Economics and Philosophy, **34** (2018) 135–163 © Cambridge University Press
doi:10.1017/S026626711700030X First published online 9 January 2018
cambridge.org/eap

HARNESSING HEURISTICS FOR ECONOMIC POLICY

Ramzi Mabsout*, **Jana G. Mourad**†

Abstract: The effectiveness of heuristics has received contradicting interpretations in the behavioural sciences. We study the policy implications of two programmes that dispute the effectiveness of heuristics – the biases and heuristics and the fast and frugal heuristics programmes. While the first blames heuristics for most errors in judgement, the second posits heuristics as simple mental algorithms that work well in a range of environments. We argue that the fast and frugal programme is less paternalistic insofar as it models humans as effective decision-makers in a range of environments. However, in the rapidly changing environments of the 21st century, both are needed to inform evidence-based policies.

Keywords: Biases and heuristics, economic policy, fast and frugal heuristics, libertarian paternalism, nudges

1. INTRODUCTION

The biases and heuristics (BH) and the fast and frugal heuristics (FFH) programmes are scientific rivals competing over which has the better model of decision-making heuristics. Their disagreement has not been limited to an obscure academic dispute. Each programme reflects a different image of individuals' reasoning abilities offering distinct justifications for policy and regulation. Our argument is that the programmes are different – they capture different aspects of human judgements – and while FFH can avoid some moral pitfalls the policy arm of BH has been subjected to, the two complement each other and ought to be considered in tandem for policy purposes.

* American University of Beirut, P.O. Box 11-0236, Department of Economics, Beirut, Lebanon. Email: rm95@aub.edu.lb. URL: https://www.aub.edu.lb/fas/economics/Pages/mabsout.aspx.
† P.O. Box 70-890, Antelias, Lebanon. Email: janagmourad@gmail.com.

The BH programme of Tversky and Kahneman (1974) was in part a response to earlier research that positively benchmarked the intuitive statistical abilities of individuals to 'statistical man' (cf. Peterson and Beach 1967). It started as a systematic descriptive search for decisional deviations from normative rules (logical and probabilistic). Their descriptive programme, Kahneman and Tversky (1984: 341) argued, 'is concerned with people's beliefs and preferences as they are not as they should be'. The culprits are so-called mental heuristics that hijack reasoning whenever a person has a difficult mental task.

As BH made its way into economics, policy implications began to emerge in the early 2000s. Economists, philosophers, as well as legal scholars formulated soft and hard paternalistic policies including libertarian paternalism (Sunstein and Thaler 2003), light paternalism (Loewenstein and Haisley 2008), asymmetric paternalism (Camerer *et al*. 2003), and coercive paternalism (Conly 2014). All of them deploy the deviations in BH to justify the necessity of interventions that save individuals from their self-harming decisions. The argument from behavioural anomalies to paternalism is straightforward: To the extent that individual behaviour can be predictively biased, decisions can result in harm to self (smoking, under-saving, over-optimism, over-eating, etc.). A paternalistic intervention is thus rendered necessary to incentivize, guide, nudge or coerce biased individuals towards welfare-enhancing choices. Over the last few years various governments across the globe have adopted the BH policy implications (Sunstein 2013*b*; Halpern 2015), including global organizations such as the World Bank (2015) and the OECD (2017).

Notwithstanding its success, BH and the paternalistic policies that ensued motivated a significant critique. The focus of this paper is what appears to be the most powerful challenger: The fast and frugal heuristics programme of Gigerenzer and the ABC group (Gigerenzer and Todd 1999; Gigerenzer *et al*. 2002). The FFH critique is effective because it provides an alternative interpretation of heuristics, alternative normative judgement rules and alternative policy prescriptions. FFH posits that even if individuals had 'paid full attention and possessed complete information, unlimited cognitive abilities, and complete self-control' (Thaler and Sunstein 2008: 6) they would not necessarily obtain better decisional outcomes.

This critique brings to the fore two questions we will pursue: First, to what extent are FFH and BH comparable descriptive programmes? And second, in changing the descriptive base, how will that reflect on policy implications? While FFH is critical of the descriptive, normative and policy implications of BH, our answer to the first question turns it around by adopting Hammond's (1996) view that FFH and BH are not comparable but complementary and should not be in (direct)

competition. As to the second question – changing the descriptive base – it has implications that may deflect the value substitution and coercion critiques which libertarian paternalism (LP) has been subjected to.

While we concede that in some circumstances paternalism is not avoidable (and may even be desirable), it is beyond the scope of our paper to argue for an optimal level of paternalism. Instead, from our point of view, if the same policy objective can be achieved with less paternalism, it ought to. We are in broad agreement with Dworkin (1983: 34) who argued that if 'there is an alternative way of accomplishing the desired end without restricting liberty, even though it may involve great expense inconvenience etc., society must adopt it'.

In the next section we discuss the links between BH and LP as well as major criticisms. Libertarian paternalists argue it is libertarian because it focuses on means, not ends, and it preserves freedom of choice (individuals, it is claimed, can still opt out from the nudges of the choice architect if they know what they want). It is paternalist because its policies steer individuals towards welfare enhancing choices. To be specific, nudges are the key policy instruments that steer individuals towards the choices that best serve their interests without restricting choice sets. To preserve freedom of choice, nudges should have little to no opportunity cost for rational individuals. But LP (and the nudge programme) were criticized and towards the end of Section 2 we review the strongest criticisms. In Section 3, a different descriptive base is introduced, namely the FFH programme which has distinct policy prescriptions. In Section 4 we compare the FFH policy implications to BH. In Section 5 Hammond's (1996) meta-theoretical taxonomy between correspondence and coherence theories of judgement is introduced to suggest that BH and FFH are in fact complementary rather than substitutes. According to Hammond (1996: 351), the tension between correspondence and coherence meta-theories cannot be eliminated because it is 'rooted in our nature just as language and tool building are rooted in our nature'. Section 6 discusses the implications of the correspondence and coherence meta-theories for policy. Section 7 concludes.

2. LIBERTARIAN PATERNALISM AND ITS CRITICS

2.1. The biases and heuristics descriptive programme (BH)

In the early 1970s Tversky and Kahneman launched the BH programme, which owes its success in large part to them (Kahneman 2011). Their 1974 review of errors in judgments, published in *Science*, opened up the field of BH to other disciplines including medicine, law, political science and economics (Gilovich and Griffin 2002). Tversky and Kahneman's (1974) major discovery was that human judgement

under certainty and uncertainty depends on heuristics that bias decisions against normative benchmarks (as set by the rules of logic and probability). It was a bleak view that led to an almost new field of study in a few years, to more anthologies (Kahneman *et al*. 1982; Gilovich *et al*. 2002), and to a dozen books for the general public arguing that human judgement is not to be trusted over a wide range of tasks.

Tversky and Kahneman define heuristics as a limited number of principles that people resort to in order to 'reduce the complex tasks of assessing probabilities and predicting values to simpler judgmental operations' (1974: 1124). They admit that despite being helpful, heuristics often yield systematic biases. These biases are deviations from the normative rules of logic – for instance the evaluation of conjunctive and disjunctive events – and probability – for example base rate negligence, insensitivity to sample size and predictability, misconception of chance and regression, and the illusion of validity and correlations, among others. The biases were also assimilated to visual illusions in assessing distance and size.[1] Kahneman and Tversky (1974: 1130) conclude that biases are difficult to correct as even incentivized individuals and trained statisticians were unable to overcome the pull of heuristics, a point we will come back to.

In his most recent book Kahneman (2011) synthesizes over 40 years of work on biases redefining heuristic in terms of 'a simple procedure that helps find adequate, though often imperfect answers to difficult questions' (2011: 98); persists in arguing that 'even when the heuristic has some validity, exclusive reliance on it is associated with grave sins against statistical logic' (2011: 151); explains that emotions are heuristics too and that the importance of the affect heuristic in particular has been neglected (2011: 12); and embeds heuristics in system 1 – automatic, quick, hardly trainable, unconscious mental processes – to be contrasted from system 2 which are slower, computational, conscious and cognitively demanding. Libertarian paternalism – the policy arm of BH also endorsed by Kahneman (2011: 413–14) – adopted this taxonomy as its proponents posit biases 'are firmly rooted in the operations of system 1' (Sunstein 2013*a*: 1842; see also Thaler and Sunstein 2008: 21–4).

In economics, the work of Tversky and Kahneman was a major impetus in the re-introduction of psychology after Herbert Simon's

[1] In fact, these errors must be close to incorrigible or else the programme could not attract more interest than observing that English native speakers commit errors and fallacies when speaking Japanese (Cosmides and Tooby 1996: 10) or how the discovery that 'persons untrained in differential calculus make errors when asked to solve calculus problems, whereas those trained in calculus do not … would hardly have attracted attention' (Hammond 1996: 206).

and others' frustrated attempts (Sent 2004). The field of behavioural economics – which most libertarian paternalists use as the descriptive base for their policy prescriptions – has experienced tremendous growth and is now well integrated in economics (Thaler 1991, 2002; Conlisk 1996; Rabin 1998; Camerer *et al*. 2004; see Heukelom 2014 for historical background). Another important feature of the BH programme inherited by behavioural economists is the assimilation of errors in judgement to speech errors, forgetting and optical illusions (Thaler 1991; Camerer and Loewenstein 2004; Loewenstein and O'Donoghue 2006: 183). Thaler (1980: 40), an early adopter, argued mental illusions are the rule rather than the exception. The assimilation of reasoning errors to optical illusions has implications for learning and policy as we shall see.

While the list of biases that justify LP can be quite large, Thaler and Sunstein (2008) emphasize the following: (i) Anchoring, (ii) Availability, (iii) Representativeness, (iv) Optimism and overconfidence, (v) Gains and losses, (vi) Status quo bias, (vii) Framing, (viii) Temptation, and (ix) Following the herd. Thaler and Sunstein (2008) note that Tversky and Kahneman (1974) originally identified only the first three as heuristics, however, Kahneman's (2011) more recent definition of heuristics is broader and may now accommodate other elements including emotional ones.

2.2 Libertarian paternalism and nudges

The initial focus of BH was to catalogue deviations from normative rules of logic and probability. Thaler's (1980) initial interest, for example, was in how consumers act in ways inconsistent with economic theory. He argues that BH merely offers an alternative description (prospect theory) and 'does not necessarily imply any criticism of the axioms of rational choice as a normative ideal' (1991: 138) nor are there signs that he is interested in the policy implications of the biases. In 2003, the year Sunstein and Thaler and Camerer *et al*. published their works on libertarian and asymmetric paternalism respectively, Berg (2003: 411) noted that behavioural economics failed 'to penetrate into contemporary discourses on leading policy issues'. As late as 2008, Sugden (2008: 227) remarked, 'until quite recently the programme was an almost wholly descriptive enterprise'. However, a change in rhetoric is perceptible in the early 2000s with respect to the policy implications (but not the normative benchmarks) as behavioural economics consolidated its position in mainstream economics (Berg and Gigerenzer 2010). The rationale for the extension of BH to policy was straightforward: Deviations from normative principles – the biases – have welfare consequences and hence paternalistic interventions are necessary to protect individuals

from their own flawed decisions. More specifically, Sunstein (2013*a*: 1826) argues that the biases create behavioural market failures that justify paternalism.

Paternalism is encountered in a variety of domains, in both the private and public spheres. Most definitions of paternalism state that it involves an action by some agent A (parent, teacher, government, choice architect) on another agent B to protect or promote the well-being of B. Dworkin (2016) defines a paternalistic action as one that involves 'the interference of a state or an individual with another person, against their will, and defended or motivated by a claim that the person interfered with will be better off or protected from harm'. Not all definitions of paternalism adopt 'against their will' since many, including LP, assume that B would have wanted the intervention to be made if it knew it would promote own well-being.

Sunstein and Thaler (2003: 1162) define a policy as paternalistic if it 'attempts to influence the choices of affected parties in a way that will make choosers better off'. It follows that an intervention by A to modify the means or ends of B because they cause an externality (but no self-harm) is not paternalistic. Paternalists argue that if we accept interventions to correct market failures caused by externalities we should also accept interventions caused by a different type of market failure, namely, behavioural market failures (Sunstein 2014: 16, 20). According to LP, an intervention is morally justified when individual actions cause harm to self or when choices do not promote the achievement of own ends (Sunstein 2013*a*: 1845; 2014: 34).

Biases in judgements are used to support a range of paternalistic interventions, from the coercive to the soft. Conly (2014: 43), for instance, argues that paternalistic interventions are justified when 'people's choices of instrumental means are confused, in a way that means they will not achieve their ultimate ends'. She defends a coercive form paternalism that bans self-harming choices. Such bans are justified even when individuals are fully informed about the consequences of their actions. For Conly, autonomy can be bypassed when cost benefit analysis reveals that the harms to self are more significant than preserving autonomy and freedom of choice. Though they differ on the policy tools, soft paternalism and coercive paternalism are not different kinds of paternalisms since both use the same criteria to justify paternalistic interventions (Sunstein 2014: 18, 22).

More specifically, to pre-empt the classic Millian critiques of paternalism – that it is coercive, that it undermines autonomy, freedom and responsibility – Sunstein and Thaler (2003) claim that LP avoids full-fledged paternalism by occupying a middle ground on a continuum with paternalism occupying one extreme and libertarianism the other. Even though the concepts of libertarianism and paternalism seem to be

oxymorons, with the former purporting freedom of choice and the latter denying it, Thaler and Sunstein claim to have reconciled them: LP is soft because it does not use bans and it targets means not ends to help preserve freedom of choice (Sunstein 2014: 19). Accordingly, libertarian paternalists do not endorse policies that constrain individual choice, advocating instead policies that engender a choice architecture in which it is easier for individuals to achieve their own ends. It is, nevertheless, paternalistic since it grants public institutions power to design a choice architecture that steers peoples' choices towards (reconstructed) welfare-promoting ends.

Nudges are an important if controversial tool in the policy kit of the libertarian paternalist. Nudges are not necessarily paternalistic as they can be used to change a behaviour engendering an externality (Hausman and Welch 2010; Rebonato 2012). A nudge is 'any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives' (Thaler and Sunstein 2008: 6). More recently, Sunstein (2014: 17) re-defines nudges as 'initiatives that maintain freedom of choice while also steering people's decisions in the right direction (as judged by people themselves)'. Thaler and Sunstein (2008) claim that since LP imposes insignificant material and psychological costs on rational individuals who wish to diverge from the planner's welfare enhancing choice, its interventions qualify as nudges. Nudges combine means and soft paternalism and must be costless or nearly costless to be avoided for otherwise they risk constraining individual choice. Nudges should be used when choices are separated from their consequences (and self-control problems arise), when tasks are difficult (selecting a bank mortgage), when choices are infrequent (buying a car, choosing a spouse), when feedback is irregular, and when we don't know what we like (the problem of translating choices into experiences) (Thaler and Sunstein 2008: 79–84).

In free market democracies, people continuously decide what and when to buy, how much to save, where to work, or whom to vote for. According to Sunstein (2013*a*), most of these decisions will be affected by the existing choice architecture. The intervention, for Sunstein (2013*a*) as for Conly (2014), is based on an evaluation of the total welfare effects or the overall costs and benefits of an intervention: If the total welfare effect is positive, libertarian paternalists argue, an intervention based on means rather than ends and which safeguards liberties is justified.

Libertarian paternalists have been subjected to criticisms on several fronts including the descriptive programme, the normative benchmarks and the policy prescriptions. Since in Section 3 we consider an alternative descriptive programme, we focus next on the value substitution and coercion critiques.

2.3. The value substitution and coercion critiques

The emergence of new forms of paternalism supported by BH sparked many criticisms.[2] There is quite some overlap and repetition between these critiques so we focus on the strongest arguments against LP. The basic claim of libertarian paternalists is that nudges steer people towards choices that they would have made had they been unboundedly rational – employees reporting that their savings rate is too low and who wish to save more for retirement, will be nudged by Save-More-Tomorrow to increase savings. This claim is based on assuming (i) People's interests are disclosed to the libertarian paternalist and (ii) The libertarian paternalist is a better judge of one's welfare than the individual himself. These two assumptions have focused the attention of the critics on how the choice architect comes to know the true interest of the biased individual about to be nudged. White (2013) argues that an outside observer (planner, choice architect) can never uncover the real motives behind people's choices. In fact, individuals themselves might not know their own motives. If individuals' preferences are inconsistent or unknown, how can the libertarian paternalist guess what they are? White (2013) states that since the only information available to libertarian paternalists regarding individual's true interests are actual choices, the libertarian paternalist dismisses their diverse and private conception of their own welfare to impose a reconstruction of what promotes it (see Sugden 2008 for an earlier argument along similar lines). Thus, welfare substitution – which entails the 'usurping of judgment about the way individuals structure and run their lives' (White 2013: 74) – arises rendering LP morally problematic. A somewhat similar critique is formulated by Rebonato (2012) who argues that the libertarian paternalist must reconstruct the unbiased preferences of system 2 based on observations generated from the biased choices of system 1. Rebonato claims this is not a feasible task because the preferences of biased individuals change with context and these individuals are unlikely to have an empirically identifiable consistent set of preferences.

The libertarian paternalists, as we have seen, justify the restrictions on freedom with increments to welfare generated by the nudge. Cost–benefit analysis is deployed to determine if the (welfare) gains are larger than the losses (to freedom). Cost–benefit analysis, however, neglects the problem of commensurability between welfare and freedom, a problem Sunstein (1996) acknowledged in earlier works but somehow ignores

[2] See, among others, Klein (2004), Glaeser (2006), Saint Paul (2011), Goodwin (2012), Grüne-Yanoff (2012), Rebonato (2012), White (2013), Whitman and Rizzo (2015) and Kosters and Van der Heijden (2015).

when he writes about LP.[3] In fact, no framework is offered in LP to compute trade-offs between freedom and welfare. Libertarians have (for as long as they existed) argued that freedom should be valued not for its contribution to welfare but for itself as it preserves autonomy, dignity and responsibility. Buchanan (1978 [1999: 239]) thus observed 'man does not want liberty in order to maximize utility. He wants liberty to become the man he wants to become.' More recently, Saint-Paul (2011: 127), concerned with the legitimation of paternalism by behavioural economists, warns: 'Each freedom is conditional on science putting an approval on it. The day it is proven that having that freedom statistically has harmful consequences, it may be revoked.' Where the red lines of the paternalist should be drawn becomes contingent on the findings of science. Without the constitutional protection of freedom, evidence-based paternalist policies can bring about a new era of scientific tyranny.

In addition to the above, Grüne-Yanoff (2012) argues that libertarian paternalistic policies are coercive due to their interference with individual decision-making. He suggests that the definition of coercion offered by Thaler and Sunstein (2008) is too narrow as it encompasses only conditional threats or direct force. Grüne-Yanoff suggests a broader conception of coercion based on Berlin's (1969) definition whereby coercion 'implies the deliberate interference of other human beings within the area in which I could otherwise act' (quoted in Grüne-Yanoff 2012: 638). The coercion critique entails that the libertarian credentials of LP are misleading. Biased individuals whose preferences are constructed by the environment will not be aware that their choices are being steered towards the ends the libertarian paternalist claims they would have wanted. Nudges exploit biased minds to achieve that goal. Rational individuals, who know what they want, are not affected by the nudge. But for biased individuals this is a coercion towards the choice preferred by the choice architect who, according to the value substitution critique, has no informational access to the welfare enhancing preferences of the individual. It follows that rather than safeguarding the freedom of biased individuals, LP transforms their original choice set into one where the nudged choice is the only one left. This is done without firm grounds to believe that the preferences have been discovered rather than constructed by the choice architect. Because of the value substitution

[3] Sunstein (1996: 780) wrote 'human values are plural and diverse … In a way that are not reducible to some larger and more encompassing value … human goods are not commensurable … such goods are not assessed along a single metric". These claims are hard to reconcile with his more recent argument that welfare trumps freedom or that 'some sacred values, and some refusals to consider tradeoffs, are best seen as moral heuristics, simplifying decision tasks' (Sunstein 2005: 568).

critique and the absence of consistent identifiable preferences, LP failed to convince us that this is what biased individuals would have wanted anyway.

Thaler and Sunstein (2008) and Sunstein (2013*a*, 2014) acknowledge the power of these critiques and offer responses. For Sunstein (2014), the epistemic principle (his equivalent term for the value substitution critique) cannot dismiss all paternalistic policy interventions, especially those that have unambiguous positive welfare effects and that are not too costly to implement. For more complex interventions, he recommends tailoring them for individual preferences or what he calls personalized paternalism. For example, to help people optimize savings, he recommends that policies should be designed for specific age groups since saving needs differ with age in predictable ways. Much more empirical work will be required, Sunstein acknowledges, to uncover what those individual preferences are. Sunstein's response, however, does not account for the quickly rising cost of personalized paternalism. Furthermore, most case studies that demonstrate the effectiveness of LP, including seatbelts and smoking, are cases in which the well-being of the individual – or what is in her or his interest – is unambiguous. But in other settings – including savings, nutrition, recreative drug use or physical activity – interventions will be increasingly and controversially personal, meshing ends and means and loosening the soft edge of LP. On the coercion critique, Thaler and Sunstein (2008: 251) admit that with respect to mandates and bans 'deciding where to stop, and when to call a nudge a shove (much less a prison) is tricky' (2008: 251) and 'that there are no hard-and -fast stopping points' (2008: 249). These conclusions re-emphasize our earlier claim that the difference between coercive paternalism and LP is one of degree not kind. And with respect to a range of different paternalist regulations it follows that whether they will be classified as libertarian or coercive is not without ambiguity.

Can changing the descriptive programme deflect the critiques of LP? This is the question we consider next. We begin with an outline of the fast and frugal heuristics programme followed by the kind of policy prescriptions it is able to generate (Section 3). Section 4 compares those policy prescriptions to LP.

3. THE FAST AND FRUGAL HEURISTICS PROGRAMME (FFH)

Thaler and Sunstein base LP on the findings of the BH programme. In what follows, we introduce a competing programme that provides a different descriptive base, namely, FFH. Both programmes trace their origins to Simon's influential work on bounded rationality. The FFH programme does not deny individuals have limited computational

abilities – their bounded rationality – though its object of study are heuristics that BH blames for most errors in judgments. Heuristics, according to Todd and Gigerenzer (2007: 167), are 'simple decision making algorithms that can work well in appropriate environments'. For Tversky and Kahneman (1974: 1124), we may recall, heuristics such as representativeness, availability and adjustment and anchoring are used in assessing probabilities and predicting values and, while useful, they 'lead to severe and systematic errors'.

Before the success of BH, Simon (1955: 99) argued that heuristics such as satisficing – 'that require only that the pay-off exceed some given amount' set by the aspiration level – in humans and computers are necessary for effective decisions. He thought so for two reasons, namely, humans have limited computational abilities, memory etc., and they live in complex environments where computation is often impossible and optimal solutions intractable (Gigerenzer and Selten 2001: 4). Simon (1955: 104) wrote,

> the task is to replace the global rationality of economic man with a kind of rational behavior that is compatible with access to information and the computational capacities that are actually possessed by organisms, including man, in the kinds of environments in which such organisms exist … Our procedure will be to introduce some modifications that appear to correspond to observed behavior processes in humans, and that lead to substantial computational simplifications all the time.

According to Gigerenzer and co-authors, Kahneman and Tversky tweaked this interpretation of heuristics from processes that make computers smart 'to one that explains why people are not smart' (Gigerenzer and Brighton 2009: 109; see also Gigerenzer and Todd 1999: 27). Hammond (1996: 107) notes in this regard that 'although [Kahneman and Tversky] deny the charge that they have denigrated the rationality and competence of human judgements, the charge remains, for the denial seems disingenuous in view of the persistent demonstration – and celebration – of ineptness'. The BH programme adopts an ambivalent stance towards heuristics, sometimes suggesting they are effective but more often systematically misleading. With the exception of work done by Tversky (1972) on the elimination by aspects heuristic, there is no test in Tversky and Kahneman's (1974) article where heuristics are shown to work (Lopes 1991: 77). With few exceptions, most subsequent experiments pursue the emphasis on how heuristics violate the normative laws of probability and logic (Kahneman 2011).

To Gigerenzer and Brighton (2009: 109), FFH sheds light on three misconceptions in BH: (i) heuristics are always second best, (ii) heuristics are used because of human cognitive limitations and (iii) more time,

information and computation is always better. FFH's major discovery is that less is more and simple heuristics, rather than causing systematic errors, can be ecologically more efficient than the more complex normative benchmarks of BH which are based on probability principles and logic.

Whether one adopts the findings of BH or not, FFH is an alternative research programme that appears to offer a radically different view of heuristics. The understanding of heuristics in FFH purports faithfulness to the original definition of heuristics as intelligent computer programs. Heuristics evolved over the Pleistocene (that lasted from about 2.5 million to 12,000 years ago) and are constituted from evolved capacities – 'the construction materials' – such as language, imitation, recognition memory and object tracking, among others. They are distinguished from traits, attitudes, preferences and other internal mental concepts. They are fast since they are simple; and are frugal because they use little information (Gigerenzer *et al*. 2002). Heuristics do not employ complex computation and most of them are one-reason decision algorithms – they are 'amazingly simple and efficient' (Gigerenzer and Brighton 2009: 108).

All heuristics are constructed from three building blocks (search, stop and decide) and can outperform more complex modes of judgements in a range of environments. Which heuristic is more effective depends on the environment – the problem of ecological rationality. Berg (2014*a*) argues that ecological rationality does not derive its normative benchmarks from an independent class of choice axioms, as the consistency approach to bounded rationality does, but 'adopts multiple, domain-specific performance metrics, reflecting the view that adequate descriptions of well-being are irreducibly multivariate' (2014*a*: 375). Let us consider two examples of fast and frugal heuristics and compare them to normative benchmarks when possible.

Example 1. The recognition heuristic entails that the recognized object has a higher value than the one that is not recognized. The recognition heuristic implies the selection of the familiar choice. Ecologically, 'the recognition heuristic is successful when ignorance is systematic rather than random that is when the recognition is highly correlated with the criterion' (Gigerenzer 2008: 25). In guessing which cities have more inhabitants, it states that 'If you recognize the name of one city but not that of the other, then infer that the recognized city has a larger population' (Gigerenzer 2014: 48). For this heuristic to work, however, subjects must recognize one of the two choices only. The effectiveness of the recognition heuristic was demonstrated in an experiment where German students had better prediction rates of the number of inhabitants of US cities than US students (Goldstein and Gigerenzer 1999). The recognition heuristic also successfully selected a portfolio of stocks that performed better than

the portfolios of financial experts (Ortmann *et al*. 2008). For behavioural economists, however, this heuristic generates a bias, namely, the home country bias where 'people in every country overinvest in the country they are familiar with – their own' (Camerer and Loewenstein 2004: 21). French and Poterba (1991) report that equity investors over-invest in their home markets. To rationalize the current bias in home country investments, local returns ought to be much higher than the actual level and – by not diversifying – investors are missing on the opportunity to increase their returns. The FFH programme would not interpret the home country bias as one but instead as a fast and frugal heuristic capable of beating optimal decision processes. Maybe Adam Smith (1776: sec. IV.2.9) had the recognition heuristics in mind when he wrote in *The Wealth of Nations* that 'by preferring the support of domestic to that of foreign industry, he intends only his own security; and by directing that industry in such a manner as its produce may be of the greatest value, he intends only his own gain, and he is in this, as many other cases, led by an invisible hand to promote an end which was no part of his intention'.

Example 2. The 1/N heuristic rule states 'Allocate your money equally to each of N funds' (Gigerenzer 2007: 26). This heuristic is quite interesting because Benartzi and Thaler (2001) had shown that individuals construct their investment portfolios based on the 1/N heuristics ('As the number of stock funds increases, so does the allocation to equities' (p. 96) as opposed to bonds) but they considered this strategy costly leading to suboptimal investment decisions away from the efficiency frontier or on the wrong point along the frontier. Demiguel *et al.* (2009), however, found that 1/N outperformed 14 Bayesian and non-Bayesian models of optimal portfolio allocations. Optimization models would need 3000 months (with N = 25 assets) and 6000 months (with N = 50 assets) to outperform 1/N according to their findings. As Gigerenzer (2008) explains, 1/N also predicts better out of sample because optimization models over fit the data. Which will outperform the other depends on three criteria: (i) the predictive uncertainty of the problem, (ii) the number of N assets and (iii) the size of the learning sample. The heuristic outperforms optimization models when uncertainty and N are large and when the learning sample – the number of years over which the analysis is conducted – is small.

There is no prima facie compelling empirical evidence that normative modes of judgement fare better in all environments. Gigerenzer and Brighton (2009) report that irrational players make more money than rational players in tit-for-tat games. Berg and Gigerenzer (2010) observe that deviations from rational decisions are not corroborated with worse information, lower well-being or profits. If the normative standards of BH across real world environments do not generate better outcomes, they are

not normative standards.[4] The discovery that fast and frugal heuristics could be more effective than the benchmarks provided by normative decision models led proponents of FFH to reassess the ideals of rationality away from a narrow and unique set of rules (Berg 2014*a*). Gilboa *et al*. (2009: 189), for example, reject the view that rationality is axiomatic and dichotomous (rational/irrational) arguing that it ought to be assessed using multiple criteria including correspondence with data. We should note, nevertheless, that since the normative benchmarks of FFH are not axiomatically derived they must be empirically discoverable (Todd and Gigerenzer 2007). FFH advocates the adoption of plural normative benchmarks that cannot be known a priori and whose effectiveness is gauged with respect to fallible outcome indicators (correct beliefs, economic payoffs, health, well-being etc.) and the environment. It follows from this that the choice of outcome indicators, alongside the environment in which decisions are made, are the appropriate variables to study the empirically determined normative benchmarks of heuristics.

4. THE FFH DESIGN-POLICY STAGE

After reviewing and comparing the descriptive and normative claims of FFH, we focus now on its policy prescriptions or the design stage (Gigerenzer 2008). The first and most obvious policy implication from FFH is that individuals should not benchmark judgements against normative rules. Since heuristics are ecologically efficient they help individuals reach goals in effective ways supporting policy prescriptions that leave things as they are. For Gigerenzer (2007: 151), 'the lesson is to trust your intuition when thinking about things that are difficult to predict and when there is little information'. In other words, everyday heuristics are doing a good job, so individuals ought not to be nudged in the direction of better choices.

While no policy is a form of policy, heuristics as effective mental tools require no external intervention and – accepting the discoveries of FFH – generate non-paternalistic policies in so far as humans already use effective heuristics in decision-making. Still, when it is discovered that a specific heuristic is mis-matched to its environment, FFH 'does not lay blame exclusively on badly behaved individuals or external causes in the environment … rather problems are diagnosed in terms of mis-matched decision process and environment … which suggests more degrees of freedom when prescribing corrective policy and new institutional design' (Berg and Gigerenzer 2010: 149). A policy intervention then may be

[4] See Hodgson (1993), Radner (1995), Berg and Gigerenzer (2007), Smith (2008), Sugden (2008), Gilboa *et al*. (2009), Rebonato (2012: sec 7), Berg *et al*. (2016). See also references and discussion in Conlisk (1996).

needed when a heuristic is consistently mis-matched by some individuals to an environment. When such a mis-match has been identified, FFH can prescribe the following: (i) use empirical research to identify effective heuristics – that is heuristics that work ecologically well in specific environments to solve specific tasks, (ii) identify instances in which individuals ineffectively use heuristics and (iii) help individuals adopt an effective heuristic matched to the environment.

More generally, we can deduce two types of FFH policy interventions. The first results from a mismatch between the heuristic and the environment whereas the second arises when the environment frames the problem in terms that mislead individuals, for example, when the problem is described in probabilistic (or conditional probabilistic) terms instead of frequencies.

If the environment is known (and relatively stable), FFH can identify an alternative heuristic when the current one is ecologically mismatched to the environment. An individual using a mismatched heuristic, for example, can drop it and adopt another: 'Individuals can certainly be led to use particular heuristics in inappropriate environments and consequently make errors, as the heuristics and biases tradition emphasized' (Todd and Gigerenzer 2007: 168). This replacement will, arguably, lead to ecologically more effective decisions. But how can the recommendation to change the heuristic be made while deflecting the critiques LP has been subjected to? At first, one can sense that the critiques cannot be avoided including the value substitution critique. Yet FFH may avoid the value substitution critique if it merely replaces a mismatched heuristic without encroaching on unobservable preferences.

This possibility raises two difficulties. First, if the policymaker is perceived as imposing a change in heuristic, this could be interpreted as a coercion (see the Berlin quote in Section 2). Whether freedom or autonomy will be diminished will depend on how the intervention is implemented and on its relationship with individual ends. This brings us to the second concern, namely, that the change of heuristic can only be recommended if there is some end or preference towards which the heuristic is mis-adapted. Here, the value substitution critique rears its head, if indirectly, since a change in heuristic is advisable if the current heuristic is unsuccessful at realizing some set of preferences. But the value substitution critique states that consistent welfare enhancing preferences are not accessible to the policy-maker (maybe even to the person herself), hence the change in heuristic cannot be implemented because no policy can be tailored to the specific and true preferences of individuals.

Thus the problem of the correct identification of individual preferences brings back the value substitution critique. If the decision-maker – whoever that may be – is advised to adopt a different heuristic that better serves her preferences (based on what the policy-maker

believes her preferences are), then, indeed, the value substitution has not been disarmed. Nevertheless, in FFH, the reasoning is not faulty but mismatched to the environment whereas in BH the reasoning is at fault – it does not follow logical or probabilistic normative rules. The intervention required from FFH remains less coercive in the sense that it is more transparent in comparison to its libertarian paternalistic counterpart, which requires the decision-maker not to be aware of the changes, partly because learning logical and probabilistic principles is more difficult and partly because the biases are like visual illusions and cannot be dislodged.

The policy design stage of FFH could be immune from another related difficulty in LP. This possibility, suggested by McQuillin and Sugden (2012), is that some biases, such as the status quo bias, are deep seated and may 'reflect a fundamental asymmetry in human desires' (see also Sugden 2008). If so, informing individuals about it would not correct behaviour which 'would continue to reveal incoherent preferences' (McQuillin and Sugden 2012: 560). While, on the one hand, this brings back the analogy that mental biases are like optical illusions, it is not very effective against LP because nudges do not seek to teach people how to reason correctly. FFH, on the other hand, rejects the optical illusion analogy and may not interpret the status quo bias as a bias but a fast and frugal heuristic. If there is no mis-match between individuals' ends, the heuristics and the environment, a paternalistic intervention is not necessary. FFH prescribes a policy that involves the adoption of a different heuristic only if the current heuristic is deemed ecologically mismatched.

Can the FFH design stage be less paternalistic than LP when it does not seek to modify the heuristic already in use but instead change the framing of the task from probabilities to frequencies? Rather than assuming that people are making incorrect judgements based on some normative logical or probabilistic benchmark, FFH blames the environment of the decision task for not being suitable to the heuristics people use in everyday life. In FFH too, biases vanish if information is offered in a format that suits the human mind (for example, frequencies instead of probabilities and percentages cf. Cosmides and Tooby 1996; Hoffrage and Gigerenzer 2004). Experiments that modify the format of tasks from numbers and letters to terms used in everyday reasoning dramatically changed the proportion of correct answers in the Wason selection task (Evans 2003) whereas reframing problems in terms of frequencies has been shown to help children solve Bayesian problems (Gigerenzer 2008). This is not very distinct from certain policy proposals in LP that recommend simplifying the choice architecture in complex problems into formats that facilitate effective decision-making and learning (Thaler and Sunstein 2008: 103–6). Kahneman and Tversky (1996) were also not oblivious to this finding and argued that, before FFH, they discovered the frequency effects (Tversky and Kahneman 1983). We

cannot then conclude that the design stage of FFH is less paternalistic than LP when it comes to changing the framing since both the FFH design stage and LP can support similar interventions that increase the proportion of correct answers.

Whether a change in heuristic or a modification in the information format of the environment should be adopted is a complex question to be further explored in Section 6. In some cases, it may not be possible to change the wording of the choice problems and hence it would be necessary to modify the heuristic individuals use when navigating different environments. For example, if there is uncertainty as to which environment individuals will face, as is likely to be the case in the world outside the lab, it would not be possible beforehand to modify all frames to trigger the adapted heuristic.

Both BH and FFH are blocked from policy prescriptions if their implications are deemed interferences with individual freedom. There is nothing to be changed if preferences are not to be tampered with. In our view such claims require a different kind of response, namely, one to be found in moral philosophy, not in the descriptive programmes of FFH or BH. Although such arguments would take us beyond our stated aims, policy interventions have to be contained in a moral framework (not just CBA) that argumentatively allows interference with freedom. Even though FFH avoids some pitfalls in the BH policy interventions, it may nevertheless not be immune from the value substitution and coercion critiques.

5. SORTING OUT THE CONFLICT: COHERENCE AND CORRESPONDENCE META-THEORIES

In Section 4 we argued that the policy implications of FFH can be less paternalistic than BH, avoiding some (but not all) of the criticisms that dog the LP policy agenda. FFH, for example, generates fewer paternalistic interventions because the heuristics already in use are effective mental tools, on the one hand, whereas, on the other hand, both LP and FFH support policies that would reformulate decision problems to improve performance.

The dispute between FFH and BH appears to be quite intractable and both theories are entrenched, holding on to the empirical evidence they have accumulated over the years. An empirical breakthrough is unlikely soon to favour one side over the other. With this deadlock in mind, Hammond (1996) sought to bridge the communication gap between FFH and BH. More specifically, he argues that all judgement theories fall into two categories, the coherence and the correspondence meta-theories of truth. While the former tradition started with Edwards (1954), Brunswik (1952) initiated the latter and Hammond (1955) adopted it

to decision theory. Hammond's central claim is that both BH and FFH are meta-theories of truth – they are not dichotomous programmes (one wrong, the other correct) but complementary. More specifically, FFH follows the evolutionary (and older) correspondence principle whereas BH follows the coherence principle that emerged with language, science and technology. Correspondence meta-theories seek empirical accuracy between subjective judgements and some indicators whereas coherence meta-theories are concerned with the coherence of subjective judgements with respect to logical, mathematical and statistical theory (Hammond 1996: 95). The accuracy of judgements (as appraised in correspondence meta-theories) refer to events in the real world in contrast to the validity of judgements (as appraised in coherence meta-theories) which do not refer to reality (Hastie and Rasinski 1988: 204). Thus, while coherence offers a criterion to decide on whether to accept some theory (the coherence of ideas with ideas), correspondence is not as much concerned with theories and their internal consistency, as with the accuracy of our subjective judgements with some facts about the world. It follows that successful coherent theories may not be refuted when they do not correspond with some facts about the world whereas correspondence theories sideline coherence to assess success in terms of correspondence of judgements with facts.

Hammond (1996: 203) remarks that each approach has its own standard to assess competence in judgement. However, coherence theorists generated a vast literature on cognitive illusions and the incompetence of human judgement. He argues that 'judgments based on coherence have an all or none character … coherence based systems demanded wholehearted acceptance or rejection; tinkering destroys coherence' (1996: 218). Competence in coherence theories start

> with principles and axioms, deduce mathematically coherence processes (equations) and correct answers (when specific numbers are inserted in the equations), then compare processes used by people to produce answers, as well as their answers, with those of the logically impeccable process chosen by the researcher. Human competence is thus compared with a mathematical "standard" (of which there may be more than one). Current researchers who study coherence competence have concluded that human judgment and decision-making are deeply flawed; coherence, or rationality, is seldom achieved. (1996: 217)

The findings of coherence researchers are of course not uncontested. They are refuted from within their camp by those who argue 'people are generally competent in achieving coherence judgments over different task conditions and that they achieve competence through the use of different – that is, multiple – cognitive strategies' (Hammond 1996: 217; cf. Payne

et al. 1993) and by correspondence theorists who, as we have seen, counter that heuristics succeed in achieving correspondence with facts (FFH).

Hammond (1996: 201) suggests that the two meta-theories should co-exist notwithstanding their opposing findings because their differences are complementary. He thinks that this tension between coherence and correspondence theories is here to stay and a compromise is not possible. But in lieu of an all-out war, they can peacefully and constructively coexist without the need to deny the achievements of the other.

Popper's (1972) three Worlds as described in his book *Objective Knowledge: An Evolutionary Approach* offer a framework to support this meta-theoretical argument. Popper defends a form of realism – the common sense theory of the world that affirms the existence of a physical world that does not depend on minds. The physical world contains things such as 'stones, trees, and humans'. Reality not being limited to physical things, it also contains the 'subjective decoding of our experiences' (Popper 1972: 37). Accordingly, there is the world of physical things (World 1) and the world of states of consciousness (World 2). Although World 1 exists independently of thoughts, World 1 and 2 are porous and interact. Furthermore, World 2 contains inborn dispositions to act which could be interpreted as heuristics.[5]

The third world is where the objective contents of thought reside. Popper compares World 3 to Plato's world of ideas except that it is 'man-made and changing [and unlike Plato's] it contains not only true theories but also false ones, and especially open problems, conjectures, and refutations' (Popper 1972: 121). While World 3 influences World 2, it is independent from physical things and states of consciousness. Scientific knowledge is the world of objective theories, objective problems and objective arguments all of which are in World 3. Even if humans produce the objective knowledge of World 3, once it is produced it becomes 'knowledge without a knower: it is knowledge without a knowing subject' (Popper 1972: 109). Popper insists that World 3 is autonomous 'despite the fact that it is our product and that it has a strong feedback effect upon us (as inmates of Worlds 1 and 2)' (1972: 112). Knowledge grows as World 1 and 2 interact with World 3. Humans constructed World 3 theories from the higher functions of language yet subjective beliefs have no implications on these theories. In other words, the psychological mental states of World 2 should not be confused with the objects – or the scientific theories – of World 3.

According to Hammond, BH and FFH entertain different relationships in Popper's three worlds: FFH seeks correspondence between World

[5] To illustrate how World 1 and World 2 interact, Berg (2014*b*) tests how successful businesses decide on location (World 1) using the satisficing heuristic (World 2) instead of computing marginal costs and benefits or probabilities (see also Klein 2009: 74).

1 and World 2 whereas BH seeks coherence between World 2 and World 3: 'The relationship between World 3 and 2 informs us about the coherence of a person's judgements process, but it does not directly inform us about its correspondence with World 1. One's judgement may be coherent but inaccurate, or the reverse' (Hammond 1996: 220). Correspondence entails empirical fit, or predictive power, between judgements and states of the world 'irrespective of whether the cognitive activity of the judge can be justified or even described' (Hammond 1996: 106) while coherence entails reasoning that satisfies some logical test in a closed system 'irrespective of whether the judgment is empirically accurate' (1996: 106). Accordingly, correspondence theorists focus on how the mind works in relation to the way the world works whereas coherence theorists compare how the mind works with how it ought to work.

Evolution, Hammond (1996: 221) argues, linked Worlds 1 and 2 while education and training offer access to the world of knowledge which is World 3. When we think of competence in judgement we need to ask, is it between World 1 and 2 or between World 3 and 1? Hammond argues that 'statistical principles are acquired from World 3, not from experience of World 1' (1996: 222) and we should not be surprised, as Tversky and Kahneman (1974) are, that individuals – who are mostly acquainted with World 1 – fail at solving puzzles in logic and probability. The BH programme then misinterprets the relation between World 1 and 3 as being the same as World 2 and 1. The puzzles of BH could only be answered by a trained person (and still statisticians failed to submit to the normative rules of logic and probability). Since World 1 and World 2 co-evolved, *Homo sapiens* have successfully used multiple fallible indicators to generate World 1 predictions, as corroborated by FFH's empirical findings.

6. DISCUSSION

The previous section introduced Hammond's attempt at bridging the gap between BH and FFH with his view that correspondence and coherence meta-theories are not reducible to one another. Their irreducibility engenders different research questions as each focuses on the flip side of the other programme (Hammond 1996: 224). This is corroborated by Rebonato's (2012: 50) remark that 'even when researchers belonging to the two schools look at the same phenomenon they seem to reach very different conclusions'. No winner can arise from this dispute and a constructive coexistence is necessary for both to work together in so far as their differences are complementary. As such, the long-standing tension between BH and FFH, whereby one seeks correspondence between World 1 and 2 whereas the other seeks coherence between World 2 and 3, should be used to better guide policy on the complexities of judgements.

We should extend this argument and ask whether FFH should be limited to establishing correspondence between World 1 and 2. Our view is that it should not. Consider how FFH relies on World 3 theories to explain the origin of heuristics and how they co-evolved with World 1. Heuristics have also been empirically identified and now pertain to the world of objective knowledge as a scientific programme (FFH). In World 3, however, FFH will remain context dependent and unable to achieve the level of generality attained by judgement theories that use World 3 benchmarks. Can BH in turn operate beyond achieving coherence between World 2 and World 3? Can the subjective experience of World 2 be trained to reason probabilistically and logically to achieve predictive success in World 1? If such learning is possible System 1 could be debiased. Accordingly, LP would no longer be needed as the biases are eliminated. But for Kahneman, Thaler or Sunstein such learning is unlikely – system 1 does not learn easily if at all. However, things are not so simple since even within the BH programme, as we shall see, there is no clear consensus on the matter. Second, even if subjective minds (are taught to) reason coherently, other factors that may not be easy to disentangle empirically, interfere including emotions, will power, fear, cognitive dissonance etc. Third, a coherent mind does not necessarily outperform heuristics in real-world environments. Since heuristics can be as good as, or sometimes better, than logical and probabilistic reasoning in achieving informational accuracy or economic returns, what is the rationale for investing resources in training minds to reason coherently?

Researchers in BH and FFH will agree that the normative principles of World 3 are cognitively demanding. And even if one could overcome the difficulty of teaching normative logical and probabilistic principles, biases (as optical illusions) are difficult to eliminate. Thaler, for instance, states that 'our ability to debias people is quite limited' while Kahneman considers 'it takes an enormous amount of practice to change your intuition' (both quoted in Bond 2009: 1191). Loewenstein and Haisley (2008) offer a pessimistic view ('in many cases this knowledge is insufficient to motivate behavior change', 2008: 212) and an optimistic view ('paternalistic policies do not preclude learning', 2008: 214). Wilson *et al*. (2002) are so pessimistic that their best recommendation is for individuals to avoid situations where they could be subject to biases. Fischhoff (1982) noted that even if debiasing is possible, the ensuing increase in people's faith about their judgements may lead to additional biases and may thus not be worth the trouble.

There are two concerns with the view that the biases are hard to dislodge. The first is that the optical illusion analogy is rejected by FFH. Unlike an optical illusion where a ruler can generally benchmark the deviation from the correct measure, in decision-making there may be no single normative model of choice from which deviations can be assessed.

Furthermore, the normative standard of the ruler as a corrective measure for human vision ignores how vision distorts to provide us with a better grasp of reality. This arises when the mind adds a third depth dimension to a two-dimensional picture of crossing 'parallel' rail road tracks (Berg and Gigerenzer 2010). The second point is that there is some evidence that learning World 3 reasoning theories is possible under certain conditions. Nisbett *et al*. (1987) found that while individuals were able to use statistical concepts such as the law of large numbers successfully with 'even brief formal training in inferential rules' (1987: 625), it was much more difficult for individuals to apply propositional logic. Nisbett *et al*.'s findings come close to FFH in their emphasis that learning will succeed if it taps mental rules individuals are acquainted with.

Given that learning to 'make statistical inferences turned out to be … difficult but not impossible' (Hammond 1996: 205), effective teaching remains more challenging in environments with uncertainty where feedback is irregular, information in flux and exogenous shocks unpredictable (we may speak of Keynesian–Knightian uncertainty). It is, of course, possible to create artificial closed systems where coherence can be constructed and tested but not without costs. Engineered systems can be isolated from outside disturbances generating closed loops where coherent reasoning works well if undisturbed. In such closed systems, where uncertainty is contained and quantified, it is possible to implement rule-based logical and Bayesian systems of reasoning. But these systems operate within an open system subject to exogenous shocks that cannot be controlled from within. Thus, even if the correct probabilistic or logical modes of reasoning are used, a sub-optimal outcome may still arise when the environment changes (World 1) or when a fast and frugal heuristic (World 2) exploits the environment more efficiently.

However, today's economies are changing at an unprecedented rate whereas the environment in which the heuristics evolved was stable for a few millions of years. Saint-Paul (2011: 17) remarks that 'automatic processes evolved millions of years ago to help us deal with situations that are quite different from those we encounter in the contemporary world … Inconsistency in behavior thus reflects the differences between the environment faced by our ancestors and ours'. Such beliefs have led BH advocates such as Stanovich (2004) to argue that heuristics work for replicators (the genes) not the interest of the vehicle (human well-being) which is promoted by rational thought or system 2. But for FFH, heuristics can empirically outperform normative World 3 modes of judgement even in a world characterized by incessant technological change (there would be no debate otherwise). To be a successful decision-maker, or a risk-savvy one (Gigerenzer 2015), in today's rapidly moving economies, one only needs to employ the right heuristic – there is no need to understand complex logical and probabilistic rules. Cognitive effort, accordingly,

is minimized because 'when intuition is smarter than logic, there is little need to educate people out of their "logical errors"' (Gigerenzer 2015: 370). But learning to reason coherently is also made easier 'with the help of adequate numerical and visual representations' (Gigerenzer 2015: 366). Individuals can employ the correct logical and probabilistic World 3 concepts if the problem is presented in the right format, usually frequencies and everyday terms as in the modified Wason task. Nisbett *et al.* (1982: 456) explain that 'good statistical heuristics should be learnable to the point where they can be used easily, even automatically'. They also argue that these heuristics will be increasingly adopted given that human reasoning can change quite rapidly with language, education and culture.

Outside the lab, Keynesian–Knightian uncertainty prevails and heuristics are more adapted and adaptable. However, there are contexts where it is better to let a formula decide for us, as Klein (2009) argues, including 'well-ordered domains that have enough structure to get the system organized, enough stability to make the programming economical, not too much complexity or interdependencies, little need for tacit knowledge, and clear criteria for good decisions' (2009: 118). Yet, Klein adds, statistical methods and algorithms are unreliable when conditions are likely to change. They also inhibit learning and render people passive followers of rules. He concludes that outside the lab, decision-support systems are not helpful and most judgements should be made by humans, especially in context dependent, unstable, shadowy conditions.[6]

Can a risk-savvy person in the 21st century – as a citizen, worker or consumer – solely rely on heuristics to act rationally or at least not to be fooled? Hammond (1996) claims that while correspondence competence has been with us for a longer stretch of time, coherence competence 'began to play a larger and larger role in human judgment; it has begun to dominate Western civilization and has made large inroads in the rest of the world. There seems to be no end in sight' (1996: 352). But this is not without danger, he warns, pointing out that this change is behind the largest human tragedies of the 20th century. Still, we can ask, if heuristics are as efficient as the FFH programme suggests they are, could they substitute the morally controversial soft paternalism of nudges? Since heuristics are morally less subject to critique, they should be the first policy response when there is evidence that they work. However, heuristics on their own only offer a partial view of judgement. In an open world where the pace of technical change and environmental

[6] Kahneman and Klein (2009) argue that expert intuitions are reliable when there are adequate learning opportunities and in high validity environments, i.e. when there are reliable learning cues. Algorithms are more consistent than human decision-makers and will outperform expert human judgements when validity is low or when validity is very high.

transformations are unlikely to slow down, the borderline between the adequacy of World 2 and World 3 modes of judgements cannot be fixed. Policy problems should be assessed with both perspectives in mind as Hammond advocates.

The informed policy-maker needs to know the difference between heuristics that work well in certain environments and the normative decision rules from World 3. What mental tool is more effective is contingent on the constitution of our minds and the ability to learn, the target indicator, and the environment in which the decisions are made. Policy-makers must also not neglect the importance of the time frame in which the policy will be implemented. The time frame is especially significant since heuristics could potentially be rapidly modified whereas learning coherent theories of judgement is more time and resource intensive.

7. CONCLUSION

We compared the descriptive, normative and policy implications of two programmes, the biases and heuristics and the fast and frugal heuristics. The adoption of Hammond's meta-theoretical framework entails that the two programmes are complementary and are both needed to complete our understanding of judgement and decision-making. We argued that, if the protagonists have argued past each other, it is partly because they have different ontological frameworks, namely, one based on coherence and the other on correspondence. Their boundaries are fluid and future works will have to explore where they overlap and where they are effectively offering competing descriptive, normative and policy prescriptions. In an economy driven by technological change, where work and consumption require ever more smart decision skills, the friction between coherence and correspondence meta-theories will arise more frequently.

It is pointless to deny, in conclusion, that BH has identified heuristics that in certain contexts harm decision-makers as much as there is no point in denying that, in other contexts, heuristics are quite effective mental tools. There are even cases where both programmes offer similar policy prescriptions, as is the case of some defaults – contrast Gigerenzer (2007: 182) to Thaler and Sunstein (2008: 185–7). Our view therefore is that there are synergies in working together and combining both programmes into one whose aim is to explore when heuristics work and when they fail us. This is necessary to re-focus scientific work on the effectiveness of heuristics in decision-making rather than on a single type of question. Currently, the BH programme only acknowledges that heuristics may not always bias judgements (without providing evidence) whereas the FFH programme does the opposite, working from the other end and focusing almost exclusively on confirming how heuristics are effective. Our insight

is for advocates of these complementary programmes to work together in exploring all aspects of heuristics including contentious theoretical disagreements on the effectiveness of expert judgements and intuitions (see joint work by Kahneman and Klein 2009) or the effects of frequencies on the conjunction fallacy (also known as the Linda problem; see Mellers *et al*. 2001). Furthermore, while Samuels *et al*. (2002) eliminate various disagreements between FFH and BH, they conclude that the interpretation of probability – whether it should be understood in terms of frequencies as evolutionary psychologists argue or subjectively as Bayesians suggest – is a real contentious issue that resists their attempted reconciliation.

Finally, Hammond's original contribution leaves out ethical deliberation which is necessary if paternalistic policies are to be justified. Paternalists assume and have argued that cost–benefit analysis is the way forward to assess when an intervention (that increases welfare and reduces freedom) should be implemented. But what if these trade-offs are deemed morally unacceptable? Without a moral framework to justify what trade-offs are allowed the discussion hits a dead end. We agree with Saint-Paul (2011: 153) that 'it is impossible to provide a purely logical system of foundations for the organization of society independent of an opinion about human nature'. Our understanding of heuristics, while enhanced with Hammond's meta-theoretical categories, will be insufficient for policy guidance if not complemented with ethical deliberation.

ACKNOWLEDGEMENTS

We would like to express our thanks and gratitude to Dr Erik Angner with whom we collaborated in the workshop 'Welfare economics after behavioral economics' organized by the Department of Economics at the American University of Beirut on 13 April 2016. Additionally, we would like to thank the workshop participants for their useful comments. We also acknowledge the anonymous referees for their valuable input.

REFERENCES

Benartzi, S. and R. Thaler. 2001. Naıve diversification strategies in defined contribution saving plans. *American Economic Review* 91: 79–98.

Berg, N. 2003. Normative behavioral economics. *Journal of Socio-Economics* 32: 411–427.

Berg, N. 2014a. The consistency and ecological rationality approaches to normative bounded rationality. *Journal of Economic Methodology* 21: 375–395.

Berg, N. 2014b. Success from satisficing and imitation: entrepreneurs' location choice and implications of heuristics for local economic development. *Journal of Business Research* 67: 1700–1709.

Berg, N. and G. Gigerenzer. 2007. Psychology implies paternalism? Bounded rationality may reduce the rationale to regulate risk taking. *Social Choice and Welfare* 28: 337–359.

Berg, N. and G. Gigerenzer. 2010. As-if behavioral economics: neo-classical economics in disguise? *History of Economic Ideas* 18: 133–165.

Berg, N., G. Biele and G. Gigerenzer. 2016. Consistent Bayesians are no more accurate than non-Bayesians: economists surveyed about PSA. *Review of Behavioral Economics* 3: 189–219.

Berlin, I. 1969. Two concepts of liberty. In *Four Essays on Liberty*, 118–172. Oxford: Oxford University Press.

Bond, M. 2009. Decision-making: risk school. *Nature* 461: 1189–1192.

Brunswik, E. 1952. The conceptual framework of psychology. *International Encyclopedia of Unified Science* 1: 4–102.

Buchanan, J. 1978. Natural and artifactual man. In *The Collected Works of James Buchanan, Volume I: The Logical Foundations of Constitutional Liberty* [1999]. Indianapolis, IN: Liberty Fund.

Camerer, C. and G. Loewenstein. 2004. Behavioral economics: past, present, future. In *Advances in Behavioral Economics*, ed. C. Camerer, G. Loewenstein and M. Rabin. Princeton, NJ: Princeton University Press.

Camerer, C., S. Issacharoff, G. Loewenstein, T. O'Donoghue and M. Rabin. 2003. Regulations for conservatives: behavioral economics and the case for 'asymmetric' paternalism. *University of Pennsylvania Law Review* 151: 1211.

Camerer, C., G. Loewenstein and M. Rabin. 2004. *Advances in Behavioral Economics*. Princeton, NJ: Princeton University Press.

Conlisk, J. 1996. Why bounded rationality? *Journal of Economic Literature* 34: 669–700.

Conly, S. 2014. *Against Autonomy: Justifying Coercive Paternalism*. New York, NY: Cambridge University Press.

Cosmides, L. and J. Tooby. 1996. Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition* 58: 1–73.

Demiguel, V., L. Garlappi and R. Uppal. 2009. Optimal versus naïve diversification: how inefficient is the 1/N portfolio strategy? *Review of Financial Studies* 22: 1915–1953.

Dworkin, G. 1983. Paternalism. In *Paternalism*, ed. R. Sartorius, 19–34. Minneapolis, MN: University of Minnesota Press.

Dworkin, G. 2016. Paternalism. In *The Stanford Encyclopedia of Philosophy*, ed. E. N. Zalta. https://plato.stanford.edu/archives/win2016/entries/paternalism/.

Edwards, W. 1954. The theory of decision making. *Psychological Bulletin* 51: 380–417.

Evans, J. 2003. In two minds: dual-process accounts of reasoning. *Trends in Cognitive Science* 7: 454–459.

Fischhoff, B. 1982. Debiasing. In *Judgment Under Uncertainty: Heuristics and Biases*, ed. D. Kahneman, P. Slovic and A. Tversky, 422–444. New York, NY: Cambridge University Press.

French, K. and J. Poterba. 1991. Investor diversification and international equity markets. *American Economic Review (Papers & Proceedings)* 81: 222–226.

Gigerenzer, G. 2007. *Gut Feelings: The Intelligence of the Adaptive Unconscious*. London: Penguin Books.

Gigerenzer, G. 2008. Fast and frugal heuristics. In *Rationality for Mortals: How People Cope with Uncertainty*, 20–45. New York, NY: Oxford University Press.

Gigerenzer, G. 2014. *Risk Savvy. How to Make Good Decisions*. New York, NY: Penguin Classics.

Gigerenzer, G. 2015. On the supposed evidence for libertarian paternalism. *Review of Philosophy and Psychology* 6: 361–383.

Gigerenzer, G. and H. Brighton. 2009. Homo heuristicus: why biased minds make better inferences. *Topics in Cognitive Science* 1: 107–143.

Gigerenzer, G. and R. Selten (ed.). 2001. Rethinking rationality. *In Bounded Rationality: The Adaptive Toolbox*, 1–12. Cambridge, MA: MIT Press.

Gigerenzer, G. and P. Todd. 1999. Fast and frugal heuristics: The adaptive toolbox. In *Simple Heuristics that Make us Smart: Evolution and Cognition*, ed. G. Gigerenzer, P. M. Todd and ABC Research Group, 3–36. New York, NY: Oxford University Press.

Gigerenzer, G., G. Czerlinksi and L. Martignon. 2002. How good are fast and frugal heuristics? In *Heuristics and Biases: The Psychology of Intuitive Judgment*, ed. T. Gilovich, D. Griffin and D. Kahneman, 559–581. Cambridge: Cambridge University Press.

Gilboa, T., A. Postlewaite and D. Schmeidler. 2009. Is it always rational to satisfy Savage's axioms? *Economics and Philosophy* 25: 285–296.

Gilovich, T. and D. Griffin. 2002. Introduction – heuristics and biases: then and now. In *Heuristics and Biases: The Psychology of Intuitive Judgment*, ed. T. Gilovich, D. Griffin and D. Kahneman, 1–18. New York, NY: Cambridge University Press.

Gilovich, T., D. Griffin and D. Kahneman. 2002. *Heuristics and Biases. The Psychology of Intuitive Judgments*. New York, NY: Cambridge University Press.

Glaeser, E. 2006. Paternalism and psychology. *University of Chicago Law Review* 73: 133–155.

Goldstein, D. and G. Gigerenzer. 1999. The recognition heuristics: how ignorance makes us smart. In *Simple Heuristics That Make us Smart. Evolution and Cognition*, ed. G. Gigerenzer, P. M. Todd and ABC Research Group, 37–58. New York, NY: Oxford University Press.

Goodwin, T. 2012. Why we should reject nudge. *Politics* 32: 85–92.

Grüne-Yanoff, T. 2012. Old wine in new casks: libertarian paternalism still violates liberal principles. *Social Choice and Welfare* 38: 635–645.

Halpern, D. 2015. *Inside the Nudge Unit. How Small Changes can Make a Big Difference*. London: WH Allen.

Hammond, K. 1955. Probabilistic functioning and the clinical method. *Psychological Review* 62: 255–262.

Hammond, K. 1996. *Human Judgments and Social Policy*. New York, NY: Oxford University Press.

Hastie, R. and T. Rasinski. 1988. The concept of accuracy in social judgment. In *The Social Psychology of Knowledge*, ed. D. Bar-Tal and A. Kruglanski. Cambridge: Cambridge University Press and Editions de la Maison des Sciences de l'Homme.

Hausman, D. and B. Welch. 2010. Debate: To nudge or not to nudge. *Journal of Political Philosophy* 18: 123–136.

Heukelom, F. 2014. *Behavioral Economics. A History*. Cambridge: Cambridge University Press.

Hodgson, G. 1993. *Economics and Evolution: Bringing Life Back into Economics*. Ann Arbor, MI: Michigan University Press.

Hoffrage, U. and G. Gigerenzer. 2004. How to improve the diagnostics inferences of medical experts. In *Experts in Science and Society*, ed. E. Kurze-Milke and G. Gigerenzer, 249–268. Dordrecht: Kluwer.

Kahneman, D. 2011. *Thinking, Fast and Slow*. New York, NY: Macmillan.

Kahneman, D. and G. Klein. 2009. Conditions for intuitive expertise: a failure to disagree. *American Psychologist* 64: 515–526.

Kahneman, D. and A. Tversky. 1984. Choices, values, frames. *American Psychologist*, 34: 341–350.

Kahneman, D. and A. Tversky. 1996. On the reality of cognitive illusions. *Psychological Review* 103: 582–591.

Kahneman, D., P. Slovic and A. Tversky. 1982. *Judgment Under Uncertainty: Heuristics and Biases*. New York, NY: Cambridge University Press.

Klein, D. 2004. Statist quo bias. *Economic Journal Watch* 1: 260–271.

Klein, G. 2009. *Streetlights and Shadows. Searching for the Keys to Adaptive Decision Making*. Cambridge, MA: MIT Press.

Kosters, M. and J. Van der Heijden. 2015. From mechanism to virtue: evaluating nudge theory. *Evaluation* 21: 276–291.

Lopes, L. 1991. The rhetoric of irrationality. *Theory and Psychology* 1: 65–82.

Loewenstein, G. and E. Haisley. 2008. The economist as a therapist: methodological ramifications of "light paternalism". In *The Foundations of Positive and Normative*

Economics: A Hand Book, ed. A. Caplin and A. Schotter, 210–248. New York, NY: Oxford University Press.
Loewenstein, G. and T. O'Donoghue. 2006. We can do things the easy way or the hard way: negative emotions, self-regulation, and the law. *University of Chicago Law Review* 73: 183–206.
McQuillin, B. and S. Sugden. 2012. Reconciling normative and behavioral economics: the problems to be solved. *Social Choice and Welfare* 38: 553–567.
Mellers, B., R. Hertwig and D. Kahneman. 2001. Do frequency representations eliminate the conjunction effects? *Psychological Science* 12: 269–275.
Nisbett, R., D. Krantz, C. Jepson and G. Fong. 1982. Improving inductive inference. In *Judgment Under Uncertainty: Heuristics and Biases*, ed. D. Kahneman, P. Slovic and A. Tversky, 445–462. New York, NY: Cambridge University Press.
Nisbett, R., G. Fong, D. Lehman and P. Chen. 1987. Teaching reasoning. *Science* 238: 625–631.
OECD. 2017. *Use of Behavioural Insights in Consumer Policy*. OECD Science, Technology and Innovation Policy Papers No. 36, January 2017. Paris: OECD.
Ortmann, A., G. Gigerenzer, B. Borges and D. Goldstein. 2008. The recognition heuristic: a fast and frugal way to investment choice? In *Handbook of Experimental Economics Results*, ed. C. Plott and V. Smith, 993–1003. Amsterdam: Elsevier/North Holland.
Payne, J., J. Bettman and E. Johnson. 1993. *The Adaptive Decision Maker*. New York, NY: Cambridge University Press.
Peterson, C. and L. Beach. 1967. Man as an intuitive statistician. *Psychological Bulletin* 68: 29–46.
Popper, K. 1972. *Objective Knowledge: An Evolutionary Approach*. Oxford: Clarendon Press.
Rabin, M. 1998. Psychology and economics. *Journal of Economic Literature* 36: 11–46.
Radner, R. 1995. *Economic Survival*. Nancy L. Schwartz Lecture. Lecture delivered May 1995 at the J. L. Kellogg School of Management, Northwestern University.
Rebonato, R. 2012. *Taking Liberties: A Critical Examination of Libertarian Paternalism*. Basingstoke: Palgrave Macmillan.
Samuels, R., S. Stich and M. Bishop. 2002. Ending the rationality wars: how to make the disputes about human rationality disappear. In *Common Sense, Reasoning and Rationality*, ed. R. Elio, 236–268. New York, NY: Oxford University Press.
Saint-Paul, G. 2011. *The Tyranny of Utility: Behavioral Social Sciences and the Rise of Paternalism*. Princeton, NJ: Princeton University Press.
Sent, E. 2004. Behavioral economics: how psychology made Its (limited) way back into economics. *History of Political Economy* 36: 735–760.
Simon, H. 1955. A behavioral model of rational economics man. *Quarterly Journal of Economics* 69: 99–118.
Smith, A. 1776. *An Inquiry into the Nature and Causes of the Wealth of Nations*. New York, NY: The Modern Library.
Smith, V. 2008. *Rationality in Economics. Constructivist and Ecological Forms*. New York, NY: Cambridge University Press.
Stanovich, K. 2004. *The Robot's Rebellion. Finding Meaning in the Age of Darwin*. Chicago, IL: Chicago University Press.
Sugden, R. 2008. Why incoherent preferences do not justify paternalism. *Constitutional Political Economy* 19: 226–48.
Sunstein, C. 1996. Incommensurability and valuation in law. *Michigan Law Review* 92: 779–861.
Sunstein, C. 2005. Moral heuristics. *Behavioral and Brain Sciences* 28: 531–573.
Sunstein, C. 2013a. The Storrs Lectures: Behavioral Economics and paternalism. *Yale Law Journal* 122: 1670–2105.
Sunstein, C. 2013b. *Simpler: The Future of Government*. New York, NY: Simon & Schuster.

Sunstein, C. 2014. *Why Nudge? The Politics of Libertarian Paternalism*. New Haven, CT: Yale University Press.
Sunstein, C. and R. Thaler. 2003. Libertarian paternalism is not an oxymoron. *University of Chicago Law Review* 70: 1159–1202.
Thaler, R. 1980. Towards a positive theory of consumer choice. *Journal of Economic Behavior and Organization* 1: 39–60.
Thaler, R. 1991. *Quasi-Rational Economics*. New York, NY: Russell Sage Foundation.
Thaler, R. 2002. *The Winner's Curse. Paradoxes and Anomalies of Economic Life*. Princeton, NJ: Princeton University Press.
Thaler, R. and C. R. Sunstein. 2008. *Nudge: Improving Decisions about Health, Wealth, and Happiness*. New Haven, CT: Yale University Press.
Todd, P. and G. Gigerenzer. 2007. Environments that make us smart: ecological rationality. *Current Directions in Psychological Science* 16: 167–171.
Tversky, A. 1972. Elimination by aspects: a theory of choice. *Psychological Review* 79: 281–299.
Tversky, A. and D. Kahneman. 1974. Judgments under uncertainty: heuristics and biases. *Science* 185: 1124–1131.
Tversky, A. and D. Kahneman. 1983. Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychological Review* 91: 292–315.
White, M. 2013. *The Manipulation of Choice: Ethics and Libertarian Paternalism*. New York, NY: Palgrave Macmillan.
Whitman, D. and M. Rizzo. 2015. The problematic welfare standard of behavioral paternalism. *Review of Psychology and Philosophy* 6: 409–425.
Wilson, T., D. Centerbar and N. Brekke. 2002. Mental contamination and the debiasing problem. In *Heuristics and Biases. The Psychology of Intuitive Judgment*, ed. T. Gilovich, D. Griffin and D. Kahneman, 185–200. New York, NY: Cambridge University Press.
World Bank. 2015. *World Development Report 2015: Mind, Society, and Behavior*. Washington, DC: World Bank. doi: 10.1596/978-1-4648-0342-0.

BIOGRAPHICAL INFORMATION

Ramzi Mabsout is Assistant Professor of Economics at the American University of Beirut. He served as the Chair of the Department of Economics for the academic year 2015–2016. He is the author of several publications among which 'Reconciling economics with naturalist ethical theory' (*Review of Social Economy*, 2016), 'Abduction and economics: The contributions of Charles Peirce and Herbert Simon' (*Journal of Economic Methodology*, 2015), and 'Mindful capability' (*Ecological Economics*, 2014).

Jana G. Mourad is currently a researcher at the Lebanese Center for Policy Studies (LCPS). Her forthcoming work spans a variety of economic subjects including refugees, development, voting behaviour, and poverty as a joint collaboration between LCPS and several UN agencies and international organizations. She previously conducted research at the American University of Beirut and contributed to a number of local and global reports including *Vulnerability Assessment of Syrian Refugees* (United Nations High Commissioner of Refugees, 2015) and *Rapid Poverty Assessment in Lebanon* (United Nations Development Programme, 2016).