Matteo M. Galizzi

Behavioral aspects of policy formulation: experiments, behavioral insights, nudges

Article (Accepted version) (Unrefereed)

Original citation:
Galizzi, Matteo M. (2017) *Behavioral aspects of policy formulation: experiments, behavioral insights, nudges.* In: Howlett, Michael, Mukherjee, Ishani and Fraser, Simon, (eds.) Handbook of Policy Formulation. Handbooks of Research on Public Policy. Edward Elgar Publishing, Cheltenham, UK. ISBN 9781784719319

Behavioral aspects of policy formulation: experiments, behavioral insights, nudges.

Matteo M Galizzi

London School of Economics of Political Science
Department of Social Policy
LSE Health and Social Care
LSE Behavioural Science
Centre for the Study of Incentives in Health
OLD 2.35 Old Building
Houghton Street
WC2A 2AE London
Email: m.m.galizzi@lse.ac.uk

Abstract. The chapter reviews and critically discusses several 'behavioral' aspects of policy formulation. First, it makes a conceptual distinction between 'behavioral' aspects of policy formulation related either to methods or to insights. Second, it discusses several methodological aspects of proposing experiments for policy formulation purposes. In particular, it reviews the case of conducting randomized controlled experiments in the broader context of evidence-based policy tools; refers to an influential taxonomy of experiments spanning the continuum from the lab to the field; and discusses the scope, limitations, and current directions of each type of experiments within such a flexible toolkit. Third, it provides an operational definition of behavioral economics and behavioral science, and proposes a taxonomy of policy formulation insights based on the extent to which these insights are genuinely inspired by behavioral science. Finally, within the policy formulation insights directly inspired by behavioral science, it focuses on the nudging approach to illustrate its potential and future challenges from a conceptual and practical perspective.

1. Introduction

In the last few years, 'behavioral economics', and, more generally, 'applied behavioral science' (as per Kahneman, 2012) have gained outstanding momentum among policy-makers. Several governments in developed countries have constituted 'behavioral insights teams' within their civil services, including the Nudge Unit in the UK Cabinet Office and the Office for Information and Regulatory Affairs (now the Social and Behavioral Sciences Team) in the US; analogous initiatives have also been set up within the governments of Australia, Canada, Denmark, Finland, France, Israel, the Netherlands, New Zealand, Norway and Singapore (Dolan et al., 2012; Sunstein, 2011; Dolan & Galizzi, 2014a; Annala et al., 2015). Insights from the behavioral sciences have also attracted increasing attention by international institutions: the European Commission has set up a Foresight and Behavioral Insights Unit and a Behavioral Economics Team at the Institute for Health and Consumer Protection, both at the Joint Research Centre of the European Commission; the World Bank's 2015 World Development Report—titled *Mind, Society and Behavior*—addresses the psychological, social, and cultural influences on decision making and human behaviour and the impact these have on development; and the OECD has promoted a series of high-level workshops on the applications of behavioral sciences to policy making (http://www.oecd.org/gov/behavioural-economics.htm).

Two recent books have systematically discussed the applications of behavioral insights to public policy (Shafir, 2012; Oliver, 2013), and a number of articles have explored specific areas of policy applications, including savings and pensions (Thaler & Benartzi, 2004; Beshears et al., 2011), welfare (Bernheim & Rangel, 2009; Costa-Font, 2011), and health (Loewenstein et al., 2007; Volpp et al., 2011; Loewenstein et al., 2012; Galizzi, 2014).

Policy-makers have indeed put forward an array of policies often referred to as 'behavioral policies', encompassing randomized controlled trials, financial incentives, comparison web portals, and nudges, among others. The interventions that fall under the behavioral umbrella in policy applications are, however, quite heterogeneous and diverse. It is often unclear to which aspect of a policy formulation the term 'behavioral' refers to, and practitioners, policy-makers and researchers often use or assume quite different definitions of 'behavioral'.

There are two particular, and related, sources of potential misunderstanding in the behavioral aspects of policy formulation. First, it is often not clear whether the 'behavioral' attribute refers to the methods of research or the insights obtained from the research (or to both)—is an intervention based on evidence from a randomized controlled trial automatically a behavioral policy?

Second, how closely is the conceptual core of these disparate 'behavioral' policies related to genuine insights from behavioral science? In other words, what 'behavioral' insights really are behavioral?

This çhapter attempts to conceptually dissect the two sources of misunderstanding in the behavioral aspects of policy formulation by building on, and generalizing from, the discussion in Galizzi (2014) on behavioral policies in the domain of health.

The chapter first provides an operational definition of behavioral policy formulation and of 'behavioral' economics as opposed to 'conventional' economics. To address the two sources of possible misunderstanding, a distinction is immediately made between insights and methods.

On the methods, the current emphasis on randomized controlled trials (RCTs) is related to the broader discussion on the various types of randomized controlled experiments (RCEs) in economics, political sciences, and social sciences in general.

On the insights, the focus is on the nature and content (rather than the methodology) of behavioral policy formulation. The chapter proposes a taxonomy consisting of five different clusters of behavioral policy formulation instruments: preferences-based policies; information-based policies; financial incentives; regulation-based policies, including tax- and subsidy-based policies; and 'nudges'.

The discussion then focuses on to which extent these five clusters of supposedly behavioral policy formulation instruments depart from the conventional economics view of individual behavior and decision-making. It turns out that some of the policy instruments that, in the public debate, are typically considered to be behavioral in fact have limited behavioral content, and are instead quite well-established tools in the conventional economics toolbox.

The rest of the chapter is structured as follows. Section 2 defines behavioral policy formulation, behavioral policy, and behavioral economics. Section 3 deals with the question of ‘behavioral’ methods and insights. Section 4 defines: (4.1) preferences-based policies; (4.2) information-based policies; (4.3) financial incentives; (4.4) regulation-based policies, including tax- and subsidy-based policies; and (4.5) nudges. Finally, a discussion in Section 5 briefly concludes.

2. A definition of ‘behavioral’ policy

The behavioral aspects of policy formulation are closely interlinked with what the literature calls more broadly ‘behavioral policy’, or, equivalently, ‘behavioral public policy’ (Shafir, 2012; Oliver, 2013). Behavioral policy is thus usually defined as a policy intervention that is directly inspired by, and designed on, the principles of behavioral research.

But there is no one precise disciplinary ‘label’ to attach to ‘behavioral’ research. Behavioral researchers are essentially social and cognitive psychologists, as well as a “growing minority of economists - behavioral economists” (Kahneman, 2012, p. ix). Daniel Kahneman proposes ‘applied behavioral science’ as a common label for the shared activities, methods and interests by psychologists and behavioral economists. A growing number of leading institutions, in both the academic and the policy arenas, have now adopted this comprehensive definition of ‘behavioral science’ as an interdisciplinary area of research bringing together insights and methods from social and cognitive psychology, behavioral and experimental economics, neuroscience, philosophy, marketing and consumer behavior, organizational behavior, sociology, political science, anthropology, biology, medical and health sciences, happiness, and well-being research, among others.

In the popular press, as well as among most policy practitioners, however, there is a tendency to reduce behavioral science to one of its sub-components, behavioral economics. There are several possible reasons for this, including the fact that, traditionally, economics has had a stronger influence and traction on policy-making and practice than the other social sciences. Another reason may be that, especially in recent decades, conventional economics as a discipline has proposed itself as a comprehensive, structured theory that could be applied to virtually any social phenomena. Psychology, in contrast, is typically represented as a set of ad-hoc theories applicable to specific issues and phenomena. A further reason may be related to the ‘imperialistic’ tendency of conventional economics research to expand into the domains of other disciplines such as political science, history and sociology (Baron & Hannan, 1994; Lazear, 2000).

Whatever the reason for the tendency to reduce behavioral science to its behavioral economics component, this synecdoche then requires us to define behavioral economics. A definition is provided by the Russell Sage Foundation’s influential Round Table for Behavioral Economics, established in 1992 to “devise activities designed to advance this new interdisciplinary field [of behavioral economics]”.

The Round Table defines behavioral economics as follows: “Behavioral economics uses facts, models, and methods from neighboring sciences to establish descriptively accurate findings about human cognitive ability and social interaction and to explore the implications of these findings for economic behavior. The most fertile neighboring science in recent decades has been psychology, but sociology, anthropology, biology, and other fields can usefully influence economics as well” (Russell Sage Foundation, 2016).

In essence, behavioral economics is thus defined as the application to 'conventional' economics of insights from cognitive and social psychology, as well as of cognate disciplines like biology, anthropology and sociology, to improve the understanding of economic behavior and decision-making.

This definition implicitly defines behavioral economics as departing from, and somehow challenging, the traditional or conventional view of economics, which is essentially based on the assumption of perfectly "rational individuals who engage in maximizing behavior" (Lazear, 2000). In its most stylized and popularized form, the conventional economics view relies on four main conceptual 'pillars':

1. Preferences: We have a complete and comprehensive set of preferences—spanning over all possible factors affecting our utility and well-being—and a clear, conscious, and consistent representation of those preferences: our preferences are thus stable both across domains/situations, and over time;
2. Information: Preferences drive our behavior and decision-making: when we decide, we process all available information, we rationally 'optimize' by calculating the costs and benefits of different choices or courses of actions, and deliberately pick the one that most closely matches our preferences;
3. Incentives: Our rational decisions and behavior best serve our own interests and maximize our own utility when interacting with others in markets: in equilibrium, markets aggregate individual costs/benefits values and translate them into prices;
4. Regulation: Since we always rationally act in our own best interests, public intervention is needed only when markets fail to correctly translate some costs/benefits values into prices; this typically occurs in the case of market failures such as 'externalities' where individual values do not incorporate effects on others' costs or benefits.

Conventional economics approaches essentially stick to these four conceptual pillars, while behavioral economics approaches relax some of these assumptions in light of evidence suggesting that, for instance, we may not always act on our own best interests, or we may only *try* to optimize.

3. Does 'behavioral' refer to methods or insights? Randomized controlled experiments and 'behavioral' policy.

The first source of potential misunderstanding in the behavioral aspects of policy formulation relates to the fact that practitioners and policy-makers tend to define a policy instrument under the behavioral umbrella merely because it entails the use of randomized controlled trials (RCTs) as a method of generating evidence.

Someone who attended a conference organized by the European Commission's Directorate General for Health and Consumer Affairs, for instance, could come away with the impression that the key feature of the EC's various behavioral insights teams is that they pre-test possible policy interventions using controlled experiments involving a treatment and a control group (European Commission DG SANCO, 2013). Probably the most influential report by the UK's Behavioural Insights Team when it was still within the Cabinet Office illustrates the need to conduct randomized controlled trials to develop public policy (Haynes et al., 2012).

This emphasis on the use of RCTs as a fundamental defining criterion for ‘behavioral’ policies calls for three conceptual clarifications.

First, the use of RCTs has to do with the *methods* employed to gather evidence for policy purposes, not with the content and insights of such evidence. RCTs are certainly not a distinguishing feature only of behavioral policy, behavioral science or behavioral economics. Outside the context of policy decision-making the use of RCTs is far from novel. All modern evidence-based science, medicine, and pharmacology are based on RCTs, starting from the pioneering work on scurvy by James Lind in 1747 to the first published RCT in medicine by Austin Bradford Hill and colleagues in 1948. Thanks to the ground-breaking methodological contributions of Charles Sanders Peirce, Jerzy Neyman, Ronald A. Fisher and others, modern science has since long considered the experimental method as ‘the’ scientific method. Even in the policy decision-making context, the idea of using versions of the RCTs for policy applications has been advocated from several decades (Rubin, 1974; 1980a; 1980b; 1986; Ferber & Hirsch, 1978; 1982; Hausman & Wise, 1985; Heckman, 1992; Burtless, 1995).

What is relatively novel in the policy formulation arena is that there is currently, probably for the first time ever, a diffuse and open-minded interest by decision-makers and practitioners in rolling out rigorous tests of envisaged policy interventions prior to their full-scale implementation (Ludwig et al., 2011; Dolan & Galizzi, 2014a).

Second, while the term RCT is now widely en vogue in policy circles, it is often used in a quite particular way which deserves a further set of methodological clarifications. To start with, in the current policy debate, the term RCT is explicitly or implicitly used to typically denote large-scale experiments conducted with entire organizations (e.g. schools, hospitals, villages) without necessarily involving the stakeholders in those organizations to explicitly express their views or their consent in the envisaged experiments. This is a major conceptual difference with respect to RCTs in medicine or pharmacology, where subjects are always explicitly asked to give informed consent prior to take part into RCTs, with obvious but profound ethical and political implications. The term ‘RCT’ is therefore conceptually inappropriate and practically misleading in a policy formulation context, as it conveys the impression that subjects have been made aware of being part to a policy experiment and have been consulted and given their consent to it, when actually this may not be the case in many applications of experiments conducted for policy purposes (including those run by the various ‘behavioral insights teams’).

Furthermore, in the current policy debate, RCTs (as intended in the above sense) are often improperly contrasted with other empirical methods used to gather evidence to inform policy-making.

It is true that only well-designed and well-run randomized controlled experiments provide an unbiased estimate of the average treatment effect (that is, in empiricists’ jargon, are *internally valid*). Because of the well-known issue of sample selection bias, in fact, it is only by randomly assigning subjects to a treatment or to a control group that one can identify the causal effect of a policy intervention on an observed outcome (Heckman, 1979; Burtless, 1995; Angrist & Pischke, 2009; List, 2011; Gerber & Green, 2012). This is why the alternative policy evaluation methods that do not use randomization (Ashenfelter, 1978; Lalonde, 1986) then need to overcome this ‘original sin’ in their design by exploiting ‘naturally occurring’ experiments (Ashenfelter & Krueger, 1994; Rosenzweig & Wolpin, 2000), or resorting to identification strategies such as *instrumental variables* (e.g. Angrist &

Krueger, 1991, 2001; Angrist & Imbens, 1995; Angrist et al., 1996); *propensity score* and other *matching methods* (Rubin, 1973; Rosenbaum & Rubin, 1983, 1984; Heckman et al., 1998; Imbens, 2004); *difference in differences* (e.g. Card, 1992, 1996; Card & Krueger, 1994, 2000); or *regression-discontinuity designs* (e.g. Trochim, 1984; Angrist & Lavy, 1999; Hahn et al., 2001; Cook, 2008; Imbens & Lemieux, 2008).

On the other hand, RCTs (as intended in the above sense) are only one specific and very peculiar type of experiment. It is useful to refer here to the influential taxonomy of experiments in economics, and more broadly in social sciences, originally crystalized by Harrison and List (2004): *conventional lab* experiments involve student subjects, abstract framing, a lab context, and a set of imposed rules; *artefactual field* experiments depart from conventional lab experiments in that they involve non-student samples; *framed field* experiments add to artefactual field experiments a field context in the commodity, stakes, task or information; and, finally, *natural field* experiments depart from framed field experiments in that subjects undertake the tasks in their natural environment, and subjects do not know that they take part into an experiment. The main idea beyond *natural field* experiments is von Heisenberg's 'uncertainty principle' in physics: the mere act of observation and measurement necessarily alters what is being observed and measured. In key areas for policy formulation, such as health, the environment or ethical and pro-social behavior, for instance, there are potential *randomization biases* (i.e. merely knowing that random assignment is in place causes the type of persons participating in a randomized study to differ from participants in other studies: Kramer & Shapiro, 1984; Heckman & Smith, 1995); *experimenter demand effects* (i.e. participants change behavior due to cues about what represents 'appropriate' behavior for the experimenter: Bardsley, 2005; Levitt & List, 2007a; 2007b; Zizzo, 2010); *Hawthorne effects* (i.e. simply knowing they are part of a study makes participants feel important and improves their effort and performance: Franke & Kaul, 1978; Adair, 1984; Jones, 1992; Levitt & List, 2011); and *John Henry effects* (i.e. participants who perceive that they are in the control group exert greater effort because they treat the experiment like a competitive contest and they want to overcome the disadvantage of being in the control group: Campbell & Stanley, 1963; Cook & Campbell, 1979).

Other, more recent, typologies of randomized controlled experiments are *online experiments* (Horton et al., 2011) conducted, for instance, using Amazon's *Mechanical Turk* (*MTurk*) (Paolacci et al., 2010; Horton et al., 2011); *virtual experiments* on virtual reality settings (Fiore at al., 2009); and *lab-field experiments* that consist of a first-stage intervention under controlled conditions (in the lab) linked to a naturalistic situation (in the field) where subjects are not aware that their behavior is observed. Lab-field experiments have been used to look at the unintended 'behavioral spillover' effects of interventions (Dolan & Galizzi, 2014b; 2015; Dolan et al., 2015) or at the *external validity* of lab-based behavioral measures (Galizzi & Navarro-Martinez, 2015), and are part of the growing efforts to bridge the gap between the lab and the field, especially in policy areas like health that are inherently challenging from a methodological perspective (Hennig-Schmidt et al., 2011; Kesternich et al., 2013; Hennig-Schmidt & Wiesen, 2014).

RCT is therefore a vague and misleading term to use for experiments for policy formulation purposes as it does not convey key information on the exact nature and typology of the experiment. There is not one single type of experiment for policy formulation purposes; rather, a broad spectrum of different types of experiments spanning from the lab to the field can prove useful. As randomized controlled experiments, all of the experiment types along the Harrison and List (2004) spectrum provide unbiased estimates of the average treatment effect. This continuum of randomized controlled experiments (RCEs), therefore, is the systematic, and methodologically appropriate, generalization of what popularly referred to as RCTs.

It is also worthwhile emphasizing that there is no consensus on whether lab or field experiments are superior: both have strengths and weaknesses, and their relative merits have been systematically discussed elsewhere (Loewenstein, 1999; Starmer, 1999a; 1999b; Smith, 2003; Harrison & List, 2004; Guala, 2005; Levitt & List, 2007b, 2008; Bardsley, 2005; Falk & Heckman, 2009; Bardsley et al. 2009; Camerer, 2011; Harrison, 2013; Dolan & Galizzi, 2014a; Kagel, 2015). Briefly, *lab experiments* allow for high *internal validity* because of their ability to tightly control the environment and frame, minimize confounding factors, closely simulate conditions of theoretical models, and replicate past experiments. Furthermore, they provide insights into possible patterns prior to moving into the wild, they uncover the mechanisms underlying decisions and behavior, and they require significantly fewer financial, time and logistical resources than field experiments. *Field experiments*, on the other hand, generally enhance the *external validity* of experimental results because observations are made with subjects, environments, situations, tasks, rules and stakes which are closer to the ones occurring in the real world (Brookshire et al., 1987; Galizzi & Navarro-Martinez, 2015; Kessler & Vesterlund, 2015). Field experiments, however, come with lesser control and with several other limitations when used for policy purposes (Harrison, 2014). Moreover, they are inherently more difficult, if not impossible, to replicate. This is clearly a major limitation given the increasing attention to the *replicability* of experimental results in psychology, economics and social sciences (Burman et al., 2010; Miguel et al., 2015; Open Science Collaboration, 2015).

On a related note, it is worthwhile to notice that the way in which RCTs are sometimes contrasted to statistical or econometric analysis is also misleading. In fact, running any of the randomized controlled experiments along the Harrison and List (2004) spectrum is just the first step of the data collection process that allows the behavioral scientist to then conduct in-depth econometric analysis of experimental data. There is no reason why behavioral scientists interested in 'what works' should only superficially look at the average treatment effect across the control and the treatment groups in an experiment, instead of delving deeper into the behavioral nuances and causal mechanisms. Indeed, as witnessed by the burgeoning field of 'behavioral econometrics' and 'experimetrics', randomized controlled experiments and econometric analysis are complementary, not substitute, methods (Andersen et al., 2010; Harrison et al., 2015; Moffatt, 2015).

The third and last methodological clarification on 'behavioral' methods relates to the often proclaimed superiority of RCTs in terms of *generalizability*— the question of what other populations, settings, contexts or domains the findings from an experiment can be generalized to (Al-Ubaydli & List, 2015).

There are three conceptually distinct threats to generalizability. The first threat comes essentially from participation bias. Unlike natural field experiments, conventional lab experiments (but also artefactual and framed field experiments, and, as noted above, RCTs) recruit participants through an explicit invitation to take part in an experiment. As a result, there is bias not only because the potential participants—university students—self-select into universities, but also because subjects who choose to participate in experiments may be inherently different in their underlying characteristics from subjects who choose not to take part. Policy-makers should therefore be aware that, because of the participation bias, even if the initial sample of subjects is indeed representative of the target (or the general) population, the resulting sub-sample of actual respondents may not be. Students participating in lab experiments, for instance, have been found to be more curious (Slonim et al., 2013) and more motivated by financial incentives (Krawczyk, 2011; Charness et al., 2013).

The second threat comes from the fact that the environment, context and frame of the experimental decisions and tasks in the lab may not be representative of real situations encountered by subjects in natural settings. This limitation can be overcome by redesigning tasks and contexts to more closely match naturalistic situations that subjects are more familiar with in real life—that is, to design framed field experiments in the sense of Harrison and List (2004) (e.g. Harrison et al, 2007; Harrison & List, 2008).

The last threat to generalizability is that students are, clearly, a peculiar sample of experimental subjects that is not representative of the general population (Enis et al., 1972; Cunningham et al., 1974). For instance, students act less cooperatively and trustfully in social preferences games (Bellemare & Kroger, 2007). If students behave differently, then, an extrapolation of their behavior to the general population would be biased even after controlling for socio-demographics (Levitt & List, 2007a; Exadaktylos et al., 2013).

To overcome these limitations, a small but growing number of researchers have started running *artefactual field experiments* with *representative samples* of the population rather than just students (Andersen et al., 2008, 2014; Bellemare et al, 2008; Galizzi, 2012; Galizzi et al., 2016a,b).

4. A taxonomy of policy formulation instruments

A second source of potential misunderstanding in the behavioral aspects of policy formulation is concerned with the nature and insights, rather than the methods, of the behavioral policies. When formulating policies, researchers and decision-makers are ultimately interested in knowing which type of policies work in effectively changing behavior. Rather than resolving the fundamental issue of whether the assumptions of conventional or behavioral economics are correct at a general level, economists and social scientists are increasingly embracing a so-called pragmatic approach (Galizzi, 2014; Bhargava and Loewenstein, 2015; Chetty, 2015; Laibson and List, 2015). In essence, this approach considers behavioral economics as a natural progression of conventional economics, rather than a fundamental challenge to it, and thus combines insights from both conventional and behavioral economics that have proved to work effectively for public policy purposes. To better understand the conceptual background underpinning such a pragmatic approach to behavioral insights, it is important to explicitly describe the extent to which different behavioral policies actually depart from the conventional economics paradigm.

A taxonomy is proposed of five different clusters of policy formulation instruments: preferences-based policies; information-based policies; financial incentives; regulation-based policies including tax- and subsidy-based policies; and nudges. Before going into the details of each class of policies, Figure 24.1 graphically summarizes how the different clusters of policy formulation instruments relate to, or depart from, the conventional economics model.

[Figure 24.1 here]

In the framework, preferences-based policies are directly related to the first pillar of the conventional economics model, and for this reason one can categorize them under the conventional rather than the behavioral economics umbrella. Similarly, it is possible to closely associate a specific pillar of conventional economics with information-based policies (pillar 2), financial incentives (pillar 3), and regulation-based policies, including tax- and subsidy-based policies (pillar 4).

Informational policies and financial incentives can also be inspired by behavioral economics and behavioral science. In that case, these behaviorally inspired information-based policies and financial incentives can be grouped under the 'behavioral' umbrella.

According to the above framework, policies based on nudges substantially depart from conventional economics, as they openly challenge its pillars 1-3. Nudges can thus be viewed as a cluster of policy formulation instruments that are most closely and comprehensively inspired by behavioral science research. The following sub-sections review the rationale beyond each element of the taxonomy in greater detail.

4.1. Preferences-based policies.

The first cluster of policy formulation instruments are what can be called preferences-based policies. These are essentially based on the idea of providing citizens with broad sets and menus of choices from which they select their most favorite option. The broader these sets of choices are, the larger the set of possible profiles of preferences that could be satisfied.

In the U.S. health policy context, for example, under the George W. Bush administration, the Medicare Part D website was launched in 2006 to help seniors choose among a wide variety of different drug plans provided by private healthcare companies. Bush explained his reform of Medicare Part D by saying "The more choices you have, the more likely it is you'll be able to find a program that suits your specific needs" (White House, 2006).

Similarly, in October 2013 the so-called Obamacare reform launched the exchange portal https://www.healthcare.gov/ to help the 50 million US citizens without health insurance to compare, in a systematic way, the profiles of the healthcare insurances in 36 US states. Similar policies have been implemented, mainly in the US and the UK, in the form of internet comparison and ranking websites in several policy formulation areas, such as health (hospitals and doctors rating websites: Galizzi et al., 2012), education (school ratings), and pensions and savings.

While more choice is almost always good, especially when it fosters competition on the supply side, as in the Obamacare example, there is no genuine behavioral insight in these policies. This cluster of policies is actually soundly grounded on conventional economics: they assume that people have clear preferences over clinical treatments, diagnostic tests, insurance and pension schemes, schools, and so on, and that a broader set of choices will help them find their most preferred option. This is fully consistent with conventional economics (pillar 1).

From a behavioral science perspective, the potential benefit of broadening the choice set of options can be partly, or completely, offset by the paralyzing effect of having too many options among which to choose. As the 'tyranny of choice' literature shows, having more options often leads to worse, rather than better, choices, because making choices is effortful, tiring and can generate anxiety (Iyengar & Lepper, 2005; Salecl, 2010). This is also the reason why we often seek advices and suggestions; imitate what other do or just follow the crowd; or stick to default options (Thaler & Sunstein, 2008).

The second, methodologically more profound consideration is that it is not easy to demonstrate a rigorous link between larger choice sets and better decisions or behavior. Empirically answering this question entails facing the 'curse' of the '*revealed preferences*' argument in economics. In practice, it is often impossible to empirically identify the effect of broader choice sets on individual behavior, simply because economists traditionally assume that behavior is just the manifestation of underlying preferences. Therefore, as in most cases we do not directly observe preferences and we only observe behavior, any type of behavior, even the most extravagant, can be easily justifiable in light of some latent, possibly 'exotic', preferences (Loewenstein, 2007).

From a conceptual point of view, the only way to rigorously test the effectiveness of preference-based policies would be to directly measure individual preferences prior to the policy intervention and then to directly observe decision-making and behavior under different conditions where the number of options in the choice sets is systematically manipulated. Rigorous evidence on this point is scarce in most policy formulation contexts. Galizzi (2014) further illustrates the conceptual challenges and practical intricacies of this approach for the health policy domain.

4.2. Information-based policies.

Proceeding down the list, the next cluster of policy formulation instruments is centered on the idea of providing information to citizens and consumers to enable them to formulate better decisions.

Information-based policies are quite firmly grounded on conventional economics: accessing more, or better, information enables us to make better decisions and plans (pillar 2). It is worth mentioning three considerations about the effectiveness of information-based policies. First, the bulk of research in behavioral science suggests that merely providing more information is generally effective in raising awareness, but does not necessarily lead to significant and sustained change in behavior. Second, providing more information can actually trigger unintended spillover effects (Dolan & Galizzi, 2015). Third, policy interventions seem to be particularly effective when the type and the design of the information provided is directly inspired to, and 'supercharged' by, genuine insights from the behavioral science (e.g. the 'informational nudges' in Bhargava & Loewenstein, 2015). Galizzi (2014) further illustrates these considerations in the health policy context.

4.3. Financial incentives.

Next in the list are those policy formulation instruments based on financial incentives. A premise is in order here. We consider in such a cluster only the policies based on the idea of providing monetary incentives conditional to a specific change in behavior, what economists often called Conditional Cash Transfer (CCT): for instance, paying smokers £100 when they quit smoking, or giving £50 on completion of a professional development course.

It is important to distinguish these policies from other related policy instruments: providing monetary incentives based on a predefined action or change in behavior makes the financial incentives inherently different from subsidies or taxes.

Taxes and subsidies, of course, aim to (and are often able to) cause changes in behavior. But they do so by directly interfering with market prices. Financial incentives, on the contrary, do not alter market prices.

Following the above conceptual framework, however, financial incentives rely on limited behavioral insight, and are not a distinguishing feature of behavioral economics. Incentives are actually one of the main hallmarks of conventional economics, being directly related to pillar 3 of the above framework. Economics as a social science can actually be defined largely as the study of incentives and their impact on changing behavior: according to the 'basic law of behavior' (Gneezy et al., 2011), once a well-designed incentive is introduced, behavior should change in the envisaged direction.

There is strong evidence that purely monetary conditional incentives tend to work in the short run, but mixed evidence on whether they are capable of leading to sustained changes in behavior, especially after they are removed.

There is also compelling evidence that financial incentives work effectively when their design is directly inspired, and 'supercharged', by genuine insights from behavioral sciences. In particular, incentives work when, in coherence with the 'asymmetric paternalism' approach by Camerer et al. (2003) and the 'libertarian paternalism' approach by Thaler and Sunstein (2003; 2008), they are designed around human biases, in the attempt to help people to change behavior. From this perspective, 'behaviorally' supercharged financial incentives can be seen as an application of the 'nudging' policy approach, which is directly inspired by behavioral economics (see section 4.5).

An archetypical example of these behaviorally 'supercharged' incentives are those used for healthier behaviors by George Loewenstein, Kevin Volpp and colleagues at the Centre for Health Incentives and Behavioral Economics at (CHIBE) (Loewenstein et al., 2007; Volpp et al., 2011; Loewenstein et al., 2012). In their set of experiments on weight loss, for instance, incentives proved to work when they were designed to account for, and lever on, our biases. These include the tendencies to over-evaluate small probabilities (e.g. a 10% probability of paying $100, instead of a 100% probability of paying $10); attach a greater value to losses than gains of the same amount (e.g. by asking subjects to put their own money down in deposits that are then matched 1:1, and then deducting money from these deposits any time subjects fail to change behavior, playing on subjects' aversion to lose their deposit); be over-optimistic about personal achievements (e.g. when asked to put money down as a deposit, most people believe they will succeed in losing weight and do put down the money); appreciate immediate feedback on our actions (e.g. by providing immediate, personalized, and punctual feedback by text messages for both rewards and punishments); regret the actions we did not take in the past (e.g. by informing subjects about the money they could have earned if they had indeed changed behavior any time they did not). Galizzi (2014) illustrates these points further in the broader area of financial incentives in health.

A relatively underexplored, but highly promising area for policy formulation purposes is the study of the optimal combination of financial incentives and 'nudges' (Dolan & Galizzi, 2014b; Chetty, 2015). A further area of interest for policy formulation purposes is related to the unintended consequences of financial incentives, and leads to the new, promising, field of 'behavioral spillovers' of policy interventions (Dolan & Galizzi, 2014b; Dolan & Galizzi, 2015; (Dolan et al., 2015).

4.4. Regulation-based policies, including tax- and subsidies-based policies

Next on the list are policy formulation instruments based on regulation, including those based on taxes and subsidies. It is easy to argue that these policies are firmly grounded on conventional economics, in particular on pillar 4. They are essentially interventions related to, or directly inspired by, the long and noble history of market regulation in public economics; that is, in the attempt to overcome market failures, the policy-maker directly intervenes in markets to realign market forces and prices.

The most typical example of market failures in practice are 'externalities' (Pigou, 1920), when markets fail to take into account the overall social costs and benefits of goods and services and do not adequately reflect them in prices. The classic public economics instruments to correct such externalities are taxes and subsidies, for instance, carbon taxes levied on carbon- and oil-based energy resources.

It is well-known from public economics that regulations, taxes and subsidies are, conceptually, the most suitable forms of policy interventions to address and correct externalities, and successful real-world examples of implementation of these policies is abundant in both developed and developing countries. The point here is that, although these interventions are clearly inspired by conventional rather than behavioral economics, there is an immense potential to combine these traditional but effective public economics tools with new insights from the behavioral science. For instance, how can we design specific schemes around well-known human biases so that we can 'supercharge' taxes and subsidies with behavioral insights to enhance their long-term effectiveness? This seems one of the most exciting and promising areas where more experimental evidence is currently needed in behavioral public policy.

4.5. Nudges

Finally, some 'behavioral' policies are inspired by the idea of 'nudging' (Thaler & Sunstein, 2008). 'Nudges' essentially consist of changes in the decision environment (the so-called 'choice architecture'), designed on the basis of behavioral evidence, to trigger changes in behavior occurring at an automatic, or unconscious, level. Among many possible examples, there is the well-known case of changing the default option in organ donation statements (Thaler & Sunstein, 2008). Other examples in the health context are the behavioral interventions to nudge healthy eating: simply relocating fruits and vegetables in more salient spots in high schools cafeterias significantly increases their consumption (Hanks et al., 2012), while serving food in small plates, or in plates that have a high color contrast with the served food (e.g. white plates for spaghetti with tomato sauce) lead to significantly lower food intakes (Wansink & Van Ittersum, 2006; Van Ittersum & Wansink, 2012).

Unlike the other behaviorally inspired policies discussed above, nudges do not involve any financial incentives or release any new piece of information; they merely change the environment where choices and actions are taken.

This is quite a broad category that encompasses a vast range of policy instruments levering on human decision biases such as the ones introduced above and many others: status quo and default bias, loss aversion, procrastination, sunk cost fallacy, halo effects, anchoring, overweighting of small probabilities, illusion of control, availability bias, saliency and framing effects, present bias, confirmation bias, adaptation, and the ostrich effect, to name just a few (Thaler & Sunstein, 2008; Kahneman, 2011).

Nudges are the cluster of policy formulation instruments that are most genuinely and firmly grounded on insights from the behavioral sciences. For this reason, 'nudging' interventions should be regarded as the only group of policy formulation instruments that comfortably sit under the umbrella of behavioral, rather than conventional, economics. They are, in fact, essentially based on two findings by behavioral economics and applied behavioral science.

First, a great part of human behavior is automatic and non-conscious. This is consistent with the idea that our judgment and decision-making are informed by two cognitive interacting systems: a fast and automatic (non-conscious) system ('System 1') and a slow and deliberative (conscious) system ('System 2') (Chaiken & Trope, 1999; Kahneman 2003, 2011).

Second, related to this, we often make mistakes and errors in judgment and decision-making and fall prey to a broad range of biases and influences from environmental cues, and to a large extent may even be unsure of what we actually want. According to the behavioral view, our judgments and preferences are malleable in that they can be affected and shifted, even substantially, by subtle differences in the social environment, the decision frame, and the cognitive or visual representations of alternatives. To the extreme, our evaluations and preferences are constructed on the moment in a given situation, and are thus affected by changes in the choice environment (Lichtenstein & Slovic, 2006). Such shifts and changes can actually occur even when we are not consciously aware of it.

Both ideas are at odds with the conventional economics idea that we make rational deliberations about what is optimal given our stable set of preferences, and we then undertake a full and coherent plan of action. In this view, in the long run our actions and decisions thus fluctuate around, and reveal, our stable set of preferences, so that we do not make systematic errors and biases. The very core of conventional economics as summarized in 'pillars' 1-2 is about rational deliberative decision-making.

It is mainly on this ground that nudging policies challenge the conventional economics view. Nudges, however, do not interfere with the sets of options available to individuals, nor with market mechanisms, and do not distort the behavior of those who act rationally. Nudges are thus less intrusive in the market mechanisms than taxes or subsidies.

Under the perspective of the degree of 'intrusiveness' of policy formulation instruments, it is possible to establish a parallel between taxes and subsidies, on one side, and nudges, on the other. Taxes and subsidies are levied to deal with externalities and market failures. If the aim of the policy is indeed to correct these externalities, taxes and subsidies seem the most appropriate conventional economics tools.

Nudges, on the other hand, are best employed to deal with 'internalities' (Galizzi, 2014; Bhargava & Loewenstein, 2015). Internalities are essentially costs that we impose on ourselves, and that we do not (sufficiently) take into account in our decisions (Herrnstein et al., 1993). These internalities' costs originate from our own errors and failures in judgment and decision-making, rather than from market failures.

Internalities are perhaps a more fundamental source of flaws and failures than externalities, as they pre-exist to markets and economic institutions. They also represent a bigger challenge as they cannot be removed by conventional policy formulation instruments such as taxes and subsidies. In principle, the internal failures and biases in human decision-making likely survive even when externalities are addressed by direct market intervention.

Because the application of nudges to policy formulation is relatively recent, it is perhaps premature to draw conclusions on the effectiveness of nudges based on systematic reviews of the evidence (Marteau et al., 2011; Loewenstein et al., 2012). The picture gathered by different streams of literature, however, is quite clear in suggesting that even subtle changes in the 'choice architecture' can lead to significant changes in behavior in a variety of policy formulation domains. An interesting area of investigation for policy formulation purposes is related to the recent evidence on the unintended 'behavioral spillovers' of nudges (Dolan & Galizzi, 2015; Loewenstein et al., 2015).

5. Conclusions

There is an increasing interest in applying behavioral insights to policy formulation challenges. Researchers and policy-makers have recently discussed a range of diverse policy formulation instruments whose terms are often used interchangeably and described as 'behavioral'.

We provide a critical review of two intertwined conceptual challenges related to these behavioral policies. We start by making a distinction between two main ingredients of so-called 'behavioral' public policies, namely methods and insights.

We then argue that what is often indicated as an inherent 'behavioral' aspect of policy formulation is simply the use of the experimental method to assess the effectiveness of a policy intervention. On this respect, we argue that, while the experimental method represent quite an innovation in the policy formulation arena, it is not a unique feature of 'behavioral' policies nor of behavioral economics.

We also argue that reducing the use of the experimental method for policy purposes to the so-called Randomized Controlled Trials (RCTs) is both conceptually misleading and over-simplistic, since it overlooks the richness of the toolbox of different typologies of randomized controlled experiments (RCEs) spanning the spectrum between the lab and the field in the sense of Harrison and List (2004).

We then argue that to assess the real behavioral component of a behavioral policy formulation, we should pay close attention to what extent the policy is inspired by genuine insights from the behavioral sciences. To achieve this, we propose a taxonomy to classify policy formulation instruments in five clusters: preferences-based policies; information-based policies; financial incentives; regulation-based policies, including tax- and subsidy-based policies; and nudges.

It is possible to look at these five classes of policy formulation instruments in terms of how far away they move from conventional economics. Policy formulation instruments aiming to provide broader sets of choices, more information, financial incentives, or to use taxes and subsidies (the first four 'clusters') are closer in their conception to conventional than behavioral economics.

Policy formulation instruments based on the nudging approach are directly inspired by insights from behavioral economics, although behavioral insights have also been applied to the design of information-based policies and financial incentives, and in the near future can be fruitfully combined with regulation-based policies.

While policy formulation instruments genuinely inspired by behavioral economics can successfully address internalities failures, they are unlikely to effectively deal with externalities and market failures that are better addressed by conventional economics interventions, such as taxes, subsidies, and other forms of regulation. There is, however, a widely unexplored potential to enhance the effectiveness of regulatory schemes using insights genuinely inspired by the behavioral sciences.

We conclude with two main considerations that highlight two parallel complementarities in the behavioral aspects of policy formulation, the first related to behavioral methods, the second to behavioral insights.

On methods, although randomized controlled experiments are not a distinguishing feature of behavioral economics or behavioral policy formulation, their growing employment by policy-makers should be welcome for formulating, testing, assessing, and fine-tuning policy interventions. In particular, the systematic use of a broad spectrum of complementary randomized controlled experiments in the lab and the field should be advocated as a powerful approach for finding out what works and does not work for policy formulation purposes, and for innovatively integrating the often detached phases of policy evaluation, cost-effectiveness analysis, subjective well-being measurement, and welfare analysis (Dolan & Galizzi, 2014a; Harrison, 2014; Chetty, 2015). In close association with econometric analysis and through innovative 'behavioral data linking' designs, such a flexible toolbox of complementary experiments can also be combined with the fast-growing wealth of survey data, administrative records, online panels, smart cards, and biomarkers to kick off a genuine revolution in the way evidence is used to inform policy decision-making (Jenkins at al., 2008; Dolan & Galizzi, 2014a, 2015; Galizzi et al., 2016a,b).

On insights, among economists and social scientists working in public policy areas there is now an increasing tendency to emphasize the complementarities, rather than the contradictions, between conventional and behavioral economics (Galizzi, 2014; Bhargava & Loewenstein, 2015; Chetty, 2015; Laibson & List, 2015). In this recent view, rather than challenging or rejecting conventional economics, behavioral economics is seen as a natural augmentation or progression of conventional economics. It is now widely recognized that insights from behavioral economics offer important policy tools that can be used to change behavior, from default options to nudges, to 'behaviorally supercharged' incentives. In policy formulation, these behavioral insights also provide better predictions of the conditions under which policy interventions could work, and of the mechanisms leading to behavioral change or to unintended 'behavioral spillover' effects (Dolan & Galizzi, 2015). Moreover, behavioral insights can also help to generate welfare analysis in policy formulation areas where individual choices suffer from imperfect attention or from the above described behavioral biases; where individual preferences are not consistent; or where 'decision' utilities typically differ from 'experience' utilities (Kahneman et al., 1997; Kahneman & Sugden, 2005; Dolan & Kahneman, 2008). This tendency calls for an even more systematic integration of perspectives for policy formulation purposes, by breaking down the traditional disciplinary silos, and by cross-fertilizing 'behavioral' science insights from a broad range of fields, from economics and political science to psychology, from neuroscience to biology, from philosophy to happiness and wellbeing research (Dolan & Galizzi, 2014a).

Acknowledgements

Comments and suggestions from Michael Howlett and Ishani Mukherjee and helpful editing by Libby Morgan Beri are gratefully acknowledged. Funding from the ESRC is also gratefully acknowledged (ESRC Future Research Leader Fellowship ES/K001965/1 PI: MM Galizzi).

References

Adair, G. (1984). The Hawthorne effect: a reconsideration of the methodological artefact. *Journal of Applied Psychology*, 69, 334-345.

Al-Ubaydli, O., & List, J. A. (2015). On the generalizability of experimental results in economics. In G. Frechette and A. Schotter (Eds.) *Handbook of Experimental Economic Methodology.* Oxford: Oxford University Press.

Al-Ubaydli, O., & List, J.A. (2015). Do natural field experiments afford researchers more or less control than laboratory experiments? *American Economic Review: Papers & Proceedings, 105(5)*, 462-466.

Andersen, S., Harrison, G.W., Lau, M., Rutstrom, E.E. (2008). Eliciting risk and time preferences. *Econometrica*, 76(3), 583-618.

Andersen, S., Harrison, G.W., Lau, M., Rutstrom, E.E. (2010). Behavioral econometrics for psychologists. *Journal of Economic Psychology*, 31, 553-576.

Andersen, S., Harrison, G.W., Lau, M., Rutstrom, E.E. (2014). Discounting behavior: a reconsideration. *European Economic Review*, 71, 15-33.

Angrist, J. D., & Imbens, G. W. (1995). Two-Stage Least Squares estimation of average causal effects in models with variable treatment intensity. *Journal of the American Statistical Association*, *90*, 430-442.

Angrist, J. D., Imbens, G. W., & Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association, 91,* 444-472.

Angrist, J. D., & Krueger, A.B. (1991). Does compulsory schooling attendance affect schooling and earnings? *Quarterly Journal of Economics*, *106,* 976-1014.

Angrist, J. D., & Krueger, A. B. (2001). Instrumental variables and the search for identification: from supply and demand to natural experiments. *Journal of Economic Perspectives, 15(4)*, 69-85.

Angrist, J. D., & Lavy, V. (1999). Using Mamonides' rule to estimate the effect of of class size on scholastic achievement. *Quarterly Journal of Economics, 114,* 533-575.

Angrist, J. D., Pischke, J.S. (2009). *Mostly harmless econometrics: an empiricist's companion*. Princeton: Princeton University Press.

Annala M., Kaskinen T., Lee, S., Leppanen, J., Mattila, K., Neuvonen, A., Nuutinen, J., Saarikoski, E., & Tarvainen, A. (2015). *Design for government: human-centric*

governance theory through experiments. Demos Helsinki – Avanto Helsinki Report for Prime Minister's Office's Design for Government Project.
Ashenfelter, O. (1978). Estimating the effect of training programs on earnings. *Review of Economics and Statistics*, *60*, 47-57.
Ashenfelter, O., & Krueger, A. (1994). Estimates of the economic return to schooling from a new sample of identical twins. *American Economic Review, 84(5),* 1157-1174.
Ashraf, N., Camerer, C., & Loewenstein, G. (2005). Adam Smith, behavioral economist. *Journal of Economic Perspectives*, *19(3),* 131-145.
Bardsley, N. (2005). Experimental economics and the artificiality of alteration. *Journal of Economic Methodology, 12(2),* 239-251.
Bardsley, N., Cubitt, R., Loomes, G., Moffatt, P., Starmer, C., & Sugden, R. (2009). *Experimental Economics: Rethinking the Rules*. Princeton, NJ: Princeton University Press.
Baron, J. N., & Hannan, M. T. (1994). The impact of economics on contemporary sociology. *Journal of Economic Literature*, *32,* 1111-1146.
Bellemare, C., & Kroger, S. (2007). On representative social capital. *European Economic Review*, *51*, 183-202.
Bellemare, C., Kroger, S., & van Soest, A. (2008). Measuring inequity aversion in a heterogeneous population using experimental decisions and subjective probabilities, *Econometrica*, *76,* 815-39.
Bernheim, B. D. (2009). Behavioral welfare economics. *Journal of the European Economic Association, 7(2-3),* 267-319.
Bernheim B. D., & Rangel A. (2007): Toward choice-theoretic foundations for behavioral welfare economics. *American Economic Review Papers & Proceedings, 97*, 464-470.
Bernheim, B. D., Rangel, A. (2009). Beyond revealed preference: choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics*, *124(1),* 51-104.
Beshears, J., Choi, J. J., Laibson, D., & Madrian, B.C. (2011): Behavioral economics perspectives on public sector pension plans. *Journal of Pension Economics and Finance, 10(2),* 315-336.
Bhargava, S., & Loewenstein, G. (2015). Behavioral economics and public policy 102: beyond nudging. *American Economic Review: Papers & Proceedings*, *105(5),* 396-401.
Brookshire, D., Coursey, D. L., & Schulze, W.D. (1987). The external validity of experimental economics techniques: analysis of demand behavior. *Economic Inquiry, 25,* 239-250.
Buhrmester, M. D., Kwang, T., & Gosling, S. D. (2011). Amazon's Mechanical Turk: a new source of inexpensive, yet high-quality, data? *Perspectives on Psychological Science, 6,* 3-5.
Burman, L. E., Reed, R. R., & Alm, J. (2010). A call for replication studies. *Public Finance Review*, *38(6),* 787-793.
Burtless, G. (1995). The case for randomised field trials in economic and policy research. *Journal of Economic Perspectives, 9,* 63-84.
Camerer, C. (1999). Behavioral economics: reunifying psychology and economics. *Proceedings of the National Academy of Sciences USA 96(19),* 10575-10577.
Camerer, C.F., & Hogarth, R.M. (1999): The effects of financial incentives in experiments: a review and capital-labor-production framework, *Journal of Risk and Uncertainty, 19 (1-3),* 7-42.
Camerer, C.F., Issacharoff, S., Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003): Regulation for conservatives: behavioral economics and the case for asymmetric paternalism. *University of Pennsylvania Law Review, 1151*, 1211-1254.
Camerer, C.F. (2011). The promise and success of lab-field generalizability in experimental economics: a critical reply to Levitt and List. SSRN working paper.

Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
Card, D. (1992). Using regional variation to measure the effect of the Federal minimum wage. *Industrial and Labor Relation Review, 46,* 22-37.
Card, D. (1996). The effect of unions on the structure of wages: a longitudinal analysis. *Econometrica, 64,* 957-979.
Card, D., & Krueger, A. (1994). Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania. *American Economic Review, 84,* 772-784.
Card, D., & Krueger, A. (2000). Minimum wages and employment: a case study of the fast-food industry in New Jersey and Pennsylvania: Reply. *American Economic Review, 90*, 1397-1420.
Chaiken, S., & Trope, Y. (1999). *Dual-process theories in social psychology*. New York: The Guilford Press.
Chandon P., & Wansink B. (2007). The biasing health halos of fast-food restaurant health claims: lower calorie estimates and higher side-dish consumption intentions. *Journal of Consumer Research*, *24*, 301-314.
Charness, G., Gneezy, U., & Kuhn, M. A. (2013). Experimental methods: extra-laboratory experiments – extending the reach of experimental economics, *Journal of Economic Behaviour and Organization, 91*, 93-100.
Chetty, R. (2015). Behavioral economics and public policy: a pragmatic perspective. *American Economic Review: Papers & Proceedings, 105(5),* 1-33.
Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: theory and evidence. *American Economic Review, 99(4),* 1145-1177.
Cleave, B., Nikiforakis, N., & Slonim, R. (2013). Is there selection bias in laboratory experiments? The case of social and risk preferences. *Experimental Economics*, *16*, 372-382.
Congdon, W. J., Kling, J. R., & Mullainathan, S. (2011). *Policy and choice: Public finance through the lens of behavioral economics*. Washington, DC: Brookings Institution Press.
Cook, T. D. (2008). Waiting for life to arrive: a history of the Regression-Discontinuity design in psychology, statistics, and economics. *Journal of Econometrics, 142,* 636-654.
Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design and analysis for field settings*. Chicago: Rand McNally.
Costa-Font, J. (2011). Behavioural welfare economics: does 'behavioural optimality' matter? *CESifo Economic Studies*, *57(4),* 551-559.
Croson, R., & Gachter, S. (2010). The science of experimental economics. *Journal of Economic Behaviour and Organization*, *73(1),* 122-131.
Crump, M. J .C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research, *PLoS ONE 8(3).*
Cunningham, W. H., Anderson, T., & Murphy, J. H. (1974). Are students real people? *Journal of Business, 47,* 399-409.
Davis, D. D., & Holt, C. A. (1993). *Experimental economics*. Princeton University Press: Princeton.
DellaVigna, S. (2009). Psychology and economics: evidence from the field. *Journal of Economic Literature*, *47(2),* 315-372.
Dolan, P., & Galizzi, M. M. (2014a). Getting policy-makers to listen to field experiments. *Oxford Review of Economic Policy, 30(4),* 725-752.
Dolan, P., & Galizzi, M. M. (2014b). Because I'm worth it. A lab-field experiment on spillover effects of incentives in health. LSE CEP Discussion Paper CEPDP1286.
Dolan, P., & Galizzi, M. M. (2015): Like ripples on a pond: behavioral spillovers and their consequences for research and policy. *Journal of Economic Psychology*, *47,* 1-16.

Dolan, P., Galizzi, M. M., & Navarro-Martinez, D. (2015). Paying people to eat or not to eat? Carryover effects of monetary incentives on eating behavior. *Social Science and Medicine*, *133*, 153-158.
Dolan, P., Hallsworth, M., Halpern, D., King, D., Metcalfe, R., & Vlaev, I. (2012). Influencing behaviour: the Mindspace way. *Journal of Economic Psychology, 33(1),* 264-277.
Dolan, P., & Kahneman, D. (2008). Interpretations of utility and their implications for the valuation of health. *Economic Journal, 118(525),* 215-234.
Enis, B. E., Cox, K., & Stafford, J. (1972). Students as subjects in consumer behavior experiments. *Journal of Marketing Research, 6,* 72-74.
Ermisch, J., Gambetta, D., Laurie, H., Siedler, T., & Uhrig, S. C. N. (2009) Measuring People's Trust, *Journal of the Royal Statistical Society, Series A, 172(4),* 749-69
European Commission DG SANCO (2013): Conference: Applying behavioural insights to policy-making, http://ec.europa.eu/dgs/health_consumer/information_sources/consumer_affairs_events_en.htm.
Exadaktylos, F., Espin, A. M., & Branas-Garza, P. (2013). Experimental subjects are not different. *Nature Scientific Reports, 3(1213).* Doi: 10.1038/srep01213.
Falk, A., & Heckman, J. J. (2009). Lab experiments are a major source of knowledge in the social sciences. *Science,* 326, 535-538.
Falk, A., Meier, S., & Zehnder, C. (2013). Do lab experiments misrepresent social preferences? The case of self-selected student samples. *Journal of European Economic Association, 11(4),* 839-852.
Feltovich, N. (2011). What's to know about laboratory experimentation in economics? *Journal of Economic Surveys, 25(2),* 371-379.
Ferber, R., & Hirsch, W. Z. (1978). Social experimentation and economic policy: a survey. *Journal of Economic Literature, 16(4),* 1379-1414.
Ferber, R., & Hirsch, W. Z. (1982). *Social experimentation and economic policy*. Cambridge: Cambridge University Press.
Fiore, S. M., Harrison, G. W., Hughes, C. E., & Rutstrom, E. E. (2009). Virtual experiments and environmental policy. *Journal of Environmental Economics and Management, 57,* 65-86.
Franke, R. H., & Kaul, J. D. (1978). The Hawthorne experiments: first statistical interpretation. *American Sociological Review, 43(5),* 623-643.
Galizzi, M.M.. (2012). Linking survey and experimental data: Behavioural experiments in health and wellbeing. ESRC Future Research Leader Fellowship Case for Support ES/K001965/1, London School of Economics: http://www.esrc.ac.uk/my-esrc/grants/ES.K001965.1/read; http://gtr.rcuk.ac.uk/project/612D2922-3497-42C6-BBB6-F1CF7A4FFA99
Galizzi, M. M. (2014). What is really behavioural in behavioural health policies? And, does it work? *Applied Economics Perspectives and Policy, 36(1),* 25-60.
Galizzi, M. M., & Navarro-Martinez, D. (2015). On the external validity of social preferences games: a systematic lab-field study. Barcelona GSE Working Paper 802.
Galizzi, M. M., Harrison, G. W., & Miniaci, R. (2016a). Linking experimental and survey data for a UK representative sample: structural estimation of risk and time preferences. London School of Economics.
Galizzi, M. M., Machado, S. R., & Miniaci, R. (2016b). Temporal stability, cross-validity, and external validity of risk preferences measures: experimental evidence from a UK representative survey. London School of Economics.
Gerber, A. S., & Green, D. P. (2012). *Field experiments: Design, analysis, and interpretation*, New York: W.W. Norton & Company.

Gneezy, U., Meier, S., & Rey-Biel, P. (2011): When and why incentives (don't) work to modify behaviour. *Journal of Economic Perspectives*, *25(4)*, 191-210.
Goodman, J. K., Cryder, C. E., & Cheema, A. (2013). Data collection in a flat world: the strengths and weaknesses of Mechanical Turk samples. *Journal of Behavioral Decision Making*, *26(3)*, 213-24.
Greaves, F., Pape, U. J., King, D., et al., (2012). Associations between web-based ratings and objective measures of hospital quality. *Archives of Internal Medicine*, *172*, 435-436.
Green, D. P., & Gerber, A. S. (2003). The underprovision of experiments in political science. *The Annals of the American Academy of Political and Social Science*, *589*, 94-112.
Guala, F. (2005). *The methodology of experimental economics*. Cambridge: Cambridge University Press.
Hahn, J., Todd, P., & Van Der Klaauw, W. (2001). Identification and estimation of treatment effects with a Regression-Discontinuity Design. *Econometrica*, *69*, 201-209.
Hanks, A.S., Just, D.R., Smith, L.E., Wansink, B. (2012). Healthy convenience: nudging students toward healthier choices in the lunchroom. *Journal of Public Health*, 34(3), 370-376.
Hargreaves Heap, S. (2013). What is the meaning of behavioural economics? *Cambridge Journal of Economics*, *37(5*), 985-1000.
Harrison, G. W. (2010). The behavioral counter-revolution. *Journal of Economic Behavior and Organization*, *73*, 49-57.
Harrison, G. W. (2013). Field experiments and methodological intolerance. *Journal of Economic Methodology*, *20(2)*, 103-117.
Harrison, G. W. (2014). Cautionary notes on the use of field experiments to address policy issues. *Oxford Review of Economic Policy*, *30(4)*, 753-763.
Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic Literature*, *42*, 1009-1055.
Harrison, G. W., & List, J. A. (2008). Naturally occurring markets and exogenous laboratory experiments: a case study of the winner's curse. *The Economic Journal*, *118*, 822-843.
Harrison, G. W., List, J. A., & Towe, C. (2007). Naturally occurring markets and exogenous laboratory experiments: a case study of risk aversion. *Econometrica*, *75(2)*, 433-458.
Harrison, G.W., Lau, M.I., & Rutstrom, E.E. (2015). Theory, experimental design, and econometrics are complementary (and so are lab and field experiments). In G. Frechette & A. Schotter (Eds.), *Handbook of experimental economic methodology*. Oxford: Oxford University Press
Hausman, J., Wise, D. (1985). Technical problems in social experimentation: cost versus ease of analysis. In J. Hausman, & D. Wise (Eds), *Social experimentation*. Chicago: University of Chicago Press for NBER, 187-208.
Haynes, L., Service, O, Goldacre, B., & Torgerson, D. (2012). Test, learn, adapt: developing public policy with randomized controlled trials. Cabinet Office Behavioural Insights Team: https://www.gov.uk/government/publications/test-learn-adapt-developing-public-policy-with-randomised-controlled-trials.
Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, *47*, 153-162.
Heckman, J. J. (1992). Randomization and social program evaluation. In C. Manski & I. Garfinkel. (Eds), *Evaluating welfare and training programs*. Cambridge University Press, 201-230.
Heckman, J. J., Ichimura, H., & Todd, P. E. (1998). Matching as an econometric evaluation estimator. *Review of Economic Studies*, *65*, 261-294.
Heckman, J. J., & Smith, J. A. (1995). Assessing the case for social experiments. *Journal of Economic Perspectives*, *9(2)*, 85-110.

Hennig-Schmidt, H., Selten, R., & Wiesen, D. (2011). How payment systems affect physicians' provision behavior: an experimental investigation. *Journal of Health Economics, 30(4),* 637-646.
Hennig-Schmidt, H., & Wiesen, D. (2014). Other-regarding behavior and motivation in health care provision: an experiment with medical and non-medical students. *Social Science & Medicine*, *108,* 156-165.
Herrnstein, R., Loewenstein, G., Prelec, D., & Vaughan, W. Jr. (1993). Utility maximization and melioration: internalities in individual choice. *Journal of Behavioral Decision Making, 6(3),* 149-185.
Hertwig, R., & Ortmann, A. (2001). Experimental practices in economics: a challenge for psychologists? *Behavioral and Brain Sciences*, *24,* 383-451.
Hitsch, G.J., Hortacsu, A., & Ariely, D. (2010). Matching and Sorting in Online Dating. *American Economic Review*, *100(1),* 130-163.
Horton, J., Rand, D., & Zeckhauser, R. (2011). The online laboratory: conducting experiments in a real labor market. *Experimental Economics, 14(3),* 399-425.
Imbens, G. (2004). Semiparametric estimation of average treatment effects under exogeneity: a review. *Review of Economics and Statistics, 86*, 4-29.
Imbens, G., & Lemieux, T. (2008). Regression discontinuity designs: a guide to practice. *Journal of Econometrics*, *142,* 615-635.
Iyengar, S. S., & Lepper, M. R. (2005). When choice is demotivating: can one desire too much of a good thing? *Journal of Personality and Social Psychology*, *79*, 996-1106.
Jenkins, S., Cappellari, L., Lynn, P., Jäckle, A., & Sala, E. (2004). Patterns of consent: Evidence from a general household survey. *Journal of the Royal Statistical Society: Series A, 169(4),* 701-722.
Jenkins, S., Lynn, P., Jäckle, A., & Sala, E. (2008). The feasibility of linking household survey and administrative record data: new evidence for Britain. *International Journal of Social Research Methodology, 11(1),* 29-43.
John, L., Loewenstein, G., Troxel, A., Norton, L., Fassbender, J., & Volpp, K. (2011). Financial incentives for extended weight loss: a randomized, controlled trial. *Journal of General Internal Medicine, 26(6),* 621-626.
John L., Loewenstein, G., & Volpp, K. G. (2012). Empirical observations on longer-term use of incentives for weight loss. *Preventive Medicine*, *55(1),* S68-S74.
John, P., Cotterill, S., Moseley, A., Richardson, L., Smith, G., Stoker, G., & Wales, C. (2011). *Nudge, nudge, think, think: Experimenting with ways to change civic behaviour*. London: Bloomsbury Publishing.
Jones, S. R. G. (1992). Was there a Hawthorne effect? *American Journal of Sociology*. *98,* 451-148.
Kagel, J. H. (2015). Laboratory experiments: the lab in relationship to field experiments, field data, and economic theory. In G. Frechette & A. Schotter (Eds.), *Handbook of experimental economic methodology*. Oxford: Oxford University Press.
Kahneman, D. (2003). Maps of bounded rationality: Psychology for behavioral economics. *American Economic Review*, *93(5),* 1449-14.
Kahneman, D., Sugden, R. (2005). Experienced utility as a standard of policy evaluation. *Environmental and Resource Economics*, *32(1),* 161-181.
Kahneman, D. (2011). *Thinking, Fast and Slow*. London: Penguin.
Kahneman, D. (2012). Foreword, in E. Shafir (Ed.), *The behavioral foundations of public policy*. Princeton: Princeton University Press.
Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica*, *47(2),* 263-291.

Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics, 112(2),* 375-405.
Kessler, J. B., & Vesterlund, L. (2015). The external validity of laboratory experiments: the misleading emphasis on quantitative effects. In G. Frechette & A. Schotter (Eds.), *Handbook of experimental economic methodology*. Oxford: Oxford University Press.
Kesternich, I., Heiss, F., McFadden, D., & Winter, J. (2013). Suit the action to the word, the word to the action: hypothetical choices and real decisions in Medicare Part D. *Journal of Health Economics*, *32*, 1313-1324.
Kramer, M., & Shapiro, S. (1984). Scientific challenges in the application of randomized trials. *Journal of the American Medical Association*, *252*, 2739-2745.
Krawczyk, M. (2011). What brings subjects to the lab? *Experimental Economics, 14(4),* 482-489.
Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics*, *114*, 497-532.
Lalonde, R. J. (1986). Evaluating the econometric evaluations of training programs using experimental data. *American Economic Review*, *76*, 602-620.
Laibson, D., & List, J. A. (2015). Principles of (behavioral) economics. *American Economic Review: Papers & Proceedings, 105(5),* 385-390.
Lazear, E. (2000): Economic imperialism. *Quarterly Journal of Economics*, *115(1),* 99-146.
Lemieux, T. (2008). The changing nature of wage inequality. *Journal of Population Economics*, *21*, 21-48.
Levitt, S. D., & List, J. A. (2007a). What do laboratory experiments measuring social preferences reveal about the real world? *Journal of Economic Perspectives*, *21(2),* 153-174.
Levitt, S. D., & List, J. A. (2007b). Viewpoint: on the generalizability of lab behaviour to the field. *Canadian Journal of Economics, 40(2),* 347-370.
Levitt, S. D., & List, J. A. (2008). Homo economicus evolves. *Science*. *319*, 909-910.
Levitt, S. D., & List, J. A. (2011). Was there really a Hawthorne effect at the Hawthorne Plant? An analysis of the original illumination experiments. *American Economic Journal: Applied Economics, 3,* 224-238.
Lichtenstein, S., & Slovic, P. (2006). *The construction of preferences*. Cambridge: Cambridge University Press.
List, J. A. (2004). Young, selfish, and male: field evidence on social preferences. *Economic Journal*, *114,* 121-149.
List, J. A. (2006). Field experiments: a bridge between lab and naturally occurring data. *The BE Journal of Economic Analysis and Policy, 6(2),* 8.
List, J. A. (2011). Why economists should conduct field experiments and 14 tips for pulling one off. *Journal of Economic Perspectives*, *25(3),* 3-16.
Loewenstein, G. (1996). Out of control: Visceral influences on behavior. *Organizational Behavior and Human Decision Processes, 65(3):* 272-292.
Loewenstein, G. (1999). Experimental economics from the vantage-point of behavioral economics. *Economic Journal*, 109, F23-F34.
Loewenstein, G. (2005a). Hot-cold empathy gaps and medical decision-making. *Health Psychology*, *24(4),* S49-S56.
Loewenstein, G. (2005b). Projection bias in medical decision-making. *Medical Decision Making, 25*, 96-105.
Loewenstein, G. (2007): *Exotic Preferences: Behavioural Economics and Human Motivation*. Oxford: Oxford University Press.
Loewenstein, G., Brennan, T., & Volpp, K. G. (2007). Asymmetric paternalism to improve health behaviors. *Journal of the American Medical Association, 298(20),* 2415-2417.
Loewenstein, G., Asch, D. A., Friedman, J. Y., Melichar, L. A., & Volpp, K. G. (2012). Can behavioural economics make us healthier? *British Medical Journal, 344,* 1-3.

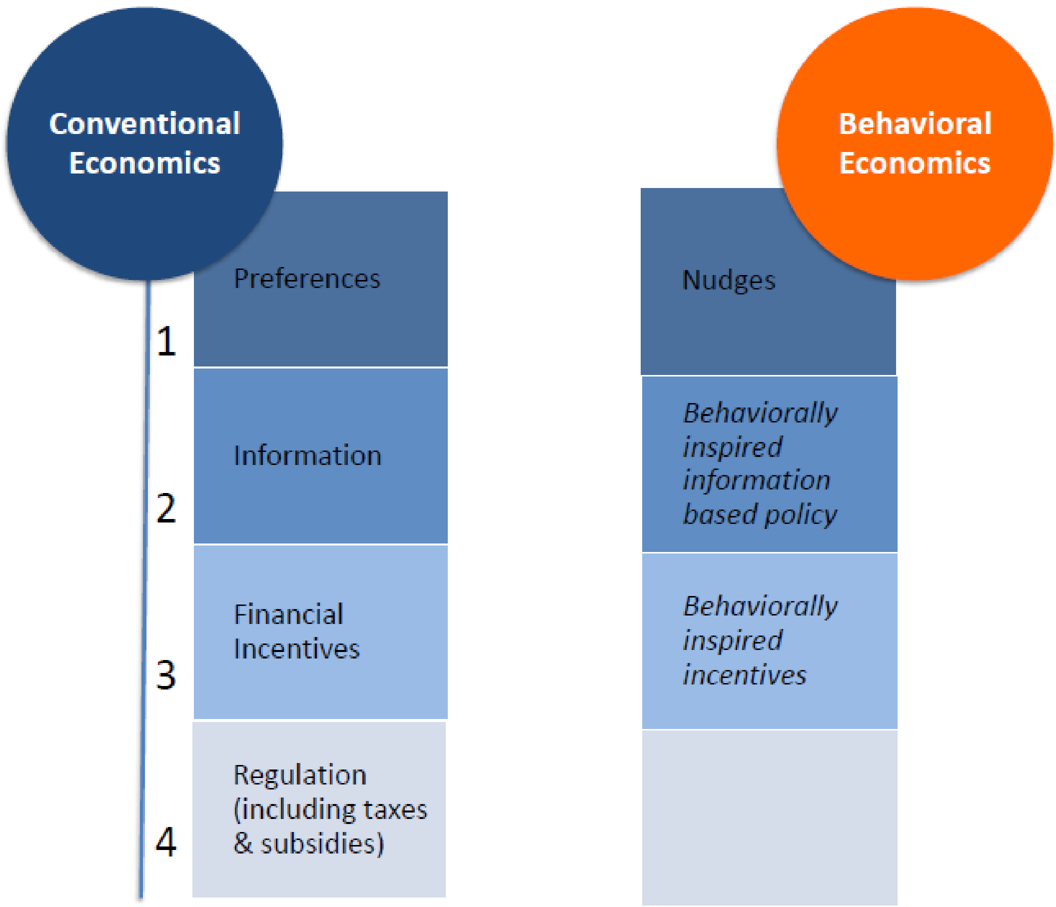
Loewenstein, G., Bryce, C., Hagmann, D., & Rajpal, S. (2015). Warning: you are about to be nudged. *Behavioral Science & Policy, 1(1),* 35-42.
Ludwig, J., Kling, J. R., Mullainathan, S. (2011). Mechanism experiments and policy evaluations. *Journal of Economic Perspectives, 25(3),* 17-38.
Madrian, B. C. (2014). Applying insights from behavioural economics to policy design. *Annual Review of Economics, 6(1),* 663-688.
Manning, W. D., Newhouse, J. P., Duan, N., Keeler, E. B., Leibowitz, A., & Marquis, S. M. (1987). Health insurance and the demand for medical care: evidence from a randomized experiment. *American Economic Review*. *77*, 251-277.
Marteau, T., Ogilvie, D., Roland, M., Suhrcke, M., Kelly, M.P. (2011). Judging nudging: can nudging improve population health. *British Medical Journal*, *342*, d228.
Mason, W., & Suri, S. (2012). Conducting behavioral research on Amazon's Mechanical Turk. *Behavior Research Methods*, *44(1),* 1-23.
Miguel, E., Camerer, C. F., Casey, K., Cohen, J., Esterling, K. M., Gerber, A., Glennerster, R., Green, D. P., Humphreys, M., Imbens, G., Laitin, D., Madon, T., Nelson, L., Nosek, B.A., Petersen, M., Sedlmayr, R., Simmons, J.P., Simonsohn, U., and Van der Laan, M. (2014). Promoting transparency in social science research. *Science, 343(6166),* 30-31.
Moffatt, P.G. (2015). *Experimetrics: Econometrics for Experimental Economics*, London: Palgrave MacMillan.
Mullainathan, S., Schwartzstein, J., & Congdon, W. J. (2012). A reduced-form approach to behavioral public finance. *Annual Review of Economics*, *4*, 511-540.
Oliver, A. (2013). *Behavioural public policy*. Cambridge: Cambridge University Press.
Open Science Collaboration (2015). Estimating the reproducibility of psychological science. *Science*, *349(6251).*
Paolacci, G., Chandler, J., & Ipeirotis, P. G. (2010). Running experiments on Amazon Mechanical Turk. *Judgment and Decision Making*, *5*, 411-419.
Paolacci, G., & Chandler, J. (2014). Inside the Turk: understanding Mechanical Turk as a participant pool. *Current Directions in Psychological Science*, *23(3),* 184-188.
Paserman, M. D. (2008). Job search and hyperbolic discounting: structural estimation and policy evaluation. *Economic Journal*, *118(531),* 1418-1452.
Pigou, A. C. (1920). *Economics of Welfare*. London: Macmillan.
Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, *70*, 41-55.
Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using sub-classification on the propensity score. *Journal of the American Statistical Associatio*n, *79,* 516-524.
Rosenzweig, M. R., & Wolpin, K. I. (2000). Natural 'natural experiments' in economics. *Journal of Economic Literature*, *38(4),* 827-874.
Roth, A. E. (1988). Laboratory experimentation in economics: A methodological overview. *Economic Journal*, *98,* 974-1031.
Roth, A. E. (1995). Introduction to experimental economics. In J. Kagel & A.E. Roth (Eds.), *Handbook of experimental economics* (pp. 3-110). Princeton: Princeton University Press.
Roth, A. E. (2015). Is experimental economics living up to its promise? In G. Frechette & A. Schotter (Eds.), *Handbook of experimental economic methodology*. Oxford: Oxford University Press.
Rubin, D. B. (1973). Matching to remove bias in observational studies. *Biometrics, 29,* 159-183.
Rubin, D. B. (1974). Estimating the causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology*, *66*, 688-701.
Rubin, D. B. (1980a). Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology, 66*, 688-701.

Rubin, D. B. (1980b). Randomization analysis of experimental data: the Fisher-randomization test; Comment. *Journal of the American Statistical Association*, *75*, 591-593.
Rubin, D. B. (1986). Statistics and causal inference; Comment: which ifs have causal answers. *Journal of the American Statistical Association*, *81*, 961962.
Russell Sage Foundation (2016): Behavioral Economics Roundtable, http://www.russellsage.org/research/behavioral-economics/behavioral-economics-roundtable.
Salecl, R. (2010). The tyranny of choice. London: Profile Books.
Shafir, E. (2012). *The behavioral foundations of public policy*. Princeton University Press.
Simon, H. A. (1955). A behavioural model of rational choice. *Quarterly Journal of Economics*, *69(1)*, 99-118.
Sims, C. A. (2010). But economics is not an experimental science. *Journal of Economic Perspectives*, *24(2)*, 59-68.
Slonim, R., Wang, C., Garbarino, E., & Merrett, D. (2013). Opting-in: participation bias in economic experiments. *Journal of Economic Behavior and Organization. 90*, 43-70.
Slovic, P., Finucane, M.L., Peters, E., & MacGregor, D.G. (2004). Risk as analysis and risk as feelings: some thoughts about affect, reason, risk, and rationality. *Risk Analysis*, *24(2)*, 1-12.
Smed, S., Jensen, J.D., & Denver, S. (2007). Socioeconomic characteristics and the effect of taxation as a health policy instrument. *Food Policy*, *32/5-6*, 624-639.
Smith, V. L. (2002). Method in experiment: rhetoric and reality. *Experimental Economics*, *5*, 91-110.
Smith, V. L. (2003). Constructivist and ecological rationality. *American Economic Review*, *93(2)*, 465-508.
Starmer, C. (1999a). Experimental economics: hard science or wasteful tinkering? *The Economic Journal*, *109*, F5-F15.
Starmer, C. (1999b). Experiments in economics: should we trust the dismal scientists in white coats? *Journal of Economic Methodology*, *6*, 1-30.
Sunstein, C. R. (2011). Empirically informed regulation. *University of Chicago Law Review 78*: 1349-1429.
Tanaka T., Camerer C., & Nguyen Q. (2010). Risk and time preferences: Linking experimental and household survey data from Vietnam. *American Economic Review*, *100(1)*, 557-571.
Thaler, R.H., & Benartzi, S. (2004). Save more tomorrow: using behavioral economics to increase employee saving, *Journal of Political Economy*, *112*, S164-S187.
Thaler, R.H., & Sunstein, C.R. (2003). Libertarian paternalism. *American Economics Review*. *93(2)*, 175-179.
Thaler, R.H., & Sunstein, C. (2008). *Nudge: Improving decisions about health, wealth and happiness.* New Haven, CT: Yale University Press.
Thøgersen, J. (1999). Spillover processes in the development of a sustainable consumption pattern. *Journal of Economic Psychology*, *20*, 53-81.
Trochim, W. (1984). *Research design for program dvaluation: the Regression-Discontinuity Approach.* Beverly Hills: Sage Publications.
Van Ittersum, K., Wansink, B. (2012). Plate size and color suggestibility: the Delboeuf Illusion's bias on serving and eating behavior. *Journal of Consumer Research*, 39(2), 215-228.
Volpp, K., Levy, A.G., Ash, D.A., Berlin, J.A., Murphy, J.J., Gomez, A., Sox, H., Zhu, J., & Lerman, C. (2006). A randomized controlled trial of financial incentives for smoking cessation. *Cancer Epidemiology, Biomarkers and Prevention*, *15(1)*, 12-18.
Volpp, K., John, L.K., Troxel, A.B., Norton, L., Fassbender, J., & Loewenstein, G. (2008). Financial incentive-based approaches for weight loss, *JAMA*, *300(22)*, 2631-7.

Volpp, K.G., Asch, D.A., Galvin, R, & Loewenstein, G. (2011). Redesigning employee health incentives: lessons from behavioural economics. *New England Journal of Medicine*, *365*, 388-390.
Wansink, B., Van Ittersum, K. (2006). The visual illusions of food: why plates, bowls, and spoons can bias consumption volume. *FASEB Journal*, 20(4), A618.
White House (2006): President Bush discusses Medicare prescription drug benefit. Press release, Washington DC, 9th May 2006. http:www.whitehouse.gov/news/releases/2006/05/20060509-5.html
World Development Report: Mind, Society and Behavior (2015). World Bank.
Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, *13*, 75-98.

Appendix: Figures

Figure 1. A taxonomy of policy formulation instruments


Conventional Economics
Behavioral Economics
Preferences
Nudges
1
Information
Behaviorally inspired information based policy
2
Financial Incentives
Behaviorally inspired incentives
3
Regulation (including taxes & subsidies)
4