which a nudge is considered a 'success' or a 'failure' becomes more nuanced. For instance, Beshears et al. (2021) find an impersonal nudge which exploits time discounting to encourage saving fails for a minority of individuals who misconstrue the time discounting message as a license for increased spending today. Furthermore, feedback may only occur at the point a nudge is used, and a behaviour is recorded. As such, less frequent interventions will tend to receive less feedback than very frequent interventions.

This is all to say, the role of predictive capacity is intimately tied to *dynamism*: a nudge which changes in real-time receives more feedback, while a nudge which is changing to be personalised *needs* more feedback, owing to the multitude of differences at any given moment, within any given context, and regarding any given individual. As above, as personalisation is about *optimisation*, to optimise predictive accuracy, an intervention would be expected to maximise feedback through the rapidity of (re)configuration and by drawing on a multitude of data sources.

Meeting these expectations introduces the importance of technology and invites one to place emphasis on the *capacity* aspect of *predictive capacity*. As above, traditional nudging is based on prediction, and can benefit from feedback. But real-time (re)configuration necessitates a choice environment where such (re)configuration is possible (Weinmann, Schneiders and vom Brocke, 2016; [1], while a multitude of data streams necessitates a broad, behavioural informatics infrastructure [39] and computational resources to produce useful and actionable insights [35,40]. These technological components enable a transformation in how nudges can develop.

1.4. Hiddenness

Hiddenness as a feature constitutes the least substantial part of this initial inquiry, and for this reason I will address it briefly. Nevertheless, this is a feature identified by Ref. [5]; and so I will address it for completeness. Both [1,5] discuss hiddenness. [1] echoes Bovens' (2008, p. 3) argument that "nudges work best in the dark," intermating that nudges which are hidden are optimal. The implication of this, in terms of hypernudging, is that hypernudges – through real-time (re)configuration – can blend seamlessly into the background, and facilitate what Frischmann and Selinger (2016, p. 372) dub a "frictionless world." For one concerned with individual liberty, as [1] is, such a prospect may be concerning.

Yet, Bovens' (2008) original assertion has faced challenge from various empirical results which find so-called transparent nudges still work [41–45]. While no empirical study has, to my knowledge, investigated the effectiveness of transparent *hypernudges*, questions concerning hiddenness and transparency transcend hypernudges, and have and continue to be raised about traditional nudges also.[6]

[5] has also considered the hiddenness of hypernudges in terms of the ability for individuals to understand what the hypernudge is doing, or how their actions influence the hypernudge; what one might call *ease of understanding*. While it seems reasonable to assume most people understand that companies such as Google will operate in such a way as to maximise their economic returns – perhaps even to the detriment of users – many people will struggle to understand how, 'liking,' a particular post, or searching a particular term, impacts them [40,46,47]. This is often not the case for traditional nudges, owing to their static character [27].

Finally, from a philosophical perspective, hiddenness may be understood as one's conscious interactions with technology. As [48] discusses, while one may be aware of technology, insofar as that technology is used, people stop consciously seeing it as *technology*, but some means to an end, with the end occupying their attention (also see Ref. [85]). Thus, technology can be hidden in a sense which invokes once more the notion of frictionlessness [13]. Yet, insofar as choice architecture is unavoidable [3,50], individuals can be understood as constantly navigating choice environments where their consciousness of the factors influencing them vary. Insofar as my purpose here is to identify *distinction* between nudges and hypernudges, such a discussion seems unhelpful.

2. Three burdens

The discussion above leads to what I propose as a worthwhile definition of a hypernudge, with respect to nudge theory:

Hypernudges are systems of nudges which change over time and in response to feedback.[7]

A summary of the above discussion can also be found in Table 1:

In this section, I will explore the implications of these differences outlined in Table 1, and unpack the definition given above, through a discussion of three behavioural features which seem to emerge from hypernudges as conceptualised. These features I characterise as *burdens*, and they are the *burden of avoidance,* the *burden of understanding,* and the *burden of experimentation*. Through this discussion, a fuller understanding of the hypernudge concept as it relates to nudge theory is achieved. In discussing burdens, I will make illusions to *welfare*, a concept I will purposely keep rather broad, as others (e.g. Refs. [31,32]) in the behavioural literature previously have. *Generally*, the term 'welfare' will be used to capture any harm or discomfort which comes from interfering with a person's will, by which I mean, what a person would have done had attempts to influence them not been made. This is potentially contentious given traditional nudges can be accused of doing this also (see Ref. [51]). Therefore, insofar as I describe the burdens of hypernudges, I will try to focus on *new* welfare concerns, and where not new, note as such.

2.1. The burden of avoidance

The burden of avoidance comes, primarily, from the notion of hypernudges *changing in real-time*. Both [1,5] allude to this burden, but

[6] It is possible that the emphasis [1,5] place on hiddenness arises because it is not possible to have a transparent hypernudge, owing to the technology involved, while it is possible to have a transparent nudge. In this sense, hiddenness is an important distinguishing feature. That being so, however, it seems strange that [1] chooses to discuss [79] at all, and that [5] centres their discussion on the transparency of the *motives* of those who are nudging. Both approaches strike one more as criticisms of the transparency of nudging *generally*, rather than remarks on the hiddenness of hypernudges *specifically* [13]. may offer a better account in this regard.

[7] This definition may leave an aspect of hypernudging to be misunderstood. Aside from dynamism, an alternative definition of a hypernudge may be *nudges which are connected*. This is something of a literal interpretation of the term hypernudge; just as *hyper*text describes the connections between bodies of text, and *hyper*space describes the space between celestial systems, so too might one see hyper as a description of the connection between nudges. Equally, this is also an interpretation reliant on a contemporary use of the word hyper. The original use of the word *hyper* was to describe something which was *over*, *above*, or *beyond* some normal state of being. In this sense, a hypernudge could be interpreted as being an *extensive* nudge, or something *beyond* a nudge. While I will not make explicit reference to such etymological interpretations, parallels can be seen.

Table 1
Nudge vs. Hypernudge Crib Sheet.

Feature	Traditional Nudge	Hypernudge
Personalisation	Often are not personalised and suffer from the 'problem of heterogeneity.' However, these nudges can be personalised using low intensity 'crude' methods which may reduce problems caused by heterogeneity in the population.	Always consist of personalised nudges personalised using high intensity 'sophisticated' methods such as machine learning and big data. Through personalisation, hypernudges are expected to significantly reduce the problem of heterogeneity.
Real-time (Re) Configuration	Often only change periodically, either during a pre-defined periodicity (e.g., an annual review) or a periodicity implied from the context in which the nudge operates (e.g., a school year).	Always change as quickly as possible, ideally in real-time, to reflect as much feedback, collected in the form of data, as possible.
Predictive Capacity	Often the predictive capacity may vary and will be highly influenced by the environment in which preliminary trials took place. Opportunities for feedback to evaluate predictions are determined by the rapidity of the nudge, which often spans several months or years.	Always constructed to maximise predictive capacity through an optimisation perspective compatible with various technologies such as loss functions in machine learning. Opportunities to evaluate predictions are common owing to the rapidity of the hypernudge.
Hiddenness	The role of hiddenness in improving nudges remains debatable, with empirical evidence suggesting that nudges remain effective even when transparent. Owing to the ubiquity of choice architecture, transparency is never guaranteed.	Hypernudges may be hidden owing to the technology which enables them to fade into the background. Furthermore, technology itself can easily become 'hidden' in a philosophical sense as attention turns from the technology to the ends which the technology facilitates.

neither follow it to the natural conclusion, especially as a contrast with simple nudging.

To appreciate the burden of avoidance, it is helpful to consider two examples. *Firstly*, consider the previously discussed example of a speed hump. The speed hump nudges insofar as it encourages the driver to slow down, but does not force or otherwise significantly coerce the driver. The driver could choose to drive over the hump at speed, only suffering the temporary physical discomfort of doing so. In driving over the hump at speed, the driver is essentially 'opting out' of the nudge; they are choosing the option *not* nudged towards. As they have driven down that particular road containing a speed hump, the driver must decide to follow the nudge or not. But once this discussion has been made – say, to *not* slow down – the driver is free from the hump. The hump does not chase the driver down the road, nor does the road itself contort to produce additional humps in *response* to the driver's decision.[8]

Secondly, consider the example of Google Maps once more. A GPS is typically considered a nudge insofar as it tells a person where to go but leaves the driving up to the driver [52,53]. A GPS is also a hypernudge, following [1]. For Yeung (2017, p. 122), a service like Google Maps is a hypernudge because it will automatically, and in real-time, "dynamically [reconfigure] the user's informational choice context." Furthermore, "the driver using Google Maps [is not] compelled to follow the 'suggestions' it offers. But if the driver fails to follow a suggested direction, Google Maps simply reconfigures its guidance relative to the vehicle's new location." In other words, and in contrast to the speed hump, a driver who 'opts-out' of turning left and instead goes right will be immediately prompted by the GPS with a nudge to 'do a U-turn,' or some other 'course-correcting' manoeuvre. Real-time feedback is integrated, and the hypernudging GPS immediately responds to a person *not following the nudge* by nudging them again, possibly differently, but certainly *continuously*, to achieve the *choice architect's objective*. As Morozovaite (2021, p. 117) notes, this process could occur, "as many times as needed." Hypernudges *follow*.

Of course, the decision-maker could opt-out of the hypernudge entirely – for instance, they could switch Google Maps off. But, as Lanzing (2019, p. 555) notes, this kind of extreme withdrawal is the *only means* of opting out of a hypernudge: "hypernudges cannot be opted out from without quitting the service altogether." This is to say nothing of the *cost* of opting-out itself. Hypernudges such as Facebook, YouTube or TikTok have compelling *social* elements to them. Yet, a person cannot have the social elements without also subjecting themselves to the hypernudging systems these services push. As such, the 'trivial' decision to opt-out of a service such as Google Maps is not consistently trivial across all services [54,55].

A visualisation of a hypernudging system can be seen in Fig. 1:

As seen, the hypernudging system constitutes a series of nudges connected based on the choices of the decision-maker (hence, a hypernudge is a *system*, rather than a *type* of nudge). Personalisation follows from the choices of the decision-maker, but as above, should also be understood to integrate historic data and data from *other* decision-makers. In contrast with the traditional nudge model shown on the left, the hypernudging system follows the decision-maker until one of two outcomes is reached – the decision-maker accepts the outcome they are nudged towards, or they abandon the system entirely.

Each nudge within a hypernudging system which a decision-maker wishes to avoid imposes some cost onto the decision-maker, as do nudges more generally [21]. Yet, the tendency for hypernudges to follow would be expected to lead to greater overall costs of avoidance, compared to traditional nudges, arising from the greater number of nudges and continuity of nudging within the hypernudge system. Indeed, the option to leave the hypernudging system only emerges after n iterations of nudging, presumably when the decision-maker is so exasperated at opting-out. In other words, the default position is to *always nudge* [12]. Two emerging ideas are complimentary to this perspective. *Firstly*, behavioural *sludge* has been proposed by [53,56] as a kind of friction which makes it harder for individuals to accomplish their objectives. *Secondly*, the *dark patterns* literature has identified *nagging* as a common technique found in online spaces to encourage individuals to behave in ways they otherwise would not [9,57]. Both appear as worthwhile concepts within this discussion.

In sum, the burden of avoidance can be understood as the cost of opting-out of a hypernudge, given that hypernudges *follow* decision-makers. The example given above has been Google Maps GPS, but other examples point to the burden of avoidance. The Facebook News feed algorithm automatically curates content for users, as does the TikTok algorithm; the Google search algorithm automatically filters and orders results, as does the YouTube algorithm. Any deviation from these hypernudges is not met with force (e.g., "you *must* 'Like' this page") but rather immediate reconfiguration (e.g., "you didn't like that page? Well, here's another I'm sure you'll like instead").

[8] Another example, which I have often employed for comic effect, is that of opt-out organ donation. A person may be nudged, say when passing their driving test, to become an organ donor, with the default option set to 'Yes, I want to be an organ donor.' This is a typical default option nudge which has been shown to greatly increase the number of people being registered as organ donors [80]. But consider those who choose *not* to become donors, which is to say, to not follow the nudge. As with the speed hump example, declining to become an organ donor does not result in persistent prompts to reconsider until the 'correct' (i.e., nudged) outcome is selected. Nor is there immediate 'personalisation': there is no department of transportation worker chasing you down the street yelling, "how about *just* your kidneys?' when you opt out.

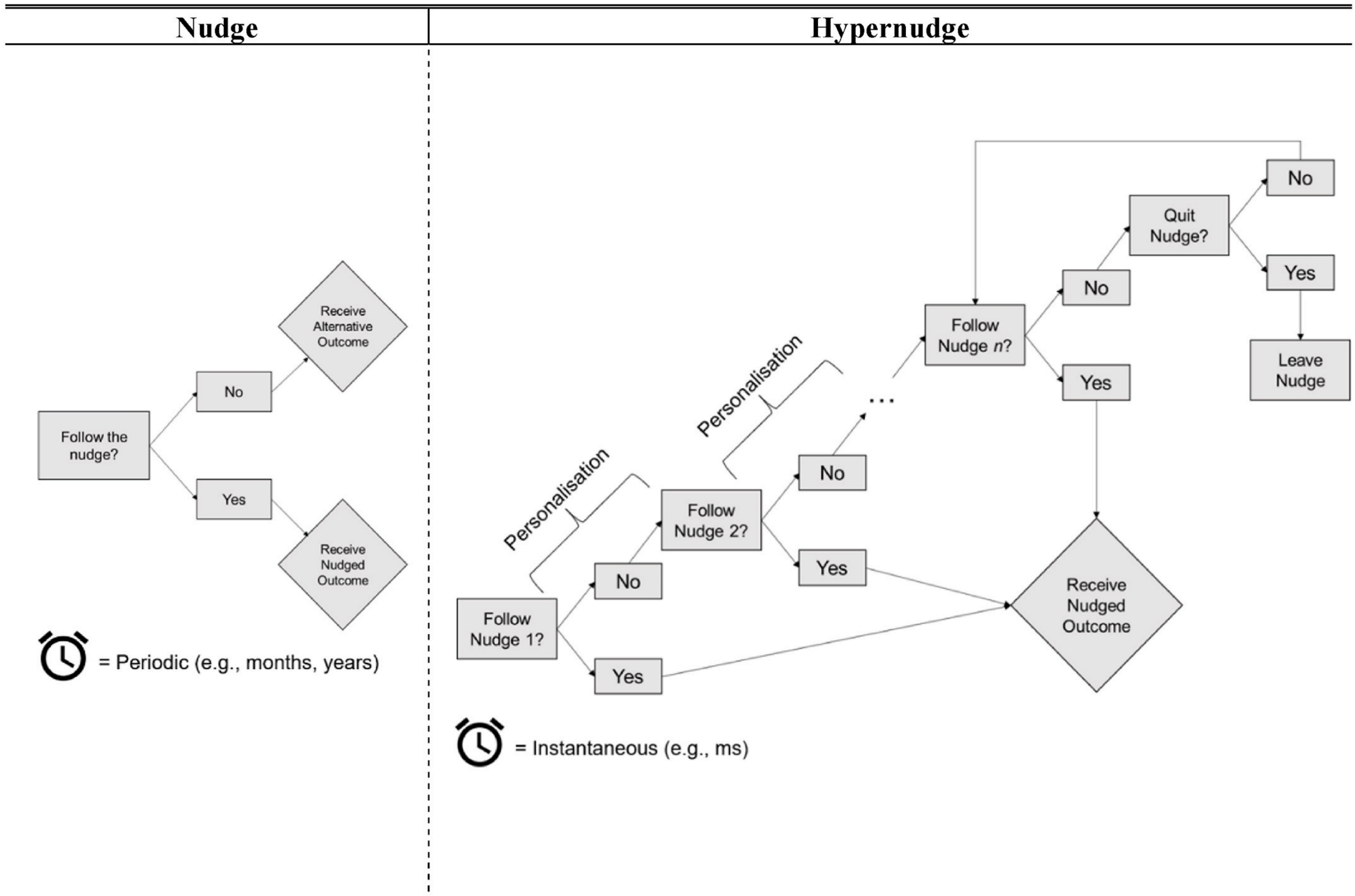

Fig. 1. Hypernudges follow.

2.2. The burden of understanding

The burden of understanding concerns an individual's ability to understand the hypernudge itself, and thus links to above discussions concerning data streams for prediction, and hiddenness. Understandability within algorithmic environments has been of interest to behavioural scientists [58,59] in recent years, and it is through considering some of these efforts that the challenges of the burden of understanding are best revealed.

[59] suggest that user understanding of algorithmic hypernudges such as the Facebook News feed algorithmic can be promoted by allowing users to decide what information should be used to curate content for them. For instance, users may want the relative recency of a post to contribute 25% to a content post's, 'score,' – with the highest scoring post appearing most prominently (e.g., first, or top) – while the number of, 'Likes,' a post receives should contribute 50%, and so on. This is but one recommendation; others include information disclosure revealing how a post's score is calculated, or engagement data revealing the source of a post, allowing the user to appraise its quality.

These measures may promote user autonomy, at least insofar as they grant users some control over the hypernudging algorithm to which they expose themselves, and thus promote understanding. However, two further considerations must be made. *Firstly*, one could consider the material interests of those who create and control hypernudging systems – e.g., large technology companies. The practicality of an intervention such as that suggested by [59] must be considered in conjunction with the interests of those who may grant *users autonomy over the algorithm*. This criticism, or a similar *digital economy* critique, has been made by many authors previously [40,60,87], including by those within the hypernudging literature [1,5]. Therefore, I will focus my attention on the second consideration.

Secondly, proposals such as that given above represent an intervention *in response to* hypernudging, rather than an explicit critique of hypernudging as a phenomenon. Much of the discussion thus far has been theoretical in nature, and so it is helpful now to consider a practical example. The Facebook News feed algorithm curates, on any given day for any typical user, around 300 posts for that user, out of a pool of around 1500 posts [1,61]. Furthermore, a 2014 investigation found that 75% of the variance in Facebook's algorithm could be explained by only five variables, including total 'Likes' and whether the post was made on a weekday or a weekend [62].[9] However, in 2013, Lars Backstrom – the engineering manager for the News feed algorithm – claimed up to 100, 000 variables were used.

In 2014, the true number of variables which would be relevant to consider probably lay between these upper and lower bounds, though as of writing, these estimates are likely rather inaccurate. For instance, in 2017 Facebook introduced 'reactions,' which allowed users to express a range of reactions to a post, an evolution of the relatively simple 'Like,'

[9] This result is based on an OLS regression performed by a journalist for a popular magazine, *The Federalist*. It likely does not withstand serious statistical scrutiny. I have chosen to use this statistic much more illustratively than definitively, as a reader will see. However, in an investigation similar to this 2014 study, conducted by a different popular magazine, the *Wall Street Journal* (2021), variance in the TikTok algorithm could be explained, to a similar degree, using only one variable – watch time. Even accepting statistical criticisms, it is likely that many hypernudging algorithms are constituted using relatively few, *core*, behavioural metrics.

button. In 2018, it is reported that Facebook started prioritising 'meaningful reaction,' such as commenting and sharing over simple reactions such as liking, and as of 2020, Facebook has begun implementing credibility assessments when recommending content [63].

The constantly changing nature of hypernudging systems, as well as the sheer size and complexity of hypernudging systems, produces a burden of understanding even before attempts are made to make these systems more accessible to users [1,5]. For instance, taking the concept suggested [59] once more, a relatively simple model for determining a post's score could be proposed:

$$Score = Ax + By + Cz\# \tag{1}$$

Here, x, y and z represent metrics established by the hypernudging choice architect (e.g., the Facebook platform), such as the number of 'Likes,' the recency of the post, and the similarity of the post to content the decision-maker has previously engaged with. Furthermore, A, B and C represent percentage weights, following [59]; such that $A + B + C = 100\%$. Decision-makers are assumed to be able to determine the value of these weights (which may be an unfair assumption [46]).

While this model appears quite simple, and the role of the weights quite understandable, this does not necessarily ensure that the *hypernudge* is understandable. For instance, how is a metric such as '*the similarity of the post to content the decision-maker has previously engaged with*,' determined? Alternatively, how can the recency of a post be determined; recent, relative to *what*? Finally, if the number of 'Likes,' is influenced by how others interact with the hypernudge, these third parties can influence how the decision-maker experiences the hypernudging algorithm without either they, or the decision-maker, knowingly *exactly* how this influence is manifested [35,64]. This is all assuming that the score really is the sum of only *three* weighted variables. As the history of the Facebook algorithm reveals, the number of variables is likely to be much larger. As [65] has argued, this raises important questions regarding *control*. For instance, if 100 variables are used to generate a score, and only five can be adjusted by the decision-maker, why only five, and why *those five in-particular*? Even when trying to promote autonomy in hypernudged-environments, it is likely a large degree of control need be retained by the choice architect simply because of the underlying complexity of hypernudging algorithms [66].

All these factors inhibit understanding and produce a burden of understanding which raises the costs of autonomous decision-making; in such an environment, it is easier to just follow the nudge [12,67,68]. The underlying complexity of hypernudges means that decision-makers may struggle to act autonomously, and as with the burden of avoidance, simply default into following the hypernudge.

2.3. The burden of experimentation

The final burden is what I call the burden of experimentation. Experimentation is an implicit but necessary component of hypernudging, and the artifacts of this component have been seen already. For instance, recall Equation (1). There is no obvious reason to assume that the product terms should be combined into a single score. Nor is there any reason to believe that the combination of these products should be additive; why not minus the first product from the second, or multiply the second by the third? In addition to this, there is no reason to believe three products are adequate. This has been acknowledged – there could be 100 variables which are used.

A somewhat more realistic mathematical expression of a hypernudge, which resolves some (not all) of these questions, could be found by adopting a machine learning perspective, as in Equation (2):

$$p(Outcome) = f(a,\ b,\ c,\ \ldots N)\# \tag{2}$$

Here, the hypernudge is not choosing, say, which road to direct a driver along, or which post to show a Facebook user, but instead which arrangement of choice architecture is *probabilistically* most likely to achieve a pre-determined outcome [69]. This perspective integrates both Lanzing's [5] argument that hypernudges are based in prediction, as well as the reality that hypernudges do not care what route a person takes, or what post a person sees, but simply that a person reaches their destination as quickly as possible, or engages with a post, or whatever else it has been programmed to optimise [70]. Equation (2) uses function notion, acknowledging that input variables may be arranged in a variety of ways, and contains N variables, recognising that the number of input variables may be arbitrarily large.

The burden of experimentation is not wholly concerned with the technical underpinnings of hypernudges which the paragraphs immediately above have begun to entertain. Rather, the burden of experimentation concerns the costs decision-makers must bear as these technical details are worked out. For instance, one may not know *a priori* what data should be collected to personalise a nudge [29]. Therefore, it may be wise to collect a tremendous amount of data simply so that experiments can be conducted to *acquire* the knowledge of how to personalise. This uncertainty also creates a kind of logic of data accumulation insofar as because data *may be useful or valuable in the future*, if it can be collected today, it should be [40,71]. Yet, this also creates costs for decision-makers in the form of privacy and surveillance [1,5,29,65]. These costs, in short, represent the burden of experimentation.

This 'experimental surveillance' may be understood as what Sætra (2020, p. 3) dubs "proactive surveillance" insofar as, "it involves [the] use of surveillance to uncover information and change the actions of individuals." An example of such proactive surveillance within a hypernudging space comes from [12], who report that Google Maps GPS will purposely send drivers on sub-optimal routes to gather data on under-mapped roads. Another famous, or perhaps *infamous*, example is Facebook's mood experiment, where the News feed algorithm was changed in such a way as to alter the mood of users, under the retroactive justification of learning how to improve Facebook services [72,73]. A final, recent example is that of the social media platform TikTok, which will periodically show a user content they may *not* like, to learn more about the user [74].

These examples underscore a philosophy of algorithm design, succinctly discussed by Russell (2019, p. 8):

> "Typically, such algorithms are designed to maximize *click-through*, that is, the probability that the user clicks on presented items. The solution is simply to present items the user likes to click on, right? Wrong. The solution is to change user's preferences so that they become more predictable."

In some instances, such experimentation may ultimately benefit the individual through, say, personalised recommendations which enhance wellbeing [23,75]. But hypernudging systems which morph or constrain individual preferences, even for a *predicted benefit to the individual*, raise serious concerns about individual self-expression and actualisation [76]. Such a concern motivates the work of Smith and Villiers-Botha (2021) on their opposition to hypernudges being used on children, but similar concerns about algorithmic experimentation have been expressed more generally. For instance, [77] argues these algorithmic mechanisms constrain cultural expression and harm social development. Beyond individual concerns about privacy, surveillance, and self-expression, these wider social costs should be understood as elements of the burden of experimentation. These three burdens are summarised in Table 2.

3. Conclusion

This article contributes to the emerging hypernudge literature by examining previous uses of the term through ideas and results found in nudge theory and behavioural science. I argue that hypernudges are systems of nudges which change over time and in response to feedback. This definition places the hypernudge concept in direct relation to the more established nudge concept. This definition also develops hypernudging in relation to the digital nudging literature; while the former is

Table 2
Three burdens.

Burden	Detail
Burden of Avoidance	• The challenge for decision-makers to 'go their own way' created by hypernudges. • As hypernudges change in real-time, and in response to immediate feedback, hypernudging systems immediately nudge a decision-maker again.
Burden of Understanding	• The challenge for decision-makers to understand how and why they are being nudged, and to therefore make an informed decision. • As hypernudges are often proprietary, and as they often use significant amounts of data in abstract ways, a decision-maker may struggle to understand how they are being nudged, and simply defer to the nudge.
Burden of Experimentation	• The challenge for decision-makers to have their preferences respected. • As hypernudges need to learn how to nudge, decision-makers may often be subject to experimentation which will not respect their preferences.

a part of the latter, the emphasis hypernudging places on the arrangement of nudges, and choice environments, distinguishes hypernudging – conceptually – from digital nudges, and in turn, enhances both.

With this definition, I offer some critique of the social implications of hypernudges which emerge beyond those which might also be attributed to traditional nudges. My critique consists of three 'burdens,' which – while not offered as exhaustive – emphasise important aspects found in both the hypernudge literature, and the nudge theory literature generally. These three burdens concern *avoidance*, *understanding*, and *experimentation*: hypernudges are harder to avoid than individual, traditional nudges; they are also likely harder to understand, owing to complexities arising via functions such as personalisation; finally, to develop this functionality, hypernudges must experiment on individuals in potentially objectionable ways.

Author statement

I declare no conflict of interest arising from the material contained within this manuscript. I take full credit and responsibility for the material contained and arranged within this manuscript. All errors are my own.

Data availability

No data was used for the research described in the article.

References

[1] K. Yeung, Hypernudge': big Data as a mode of regulation by design, Inf. Commun. Soc. 20 (1) (2017) 118–136.
[2] R.H. Thaler, C.R. Sunstein, Nudge: the Final Edition, Penguin Books, UK, 2021.
[3] R.H. Thaler, C.R. Sunstein, Nudge: Improving Decisions about Health, Wealth, and Happiness, Penguin Books, UK, 2008.
[4] A. Darmody, D. Zwick, Manipulate to empower: hyper-relevance and the contradictions of marketing in the age of surveillance capitalism, Big Data and Society 7 (1) (2020) 1–12.
[5] M. Lanzing, Strongly recommended" revisiting decisional privacy to judge hypernudging in self-tracking technologies, Philos. Tech 32 (2019) 549–568.
[6] V. Morozovaite, Two sides of the digital advertising coin: putting hypernudging into perspective, Market Compet. Law Rev. 5 (2) (2021) 105–145.
[7] H.S. Sætra, When nudge comes to shove: liberty and nudging in the era of big data, Technol. Soc. 59 (2019), 101130.
[8] J. Smith, T. de Villiers-Botha, Hey, Google, leave those kids alone: against hypernudging children in the age of big data, AI Soc. (2021), https://doi.org/10.1007/s00146-021-01314-w.
[9] A. Mathur, G. Acar, M.J. Friedman, E. Lucherini, J. Mayer, M. Chetty, A. Narayanan, Dark patterns at scale: findings from a crawl of 11K shopping websites, in: Proceedings of ACM Human-Computer Interactions, vol. 3, 2019, https://doi.org/10.1145/3359183.
[10] S.C. Matz, M. Kosinski, G. Nave, D.J. Stillwell, Psychological targeting as an effective approach to digital mass persuasion, Proc. Natl. Acad. Sci. USA 114 (48) (2017) 12714–12719.
[11] M. Weinmann, C. Schneider, J. vom Brocke, Digital Nudging [Online], SSRN, 2016, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2708250, 30/01/2020.
[12] B. Frischmann, E. Selinger, Re-engineering Humanity, Cambridge University Press, UK, 2018.
[13] B. Frischmann, E. Selinger, Utopia?: a technologically determined world of frictionless transactions, optimized production, and maximal happiness, UCLA Law Rev. 64 (2016) 372–391.
[14] W. Samuelson, R. Zeckhauser, Status quo bias in decision making, J. Risk Uncertain. 1 (1) (1988) 7–59.
[15] J.M. Jachimowicz, S. Duncan, E.U. Weber, E.J. Johnson, When and why defaults influence decisions: a meta-analysis of default effects, Behaviour. Public Pol. 3 (2) (2019) 159–186.
[16] R.B. Cialdini, M.R. Trost, Social influence: social norms, conformity, and compliance, in: D.T. Gilbert, S.T. Fiske, G. Lindzey (Eds.), The Handbook of Social Psychology, McGraw-Hill, USA, 1998.
[17] H. Allcott, T. Rogers, The short-run and long-run effects of behavioral interventions: experimental evidence from energy conservation, Am. Econ. Rev. 104 (10) (2014) 3003–3007.
[18] J. Beshears, H. Kosowsky, Nudging: progress to date and future directions, Organ. Behav. Hum. Decis. Process. 161 (2020) 3–19.
[19] P.G. Hansen, The definition of nudge and libertarian paternalism: does the hand fit the glove? European J. Risk Regulat. 7 (1) (2016) 155–174.
[20] D. Hausman, B. Welch, Debate: to nudge or not to nudge, J. Polit. Philos. 18 (1) (2010) 123–136.
[21] S. Mills, Nudge/sludge symmetry: on the relationship between nudge and sludge and the resulting ontological, normative and transparency implications, Behaviour. Public Pol. (2020), https://doi.org/10.1017/bpp.2020.61.
[22] S. Benartzi, The Smarter Screen: Surprising Ways To Influence And Improve Online Behavior, Portfolio, UK, 2017.
[23] E.J. Johnson, How Netflix's Choice Engine Drives its Business [Online], Behavioral Scientist, 2021, https://behavioralscientist.org/how-the-netflix-choice-engine-tries-to-maximize-happiness-per-dollar-spent_ux_ui/, 15/03/2022.
[24] M. Lavi, Evil nudges, Vanderbilt J Entertain. Tech. Law 1 (1) (2018) 1–95.
[25] R.H. Thaler, W. Tucker, Smarter information, smarter consumers, Harv. Bus. Rev. 91 (1–2) (2013) 44–54.
[26] M. Kaptein, Persuasion Profiling: How The Internet Knows what Makes You Tick, Business Contact, UK, 2015.
[27] S. Mills, Personalized nudging, Behaviour. Public Pol. 6 (1) (2022) 150–159.
[28] A. Porat, L.J. Strahilevitz, Personalized default rules and disclosure with big data, Mich. Law Rev. 112 (8) (2014) 1417–1478.
[29] C.R. Sunstein, Impersonal Default Rules vs. Active Choices vs. Personalized Default Rules: A Triptych [Online, SSRN, 2012, https://ssrn.com/abstract=2171343, 20/01/2020.
[30] C.R. Sunstein, The distributional effects of nudges, Nat. Human Behav. (2021), https://doi.org/10.1038/s41562-021-01236-z.
[31] B.D. Bernheim, A. Rangel, Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics, Q. J. Econ. 124 (1) (2009) 51–104.
[32] C.R. Sunstein, Behavioral welfare economics, J. Benefit-Cost Anal. 11 (2) (2020), 196-2Tha20.
[33] E. Peer, S. Egelman, M. Harbach, N. Malkin, A. Mathur, A. Frik, Nudge me right: personalizing online security nudges to people's decision-making styles, Comput. Hum. Behav. 109 (2020), 106347.
[34] C. Schöning, C. Matt, T. Hess, Personalised nudging for more data disclosure? On the adaption of data usage policies format to cognitive styles, in: Proceeding of the 52nd Hawaii International Conference on System Sciences, 2019, pp. 4395–4404.
[35] S. Viljoen, Democratic Data: A Relational Theory For Data Governance [Online], SSRN, 2020, https://ssrn.com/abstract=3627562, 11/12/2020.
[36] C.R. Sunstein, Nudges that fail, Behaviour. Public Pol. 1 (1) (2017) 4–25.
[37] A. Tor, Nudges that should fail? Behaviour. Public Pol. 4 (3) (2020) 316–342.
[38] S. Della Vigna, E. Linos [Online]*RCTs to Scale: Comprehensive Evidence from Two Nudge Units*, https://eml.berkeley.edu/~sdellavi/wp/NudgeToScale2020-03-20.pdf, 2020, 09/03/2021.
[39] L. Cao, In-depth behavior understanding and use: the behavior informatics approach, Inf. Sci. 180 (17) (2010) 3067–3085.
[40] S. Zuboff, Big other: surveillance capitalism and the prospects of an information civilization, J. Inf. Technol. 30 (2015) 75–89.
[41] H.M. Bang, S.B. Shu, E.U. Weber, The role of perceived effectiveness on the acceptability of choice architecture, Behaviour. Public Pol. 4 (1) (2020) 50–70.
[42] H. Bruns, E. Kantorowicz-Reznichenko, K. Klement, M.L. Jonsson, B. Rahali, Can nudges be transparent and yet effective? J. Econ. Psychol. 65 (2018) 41–59.
[43] F.M. Kroese, D.R. Marchiori, D.T. de Ridder, Nudging healthy food choices: a field experiment at the train station, J. Publ. Health 38 (2016) 133–137.
[44] M. Steffel, E.F. Williams, R. Pogacar, Ethically deployed defaults: transparency and consumer protection through disclosure and preference articulation, J. Market. Res. 53 (2016) 865–880.
[45] G. Loewenstein, C. Bryce, C. Hagmann, S. Rajpal, Warning: you are about to be nudged, Behavioral Sci. Pol. 1 (1) (2015) 35–42.
[46] C. Öhman, N. Aggarwal, What if Facebook Goes Down? Ethical and Legal Considerations for the Demise of Big Tech, Internet Policy Review, 2020, https://doi.org/10.14763/2020.3.1488.
[47] J. van Dijck, Dataification, dataism and dataveillance: big Data between scientific paradigm and ideology, Surveill. Soc. 12 (2) (2014) 197–208.
[48] D. Susser, Invisible influence: artificial intelligence and the ethics of adaptive choice architectures, AIES'19 (2019), https://doi.org/10.1145/3306618.3314286.
[50] C.R. Sunstein, The storrs lectures: behavioral economics and paternalism, Yale Law J. 122 (2013) 1826–1899.

[51] H.S. Sætra, S. Mills, Psychological interference, liberty, and technology, Technol. Soc. 69 (2022), 101973.
[52] C.R. Sunstein, Why Nudge? the Politics of Libertarian Paternalism, Yale University Press, USA, 2014.
[53] R.H. Thaler, Nudge, not sludge, Science 361 (6401) (2018) 431–432.
[54] M. Aiken, The Cyber Effect, John Murray, UK, 2017.
[55] S. Turkle, The Second Self: Computers and the Human Spirit, MIT Press, USA, 2004, 1984.
[56] C.R. Sunstein, Sludge audits, Behaviour. Public Pol. (2019), https://doi.org/10.1017/bpp.2019.32.
[57] C.M. Gray, Y. Kou, B. Battles, J. Hoggatt, A.L. Toombs, The dark (patterns) side of UX design, in: CHI'2018 Conference on Human Factors in Computing Systems, vol. 534, 2018, pp. 1–14.
[58] K. Hosanagar, V. Jair, We Need Transparency In Algorithms, But Too Much Can Backfire [Online], Harvard Business Review, 2018, https://hbr.org/2018/07/we-need-transparency-in-algorithms-but-too-much-can-backfire, 15/03/2022.
[59] P. Lorenz-Spreen, S. Lewandowsky, C. Sunstein, R. Hertwig, How behavioural sciences can promote truth, autonomy and democratic discourse nline, Nat. Human Behav. (2020), https://doi.org/10.1038/s41562-020-0889-7.
[60] J. Zittrain, Engineering an election, Harv. Law Rev. 127 (2014) 335–341.
[61] V. Luckerson, Here's How Facebook's News Feed Actually Works [Online], Time Magazine, 2015, https://time.com/collection-post/3950525/facebook-news-feed-algorithm/, 08/03/2021.
[62] S. Davies, We Cracked the Code on How the Facebook News Feed Algorithm Works [Online], The Federalist, 2014, https://thefederalist.com/2014/02/20/we-cracked-the-code-on-how-the-facebook-news-feed-algorithm-works/, 08/03/2021.
[63] P. Cooper, How The Facebook Algorithm Works In 2021 and How To Make It Work For You [Online], Hootsuite, 2021, https://blog.hootsuite.com/facebook-algorithm/, 08/03/2021.
[64] H.S. Sætra, 'Privacy as an Aggregate Public Good' *Technology In Society*, vol. 63, 2020, 101422.
[65] S. Zuboff, The Age of Surveillance Capitalism: the Fight for a Human Future at the New Frontier of Power, Profile Books, UK, 2019.
[66] S. Delacroix, N.D. Lawrence, Bottom-up data Trusts: distributing the 'one size fits all' approach to data governance, International Data Privacy Law 9 (4) (2019) 236–252.
[67] D. Bates, The political theology of entropy: a Katechon for the cybernetic age, Hist. Hum. Sci. 33 (1) (2020) 109–127.
[68] E. Morozov, To Save Everything Click Here: Technology, Solutionism, and the Urge To Fix Problems That Don't Exist, Allen Lane, UK, 2013.
[69] S. Mills, H.S. Sætra, The Autonomous Choice Architect, AI and Society, 2022, https://doi.org/10.1007/s00146-022-01486-z.
[70] S. Russell, Human Compatible: AI and the Problem of Control, Penguin Books, UK, 2019.
[71] N. Srnicek, Platform Capitalism, Polity Books, UK, 2016.
[72] R. Booth, Facebook Reveals News Feed Experiment to Control Emotions [Online], The Guardian, 2014, https://www.theguardian.com/technology/2014/jun/29/facebook-users-emotions-news-feeds, 08/03/2021.
[73] A.D.I. Kramer, J.E. Guillory, J.T. Hancock, Experimental evidence of massive-scale emotional contagion through social networks, Proc. Natl. Acad. Sci. USA 111 (24) (2014) 8788–8790.
[74] Wall Street Journal, 'Inside TikTok's Algorithm: A WSJ Video Investigation [Online], 2021, https://www.wsj.com/articles/tiktok-algorithm-video-investigation-11626877477, 28/10/2021.
[75] M. Schrage, '*The Transformational Power Of Recommendation* [Online], MIT Sloan Management Review, 2020, https://sloanreview.mit.edu/article/the-transformational-power-of-recommendation/, 15/03/2021.
[76] P. Meissner, C. Keding, The Human Factor in AI-Based Decision-Making [Online], MIT Sloan Management Review, 2021, https://sloanreview.mit.edu/article/the-human-factor-in-ai-based-decision-making/, 15/03/2021.
[77] G. Tanner, The Hours Have Lost Their Clock, *The Politics of Nostalgia*' Repeater Books, UK, 2021.
[78] C.R. Sunstein, Misconceptions about nudges, J Behavioral Econ. Pol. 2 (1) (2018) 61–67.
[79] L. Bovens, The ethics of nudge, in: T. Grüne-Yanoff, S.O. Hansson (Eds.), *Preference Change: Approaches From Philosophy*' (2008), Springer, 2008.
[80] E.J. Johnson, D.G. Goldstein, Defaults and donation decisions, Transplantation 78 (12) (2004) 1713–1716.
[81] Beshears, J., Dai, Hengchen, Milkman, K. L., Benartzi, S. 'Using fresh starts to nudge increased retirement savings' Organ. Behav. Hum. Decis. Process., 167, pp. 72-87.
[83] L.C. Page, B.L. Castleman, K. Meyer, Customized nudging to improve FAFSA completion and income verification, Educ. Eval. Pol. Anal. 42 (1) (2019) 3–21.
[85] M. Heidegger, Being and Time, State University of New York Press, USA, 2010.
[86] P.W. Schultz, J.M. Nolan, R.B. Ciadlini, N.J. Goldstein, V. Griskevicius, The Constructive, destructive, and reconstructive power of social norms, Psychol. Sci. 18 (5) (2007) 429–434.
[87] F. Pasquale. The Black Box Society: The Secret Algorithms that Control Money and Information, Harvard University Press, USA, 2016.
[88] Weinmann, M., Schneider, C., vom Brocke, J. (2016) '*Digital Nudging*' SSRN. [Online] [Date accessed: 30/01/2020]: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2708250.
[89] C. Mele, T.R. Spena, V. Kaartemo, M.L. Marzullo, Smart nudging: How cognitive technologies enable choice architectures for value co-creation, J. Busi. Res. 129 (2021) 949–960.