bsp

A publication of the Behavioral Science & Policy Association

volume 1 issue 1
2015

spotlight topic

Challenging assumptions about behavioral policy

A publication of the
Behavioral Science &
Policy Association

volume 1 issue 1
2015

Craig R. Fox
Sim B. Sitkin
Editors

ISSN 2379-4607 (print)
ISSN 2379-4615 (online)
ISBN 978-0-8157-2508-4 (pbk)
ISBN 978-0-8157-2259-5 (epub)

Behavioral Science & Policy is a publication of the Behavioral Science & Policy Association, P.O. Box 51336, Durham, NC 27717-1336, and is published twice yearly with the Brookings Institution, 1775 Massachusetts Avenue, NW, Washington, DC 20036, and through the Brookings Institution Press.

For information on electronic and print subscriptions, contact the Behavioral Science & Policy Association, BSP@BehavioralPolicy.org

The journal may be accessed through OCLC (www.oclc.org) and Project Muse (http://muse/jhu.edu). Archived issues are also available through JSTOR (www.jstor.org).

table of contents

volume 1 issue 1
2015

editors' note

Welcome to the inaugural issue of *Behavioral Science & Policy.* We created BSP to help bridge a significant divide. The success of nearly all public and private sector policies hinges on the behavior of individuals, groups, and organizations. Today, such behaviors are better understood than ever thanks to a growing body of practical behavioral science research. However, policymakers often are unaware of behavioral science findings that may help them craft and execute more effective and efficient policies. In response, we want the pages of this journal to be a meeting ground of sorts: a place where scientists and non-scientists can encounter clearly described behavioral research that can be put into action.

Mission of BSP

By design, the scope of BSP is quite broad, with topics spanning health care, financial decisionmaking, energy and the environment, education and culture, justice and ethics, and work place practices. We will draw on a broad range of the social sciences, as is evident in this inaugural issue. These pages feature contributions from researchers with expertise in psychology, sociology, law, behavioral economics, organization science, decision science, and marketing. BSP is broad in its coverage because the problems to be addressed are diverse, and solutions can be found in a variety of behavioral disciplines.

This goal requires an approach that is unusual in academic publishing. All BSP articles go through a unique dual review, by disciplinary specialists for scientific rigor and also by policy specialists for practical implementability. In addition, all articles are edited by a team of professional writing editors to ensure that the language is both clear and engaging for non-expert readers. When needed, we post online Supplemental Material for those who wish to dig deeper into more technical aspects of the work. That material is indicated in the journal with a bracketed arrow.

This Issue

This first issue is representative of our vision for BSP. We are pleased to publish an outstanding set of contributions from leading scholars who have worked hard to make their work accessible to readers outside their fields. A subset of manuscripts is clustered into a Spotlight Topic section

that examines a specific theme in some depth, in this case, "Challenging Assumptions about Behavioral Policy."

Our opening essay discusses the importance of behavioral science for enhanced policy design and implementation, and illustrates various approaches to putting this work into practice. The essay also provides a more detailed account of our objectives for *Behavioral Science & Policy*. In particular, we discuss the importance of using policy challenges as a starting point and then asking what practical insights can be drawn from relevant behavioral science, rather than the more typical path of producing research findings in search of applications.

Our inaugural Spotlight Topic section includes four articles. Wilson and Juarez challenge the assumption that intuitively compelling policy initiatives can be presumed to be effective, and illustrate the importance of evidence-based program evaluation. Cialdini, Martin, and Goldstein challenge the notion that large policy effects require large interventions, and provide evidence that small (even costless) actions grounded in behavioral science research can pay big dividends. Sunstein challenges the point of view that providing individuals with default options is necessarily more paternalistic than requiring them to make an active choice. Instead, Sunstein suggests, people sometimes prefer the option of deferring technical decisions to experts and delegating trivial decisions to others. Thus, forcing individuals to choose may constrain rather than enhance individual free choice. In the final Spotlight paper, Loewenstein, Bryce, Hagmann, and Rajpal challenge the assumption that behavioral "nudges," such as strategic use of defaults, are only effective when kept secret. In fact, these authors report a study in which they explicitly inform participants that they have been assigned an arbitrary default (for advance medical directives). Surprisingly, disclosure does not greatly diminish the impact of the nudge.

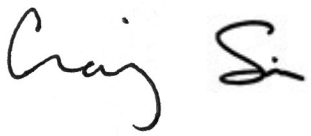
This issue also includes four regular articles. Goh, Pfeffer, and Zenios provide evidence that corporate executives concerned with their employees' health should attend to a number of workplace practices—including high job demands, low job control, and a perceived lack of fairness—that can produce more harm than the well-known threat of exposure to secondhand smoke. Knoll, Appelt, Johnson, and Westfall find that the most obvious approach to getting individuals to delay claiming retirement benefits (present information in a way that highlights benefits of claiming later) does not work. But a process intervention in which individuals are asked to think about the future before considering their current situation better persuades them to delay making retirement claims. Larrick, Soll, and Keeney identify four principles for developing better energy-use metrics to enhance consumer understanding and promote energy conservation. Finally, Manary, Staelin, Boulding, and Glickman provide a new analysis challenging the

idea that a hospital's responses to the demographic traits of individual patients, including their race, may explain disparities in quality of health care. Instead, it appears that this observation is driven by differences in insurance coverage among these groups. Hospitals serving larger numbers of patients with no insurance or with government insurance receive less revenue to pay for expenses such as wages, training, and equipment updates. In this case, the potential behavioral explanation does not appear to be correct; it may come down to simple economics.

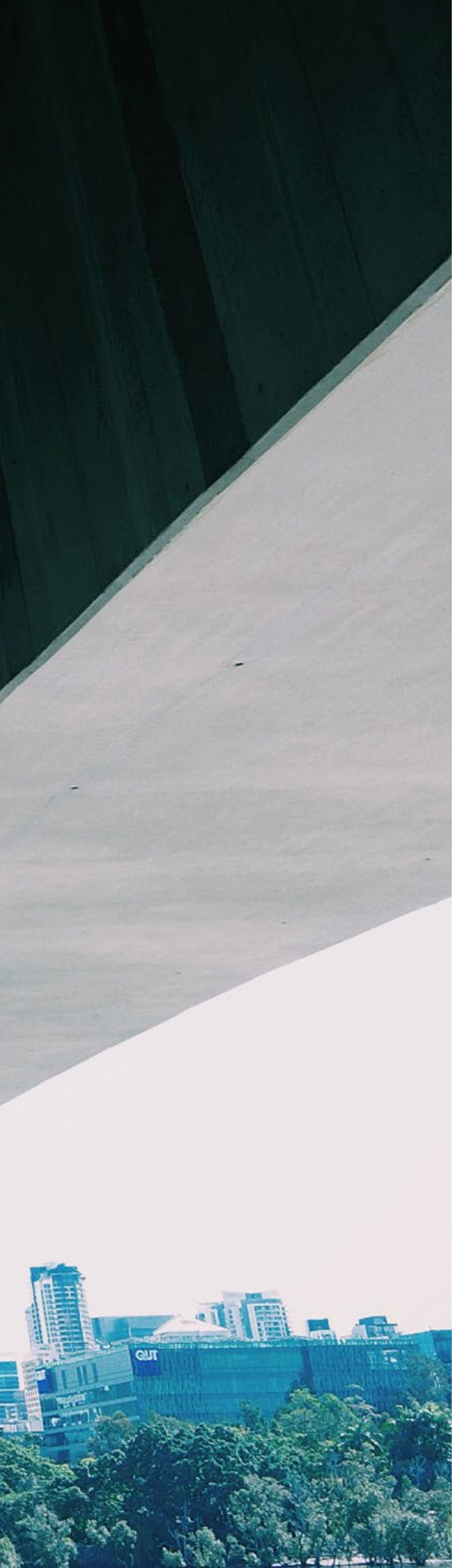
In Summary

This publication was created by the Behavioral Science & Policy Association in partnership with the Brookings Institution. The mission of BSPA is to foster dialog between social scientists, policymakers, and other practitioners in order to promote the application of rigorous empirical behavioral science in ways that serve the public interest. BSPA does not advance a particular agenda or political perspective.

We hope that each issue of BSP will provide timely and actionable insights that can enhance both public and private sector policies. We look forward to continuing to receive innovative policy solutions that are derived from cutting-edge behavioral science research. We also look forward to receiving from policy professionals suggestions of new policy challenges that may lend themselves to behavioral solutions. "Knowledge in the service of society" is an ideal that we believe should not merely be espoused but, also, actively pursued.



Craig R. Fox & Sim B. Sitkin
Founding Co-Editors



Essay

Bridging the divide between behavioral science & policy

Craig R. Fox & Sim B. Sitkin

abstract. Traditionally, neoclassical economics, which assumes that people rationally maximize their self-interest, has strongly influenced public and private sector policymaking and implementation. Today, policymakers increasingly appreciate the applicability of the behavioral sciences, which advance a more realistic and complex view of individual, group, and organizational behavior. In this article, we summarize differences between traditional economic and behavioral approaches to policy. We take stock of reasons economists have been so successful in influencing policy and examine cases in which behavioral scientists have had substantial impact. We emphasize the benefits of a problem-driven approach and point to ways to more effectively bridge the gap between behavioral science and policy, with the goal of increasing both supply of and demand for behavioral insights in policymaking and practice.

Better insight into human behavior by a county government official might have changed the course of world history. Late in the evening of November 7, 2000, as projections from the U.S. presidential election rolled in, it became apparent that the outcome would turn on which candidate carried Florida. The state initially was called by several news outlets for Vice President Al Gore, on the basis of exit polls. But in a stunning development, that call was flipped in favor of Texas Governor George W. Bush as the actual ballots were tallied.[1] The count proceeded through the early morning hours, resulting in a narrow margin of a few hundred votes for Bush that triggered an automatic machine recount. In the days that followed, intense attention focused on votes disallowed due to "hanging chads" on ballots that had not been properly punched. Weeks later, the U.S. Supreme Court halted a battle over the manual recount in a dramatic 5–4 decision. Bush would be certified the victor in Florida, and thus president-elect, by a mere 537 votes.

Less attention was paid to a news item that emerged right after the election: A number of voters in Palm Beach County claimed that they might have mistakenly voted for conservative commentator Pat Buchanan when they had intended to vote for Gore. The format of the ballot, they said, had confused them. The Palm Beach County ballot was designed by Theresa LePore, the supervisor of elections, who was a registered Democrat. On the Palm Beach County "butterfly ballot," candidate names appeared on facing pages, like butterfly wings, and votes were punched along a line between the pages (see Figure 1). LePore favored this format because it allowed for a larger print size that would be more readable to the county's large proportion of elderly voters.[2]

Ms. LePore unwittingly neglected an important behavioral principle long known to experimental psychologists: To minimize effort and mistakes, the response required (in this case, punching a hole in the center line) must be compatible with people's perception of the relevant stimulus (in this case, the ballot layout).[3,4] To illustrate this principle, consider a stove in which burners are aligned in a square but the burner controls are aligned in a straight line (see Figure 2, left panel). Most people have difficulty selecting the intended controls, and they make occasional errors. In contrast, if the controls are laid out in a square that mirrors the alignment of burners (see Figure 2, right panel), people tend to make fewer errors. In this case, the stimulus (the burner one wishes to light) better matches the response (the knob requiring turning).

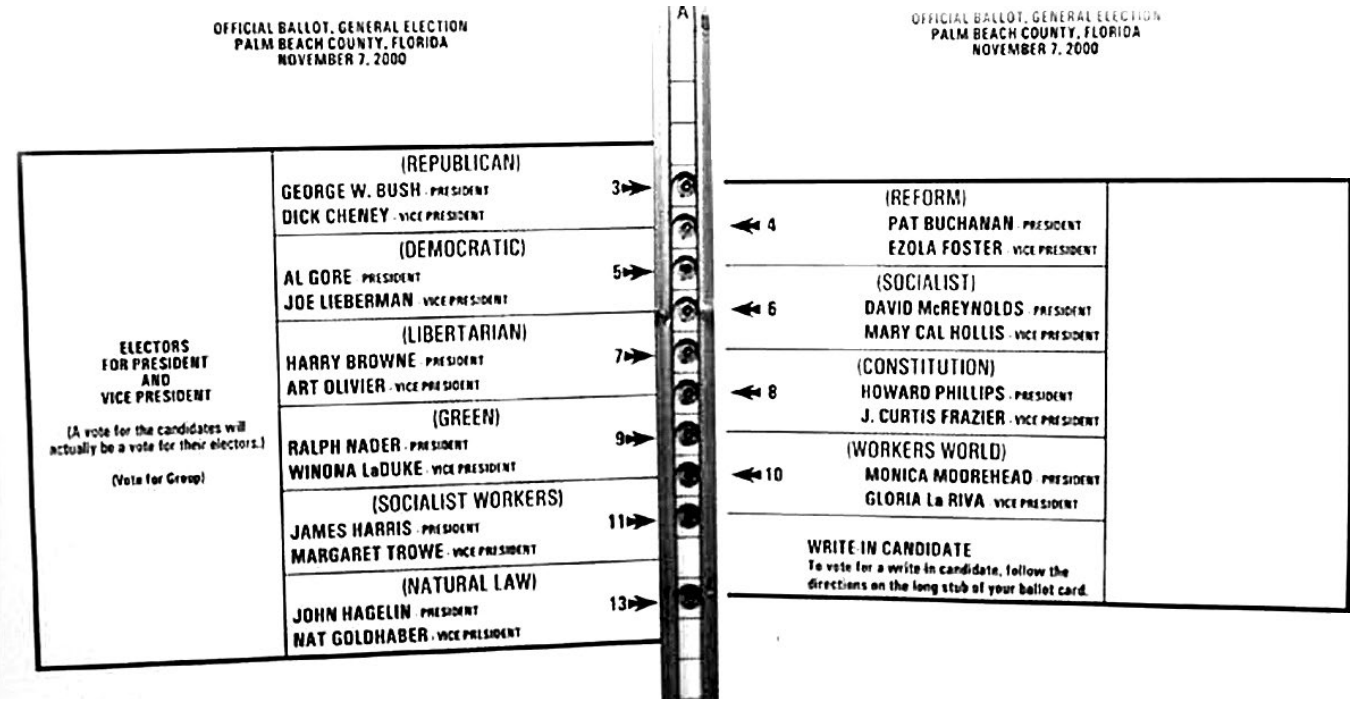
A close inspection of the butterfly ballot reveals an obvious incompatibility. Because Americans read left to right, many people would have perceived Gore as the second candidate on the ballot. But punching the second hole (No. 4) registered a vote for Buchanan. Meanwhile, because George Bush's name was listed at the top of the ballot and a vote for him required punching the top hole, no such incompatibility was in play, so no related errors should have occurred. Indeed, a careful analysis of the Florida vote in the 2000 presidential election shows that Buchanan received a much higher vote count than would be predicted from the votes for other candidates using well-established statistical models. In fact, the "overvote" for Buchanan in Palm Beach County (presumably, by intended Gore voters) was estimated to be at least 2,000 votes, roughly four times the vote gap between Bush and Gore in the official tally.[5] In short, had Ms. LePore been aware of the psychology of stimulus–response compatibility, she presumably would have selected a less confusing ballot design. In that case, for better or worse, Al Gore would almost certainly have been elected America's 43rd president.

It is no surprise that a county-level government official made a policy decision without considering a well-established principle from experimental psychology. Policymaking, in both the public and the private sectors, has been dominated by a worldview from neoclassical economics that assumes people and organizations maximize their self-interest. Under this *rational agent* view, it is natural to take for granted that given full information, clear instructions, and an incentive to pay attention, mistakes should be rare; systematic mistakes are unthinkable. Perhaps more surprising is the fact that behavioral science research has not been routinely consulted by policymakers, despite the abundance of policy-relevant insights it provides.

This state of affairs is improving. Interest in applied behavioral science has exploded in recent years, and the supply of applicable behavioral research has been increasing steadily. Unfortunately, most of this research fails to reach policymakers and practitioners in a

Fox, C. R., & Sitkin, S. B. Bridging the divide between behavioral science & policy. *Behavioral Science & Policy, 1*(1), pp.1–12.

Figure 1. Palm Beach County's 2000 butterfly ballot for U.S. president



useable format, and when behavioral insights do reach policymakers, it can be difficult for these professionals to assess the credibility of the research and act on it. In short, a stubborn gap persists between rigorous science and practical application.

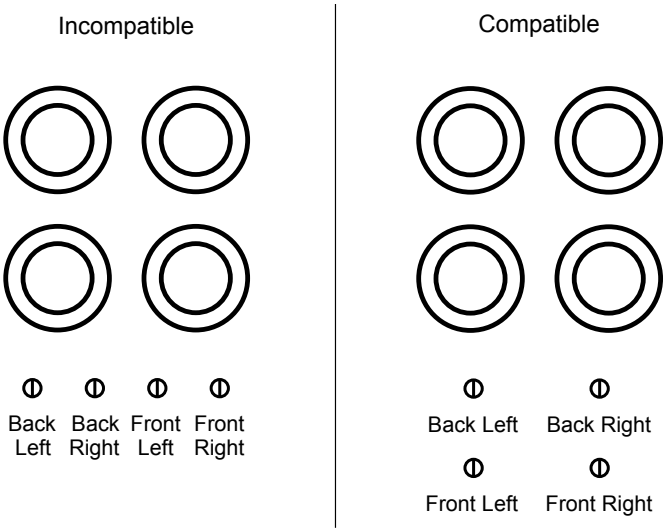
In this article, we explore the divide between behavioral science and policymaking. We begin by taking stock of differences between traditional and behavioral approaches to policymaking. We then examine what behavioral scientists can learn from (nonbehavioral) economists' relative success at influencing policy. We share case studies that illustrate different approaches that behavioral scientists have taken in recent years to successfully influence policies. Finally, we discuss ways to bridge the divide, thereby promoting more routine and judicious application of behavioral science by policymakers.

Traditional Versus Behavioral Approaches to Policymaking

According to the rational agent model, individuals, groups, and organizations are driven by an evenhanded evaluation of available information and the pursuit of self-interest. From this perspective, policymakers have three main tools for achieving their objectives: information, incentives, and regulation.

Information includes education programs, detailed documentation, and information campaigns (for example, warnings about the dangers of illicit drug use). The assumption behind these interventions is that accurate information will lead people to act appropriately.

Figure 2. Differences in compatibility between stove burners and controls



Adapted from *The Design of Everyday Things* (pp. 76–77), by D. Norman, 1988, New York, NY: Basic Books.

Incentives include financial rewards and punishments, tax credits, bonuses, grants, and subsidies (for example, a tax credit for installing solar panels). The assumption here is that proper incentives motivate individuals and organizations to behave in ways that are aligned with society's interests.

Regulation entails a mandate (for example, requiring a license to operate a plane or perform surgery) or a prohibition of a particular behavior (such as forbidding speeding on highways or limiting pollution from a factory). In some sense, regulations provide a special kind of (dis)incentive in the form of a legal sanction.

Although tools from neoclassical economics will always be critical to policymaking, they often neglect important insights about the actual behaviors of individuals, groups, and organizations. In recent decades, behavioral and social scientists have produced ample evidence that people and organizations routinely violate assumptions of the rational agent model, in systematic and predictable ways. First, individuals have a severely limited capacity to attend to, recall, and process information and therefore to choose optimally.[6] For instance, a careful study of older Americans choosing among prescription drug benefit plans under Medicare Part D (participants typically had more than 40 stand-alone drug plan options available to them) found that people selected plans that, on average, fell short of optimizing their welfare, by a substantial margin.[7,8] Second, behavior is strongly affected by how options are framed or labeled. For example, economic stimulus payments are more effective (that is, people spend more money) when those payments are described as a gain (for example, a "taxpayer bonus") than when described as a return to the status quo (for example, a "tax rebate").[9] Third, people are biased to stick with default options or the status quo, for example, when choosing health and retirement plans,[10] insurance policies,[11] flexible spending accounts,[12] and even medical advance directives.[13] People likewise tend to favor incumbent candidates,[14] current program initiatives,[15] and policies that happen to be labeled the status quo.[16] Fourth, people are heavily biased toward immediate rather than future consumption. This contributes, for example, to the tendency to undersave for retirement. It is interesting to note, though, that when people view photographs of themselves that have been artificially aged, they identify more with their future selves and put more money away for retirement.[17]

One response to such observations of irrationality is to apply traditional economic tools that attempt to enforce more rational decisionmaking. In this respect, behavioral research can serve an important role in identifying situations in which intuitive judgment and decisionmaking may fall short (for instance, scenarios in which the public tends to misperceive risks)[18,19] for which economic decision tools like cost–benefit analysis are especially helpful.[20] More important, behavioral scientists have begun to develop powerful new tools that complement traditional approaches to policymaking. These tools are derived from observations about how people actually behave rather than how rational agents ought to behave. Such efforts have surged since the publication of Thaler and Sunstein's book *Nudge*,[21] which advocates leveraging behavioral insights to design policies that promote desired behaviors while preserving freedom of choice. A number of edited volumes of behavioral policy insights from leading scholars have followed.[22–25]

Behavioral information tools leverage scientific insights concerning how individuals, groups, and organizations naturally process and act on information. Feedback presented in a concrete, understandable format can help people and organizations learn to improve their outcomes (as with new smart power meters in homes or performance feedback reviews in hospitals[26] or military units[27]) and make better decisions (for instance, when loan terms are expressed using the annual percentage rate as required by the Truth in Lending Act[28] or when calorie information is presented as a percentage of one's recommended snack budget[29]). Similarly, simple reminders can overcome people's natural forgetfulness and reduce the frequency of errors in surgery, firefighting, and flying aircraft.[30–32] Decisions are also influenced by the order in which options are encountered (for example, first candidates listed on ballots are more likely to be selected)[33] and how options are grouped (for instance, physicians are more likely to choose medications that are listed separately rather than clustered together on order lists).[34] Thus, policymakers can nudge citizens toward favored options by listing them on web pages and forms first and separately rather than later and grouped with other options.

Behavioral incentives leverage behavioral insights about motivation. For instance, a cornerstone of behavioral economics is *loss aversion,* the notion that people

are more sensitive to losses than to equivalent gains. Organizational incentive systems can therefore make use of the observation that the threat of losing a bonus is more motivating than the possibility of gaining an equivalent bonus. In a recent field experiment, one group of teachers received a bonus that would have to be returned (a potential loss) if their students' test scores did not increase while another group of teachers received the same bonus (a potential gain) only after scores increased. In fact, test scores substantially increased when the bonus was presented as a potential loss but not when it was presented as a potential gain.[35]

A behavioral perspective on incentives also recognizes that the impact of monetary payments and fines depends on how people subjectively interpret those interventions. For instance, a field experiment in a group of Israeli day care facilities found that introducing a small financial penalty for picking up children late actually *increased* the frequency of late pickups, presumably because many parents interpreted the fine as a price that they would gladly pay.[36] Thus, payments and fines may not be sufficient to induce desired behavior without careful consideration of how they are labeled, described, and interpreted.

Behavioral insights not only have implications for how to tailor traditional economic incentives such as payments and fines but also suggest powerful nonmonetary incentives. It is known, for example, that people are motivated by their needs to belong and fit in, compare favorably, and be seen by others in a positive light. Thus, social feedback and public accountability can be especially potent motivators. For example, health care providers reduce their excessive antibiotic prescribing when they are told how their performance compares with that of "best performers" in their region[37] or when a sign declaring their commitment to responsible antibiotic prescribing hangs in their clinic's waiting room.[38] In contrast, attempts to influence health care provider behaviors (including antibiotic prescribing) using expensive, traditional pay-for-performance interventions are not generally successful.[39]

Nudges are a form of soft paternalism that stops short of formal regulation. They involve designing a choice environment to facilitate desired behavior without prohibiting other options or significantly altering economic incentives.[21] The most studied tool in this category is the use of defaults. For instance, European countries with opt-out policies for organ donation (in which consent to be a donor is the default) have dramatically higher rates of consent (generally approaching 100%) than do countries with opt-in policies (whose rates of consent average around 15%).[40]

Well-designed nudges make it easy for people to make better decisions. Opening channels for desired behavior (for instance, providing a potential donor to a charity with a stamped and pre-addressed return envelope) can be extremely effective, well beyond what would be predicted by an economic cost–benefit analysis of the action.[41] For instance, in one study, children from low-income families were considerably more likely to attend college if their parents had been offered help in completing a streamlined college financial aid form while they were receiving free help with their tax form preparation.[42] Conversely, trivial obstacles to action can prove very effective in deterring undesirable behavior. For instance, secretaries consumed fewer chocolates when candy dishes were placed a few meters away from their desks than when candy dishes were placed on their desks.[43]

Beyond such tools, *rigorous empirical observation* of behavioral phenomena can identify public policy priorities and tools for most effectively addressing those priorities. Recent behavioral research has made advances in understanding a range of policy-relevant topics, from the measurement and causes of subjective well-being[44,45] to accuracy of eyewitness identification[46] to improving school attendance[47] and voter turnout[48] to the psychology of poverty[49,50] to the valuation of environmental goods.[51,52] Rigorous empirical evaluation can also help policymakers assess the effectiveness of current policies[53] and management practices.[24,54]

Learning from the Success of Economists in Influencing Policy

Behavioral scientists can learn several lessons from the unrivaled success of economists in influencing policy. We highlight three: Communicate simply, field test and quantify results, and occupy positions of influence.

Simplicity

Economists communicate a simple and intuitively compelling worldview that can be easily summed up: Actors pursue their rational self-interest. This simple model also provides clear and concrete prescriptions:

Provide information and it will be used; align incentives properly and particular behaviors will be promoted or discouraged; mandate or prohibit behaviors and desired effects will tend to follow.

In contrast, behavioral scientists usually emphasize that a multiplicity of factors tend to influence behavior, often interacting in ways that defy simple explanation. To have greater impact, behavioral scientists need to communicate their insights in ways that are easy to absorb and apply. This will naturally inspire greater credence and confidence from practitioners.[55]

Field Tested and Quantified

Economists value field data and quantify their results. Economists are less interested in identifying underlying causes of behavior than they are in predicting observable behavior, so they are less interested in self-reports of intentions and beliefs than they are in consequential behavior. It is important to note that economists also quantify the financial impact of their recommendations, and they tend to examine larger, systemic contexts (for instance, whether a shift in a default increases overall savings rather than merely shifting savings from one account to another).[56] Such analysis provides critical justification to policymakers. In the words of Nobel Laureate Daniel Kahneman (a psychologist by training), economists "speak the universal language of policy, which is money."[57]

In contrast, behavioral scientists tend to be more interested in identifying causes, subjective understanding and motives, and complex group and organizational interactions—topics best studied in controlled environments and using laboratory experiments. Although controlled environments may allow greater insight into mental processes underlying behavior, results do not always generalize to applied contexts. Thus, we assert that behavioral scientists should make use of in situ field experiments, analysis of archival data, and natural experiments, among other methods, and take pains to establish the validity of their conclusions in the relevant applied context. In addition, we suggest that behavioral scientists learn to quantify the larger (systemic and scalable) impact of their proposed interventions.

Positions of Influence

Economists have traditionally placed themselves in positions of influence. Since 1920, the nonprofit and nonpartisan National Bureau of Economic Research has been dedicated to supporting and disseminating "unbiased economic research . . . without policy recommendations . . . among public policymakers, business professionals, and the academic community."[58] The Council of Economic Advisors was founded in 1946, and budget offices of U.S. presidential administrations and Congress have relied on economists since 1921 and 1974, respectively. Think tanks populate their ranks with policy analysts who are most commonly trained in economics. Economists are routinely consulted on fiscal and monetary policies, as well as on education, health care, criminal justice, corporate innovation, and a host of other issues. Naturally, economics is particularly useful when answering questions of national interest, such as what to do in a recession, how to implement cost–benefit analysis, and how to design a market-based intervention.

In contrast, behavioral scientists have only recently begun assuming positions of influence on policy through new applied behavioral research organizations (such as ideas42), standing government advisory organizations (such as the British Behavioral Insights Team and the U.S. Social and Behavioral Sciences Team), and corporate behavioral science units (such as Google's People Analytics and Microsoft Research). Behavioral scientists are sometimes invited to serve as ad hoc advisors to various government agencies (such as the Food and Drug Administration and the Consumer Financial Protection Bureau). As behavioral scientists begin to occupy more positions in such organizations, this will increase their profile and enhance opportunities to demonstrate the utility of their work to policymakers and other practitioners. Many behavioral insights have been successfully implemented in the United Kingdom[59] and in the United States.[60] For example, in the United States, the mandate to disclose financial information to consumers in a form they can easily understand (Credit Card Accountability and Disclosure Act of 2009), the requirement that large employers automatically enroll employees in a health care plan (Affordable Care Act of 2010), and revisions to simplify choices available under Medicare Part D were all designed with behavioral science principles in mind.

Approaches Behavioral Scientists Have Taken to Impact Policy

Although the influence of behavioral science in policy is growing, thus far there have been few opportunities for the majority of behavioral scientists who work at universities and in nongovernment research organizations to directly influence policy with their original research. Success stories have been mostly limited to a small number of cases in which behavioral scientists have (a) exerted enormous personal effort and initiative to push their idea into practice, (b) aggressively promoted a research idea until it caught on, (c) partnered with industry to implement their idea, or (d) embedded themselves in an organization with connections to policymakers.

Personal Initiative (Save More Tomorrow)

Occasionally, entrepreneurial behavioral scientists have managed to find ways to put their scientific insights into practice through their own effort and initiative. For instance, University of California, Los Angeles, professor Shlomo Benartzi and University of Chicago professor Richard Thaler were concerned about Americans' low saving rate despite the ready availability of tax-deferred 401(k) saving plans in which employers often match employee contributions. In 1996, they conceived of the Save More Tomorrow (SMarT) program, with features that leverage three behavioral principles. First, participants commit in advance to escalate their 401(k) contributions in the future, which takes advantage of people's natural tendency to heavily discount future consumption relative to present consumption. Second, contributions increase with the first paycheck after each pay raise, which leverages the fact that people find it easier to forgo a gain (give up part of a pay raise) than to incur a loss (reduce disposable income). Third, employee contributions automatically escalate (unless the participant opts out) until the savings rate reaches a predetermined ceiling, which applies the observation that people are strongly biased to choose and stick with default options.

Convincing a company to implement the program required a great deal of persistence over a couple of years. However, the effort paid off: In the first application of Save More Tomorrow, average saving rates among participants who signed up increased from 3.5% to 13.6% in less than four years. Having proven the effectiveness of the program, Benartzi and Thaler looked for a well-known company to enhance its credibility, and they eventually signed up Philips Electronics, again with a successful outcome.

Results of these field experiments were published in a 1994 issue of the *Journal of Political Economy*[61] and subsequently picked up by the popular press. Benartzi and Thaler were soon invited to consult with members of Congress on the Pension Protection Act of 2006, which endorsed automatic enrollment and automatic savings escalation in 401(k) plans. Adoption increased sharply from there, and, by 2011, more than half of large American companies with 401(k) plans included automatic escalation. The nation's saving rate has increased by many billions of dollars per year because of this innovation.[62]

Building Buzz (the MPG Illusion)

Other researchers have sometimes managed to influence policy by actively courting attention for their research ideas. Duke University professors Richard Larrick and Jack Soll, for instance, noticed that the commonly reported metric for automobile mileage misleads consumers by focusing on efficiency (miles per gallon [MPG]) rather than consumption (gallons per hundred miles [GPHM]). In a series of simple experiments, Larrick and Soll demonstrated that people generally make better fuel-conserving choices when they are given GPHM information rather than MPG information.[63] The researchers published this work in the prestigious journal *Science* and worked with the journal and their university to cultivate media coverage.

As luck would have it, days before publication, U.S. gasoline prices hit $4 per gallon for the first time, making the topic especially newsworthy. Although Larrick and Soll found the ensuing attention gratifying, it appeared that many people did not properly understand the MPG illusion. To clarify their point, Larrick and Soll launched a website that featured a video, a blog, and an online GPHM calculator. *The New York Times Magazine* listed the GPHM solution in its "Year in Ideas" issue. Before long, this work gained the attention of the director of the Office of Information and Regulatory Affairs and others, who brought the idea of using GPHM to the U.S. Environmental Protection Agency and U.S. Department of Transportation. These agencies

ultimately took actions that modified window labels for new cars beginning in 2013 to include consumption metrics (GPHM, annual fuel cost, and savings over five years compared with the average new vehicle).[60]

Partnering with Industry (Opower)

Of course, successful behavioral solutions are not only implemented through the public sector: Sometimes policy challenges are taken up by private sector businesses. For instance, Arizona State University professor Robert Cialdini, California State University professor Wesley Schultz, and their students ran a study in which they leveraged the power of social norms to influence energy consumption behavior. They provided residents with feedback concerning their own and their neighbors' average energy usage (what is referred to as a *descriptive social norm*), along with suggestions for conserving energy, via personalized informational door hangers. Results were dramatic: "Energy hogs," who had consumed more energy than average during the baseline period, used much less energy the following month. However, there was also a boomerang effect in which "energy misers," who had consumed less energy than average during the baseline period, actually consumed more energy the following month. Fortunately, the researchers also included a condition in which feedback provided not only average usage information but also a reminder about desirable behavior (an *injunctive social norm*). This took the form of a handwritten smiley face if the family had consumed less energy than average and a frowning face if they had consumed more energy than average. This simple, cheap intervention led to reduced energy consumption by energy hogs as before and also kept energy misers from appreciably increasing their rates of consumption.[64] Results of the study were reported in a 2007 article in the journal *Psychological Science.*

Publication is where the story might have ended, as with most scientific research. However, as luck would have it, entrepreneurs Dan Yates and Alex Laskey had been brainstorming a new venture dedicated to helping consumers reduce their energy usage. In a conversation with Hewlett Foundation staff, Yates and Laskey were pointed to the work of Cialdini, Schultz, and their collaborators. Yates and Laskey saw an opportunity to partner with utility companies to use social norm feedback to help reduce energy consumption among their customers, and they invited Cialdini to join their team as chief scientist. Eventually, the Sacramento Municipal Utility District agreed to sponsor a pilot test in which some of its customers would be mailed social norm feedback and suggestions for conserving energy. The test succeeded in lowering average consumption by 2%–3% over the next few months. Further tests showed similar results, and the company rapidly expanded its operations.[65] Independent researchers verified that energy conservation in the field and at scale was substantial and persistent over time.[66] As of this writing, Opower serves more than 50 million customers of nearly 100 utilities worldwide, analyzing 40% of all residential energy consumption data in the United States,[67] and has a market capitalization in excess of $500 million.

Connected Organizations

The success of behavioral interventions has recently gained the attention of governments, and several behavioral scientists have had opportunities to collaborate with "nudge units" across the globe. The first such unit was the Behavioral Insights Team founded by U.K. Prime Minister David Cameron in 2010, which subsequently spun off into an independent company. Similar units have formed in the United States, Canada, and Europe, many at the provincial and municipal levels. International organizations are joining in as well: As of this writing, the World Bank is forming its own nudge unit, and projects in Australia and Singapore are underway. Meanwhile, research organizations such as ideas42, BE Works, Innovations for Poverty Action, the Center for Evidence-Based Management, and the Greater Good Science Center have begun to facilitate applied behavioral research. A diverse range of for-profit companies have also established behavioral units and appointed behavioral scientists to leadership positions—including Allianz, Capital One, Google, Kimberly-Clark, and Lowe's, among others—to run randomized controlled trials that test behavioral insights.

Bridging the Divide between Behavioral Science and Policy

The stories above are inspiring illustrations of how behavioral scientists who are resourceful, entrepreneurial, determined, and idealistic can successfully push

their ideas into policy and practice. However, the vast majority of rank-and-file scientists lack the resources, time, access, and incentives to directly influence policy decisions. Meanwhile, policymakers and practitioners are increasingly receptive to behavioral solutions but may not know how to discriminate good from bad behavioral science. A better way of bridging this divide between behavioral scientists and policymakers is urgently needed. The solution, we argue, requires behavioral scientists to rethink the way they approach policy applications of their work, and it requires a new vehicle for communicating their insights.

Rethinking the Approach

Behavioral scientists interested in having real-world impact typically begin by reflecting on consistent empirical findings across studies in their research area and then trying to generate relevant applications based on a superficial understanding of relevant policy areas. We assert that to have greater impact on policymakers and other practitioners, behavioral scientists must work harder to first learn what it is that practitioners need to know. This requires effort by behavioral scientists to study the relevant policy context—the institutional and resource constraints, key stakeholders, results of past policy initiatives, and so forth—before applying behavioral insights. In short, behavioral scientists will need to adopt a more problem-driven approach rather than merely searching for applications of their favorite theories.

This point was driven home to us by a story from David Schkade, a professor at the University of California, San Diego. In 2004, Schkade was named to a National Academy of Sciences panel that was tasked with helping to increase organ donation rates. Schkade thought immediately of aforementioned research showing the powerful effect of defaults on organ donation consent.[40] Thus, he saw an obvious solution to organ shortages: Switch from a regime in which donors must opt in (for example, by affirmatively indicating their preference to donate on their driver license) to one that requires people to either opt out (presume consent unless one explicitly objects) or at least make a more neutral forced choice (in which citizens must actively choose whether or not to be a donor to receive a driver's license).

As the panel deliberated, Schkade was surprised to learn that some states had already tried changing the choice regime, without success. For instance, in 2000, Virginia passed a law requiring that people applying for driver's licenses or identification cards indicate whether they were willing to be organ donors, using a system in which all individuals were asked to respond (the form also included an undecided category; this and a nonresponse were recorded as unwillingness to donate). The attempt backfired because of the unexpectedly high percentage of people who did not respond yes.[68,69]

As the expert panel discussed the issue further, Schkade learned that a much larger problem in organ donation was *yield management*. In 2004, approximately 13,000–14,000 Americans died each year in a manner that made them medically eligible to become donors. Fifty-nine different organ procurement organizations (OPOs) across the United States had conversion rates (percentage of medically eligible individuals who became donors in their service area) ranging from 34% to 78%.[68] The panel quickly realized that getting lower performing OPOs to adopt the best practices of the higher performing OPOs—getting them to, say, an average 75% conversion rate—would substantially address transplant needs for all major organs other than kidneys. Several factors were identified as contributing to variations in conversion rates: differences in how doctors and nurses approach families of potential donors about donation (family wishes are usually honored); timely communication and coordination between the hospitals where the potential donors are treated, the OPOs, and the transplant centers; the degree of testing of the donors before organs are accepted for transplant; and the speed with which transplant surgeons and their patients decide to accept an offered organ. Such factors, it turned out, provided better opportunities for increasing the number of transplanted organs each year. Because almost all of the identified factors involve behavioral issues, they provided new opportunities for behavioral interventions. Indeed, since the publication of the resulting National Academy of Sciences report, the average OPO conversion rate increased from 57% in 2004 to 73% in 2012.[70]

The main lesson here is that one cannot assume that even rigorously tested behavioral scientific results will work as well outside of the laboratory or in new contexts. Hidden factors in the new applied context may blunt or reverse the effects of even the

most robust behavioral patterns that have been found in other contexts (in the Virginia case, perhaps the uniquely emotional and moral nature of organ donation decisions made the forced choice regime seem coercive). Thus, behavioral science applications urgently require proofs of concept through new field tests where possible. Moreover, institutional constraints and contextual factors may render a particular behavioral insight less practical or less important than previously supposed, but they may also suggest new opportunities for application of behavioral insights.

A second important reason for field tests is to calibrate scientific insights to the domain of application. For instance, Sheena Iyengar and Mark Lepper famously documented *choice overload,* in which too many options can be debilitating. In their study, they found that customers of an upscale grocery store were much more likely to taste a sample of jam when a display table had 24 varieties available for sampling than when it had six varieties, but the customers were nevertheless much less likely to actually make a purchase from the 24-jam set.[71] Although findings such as this suggest that providing consumers with too many options can be counterproductive, increasing the number of options generally will provide consumers with a more attractive best option. The ideal number of options undoubtedly varies from context to context,[72] and prior research does not yet make predictions precise enough to be useful to policymakers. Field tests can therefore help behavioral scientists establish more specific recommendations that will likely have greater traction with policymakers.

Communicating Insights

Although a vast reservoir of useful behavioral science waits to be repurposed for specific applications, the kind of research required to accomplish this goal is typically not valued by high-profile academic journals. Most behavioral scientists working in universities and research institutes are under pressure to publish in top disciplinary journals that tend to require significant theoretical or methodological advances, often requiring authors to provide ample evidence of underlying causes of behavior. Many of these publications do not reward field research of naturally occurring behavior,[73] encourage no more than a perfunctory focus on practical implications of research, and usually serve a single behavioral discipline. There is therefore an urgent need for new high-profile outlets that publish thoughtful and rigorous applications of a wide range of behavioral sciences—and especially field tests of behavioral principles—to increase the supply of behavioral insights that are ready to be acted on.

On the demand side, although policymakers increasingly are open to rigorous and actionable behavioral insights, they do not see much research in a form that they can use. Traditional scientific journals that publish policy-relevant work tend to be written for experts, with all the technical details, jargon, and lengthy descriptions that experts expect but busy policymakers and practitioners cannot decipher easily. In addition, this work often comes across as naive to people creating and administering policy. Thus, new publications are needed that not only guarantee the disciplinary and methodological rigor of research but also deliver reality checks for scientists by incorporating policy professionals into the review process. Moreover, articles should be written in a clear and compelling way that is accessible to nonexpert readers. Only then will a large number of practitioners be interested in applying this work.

Figure 3. A problem-driven approach to behavioral policy

1. Identify timely problem.
2. Study context and history.
3. Apply scientifically grounded insights.
4. Test in relevant context.
5. Quantify impact and scalability.
6. Communicate simply and clearly.
7. Engage with policymakers on implementation.

Summing Up

In this article, we have observed that although insights from behavioral science are beginning to influence policy and practice, there remains a stubborn divide in which most behavioral scientists working in universities and research institutions fail to have much impact on policymakers. Taking stock of the success of economists and enterprising behavioral scientists, we argue for a problem-driven approach to behavioral policy research that we summarize in Figure

3. We hasten to add that a problem-driven approach to behavioral policy research can also inspire development of new behavioral theories. It is worth noting that the original theoretical research on stimulus–response compatibility, mentioned above in connection with the butterfly ballot, actually originated from applied problems faced by human-factors engineers in designing military-related systems in World War II.[74] The bridge between behavioral science and policy runs in both directions.

The success of public and private policies critically depends on the behavior of individuals, groups, and organizations. It should be natural that governments, businesses, and nonprofits apply the best available behavioral science when crafting policies. Almost a half century ago, social scientist Donald Campbell advanced his vision for an "experimenting society," in which public and private policy would be improved through experimentation and collaboration with social scientists.[75] It was impossible then to know how long it would take to build such a bridge between behavioral science and policy or if the bridge would succeed in carrying much traffic. Today, we are encouraged by both the increasing supply of rigorous and applicable behavioral science research and the increasing interest among policymakers and practitioners in actionable insights from this work. Both the infrastructure to test new behavioral policy insights in natural environments and the will to implement them are growing rapidly. To realize the vast potential of behavioral science to enhance policy, researchers and policymakers must meet in the middle, with behavioral researchers consulting practitioners in development of problem-driven research and with practitioners consulting researchers in the careful implementation of behavioral insights.

author affiliation

Fox, Anderson School of Management, Department of Psychology, and Geffen School of Medicine, University of California, Los Angeles; Sitkin, Fuqua School of Business, Duke University. Corresponding author's e-mail: cfox@anderson.ucla.edu

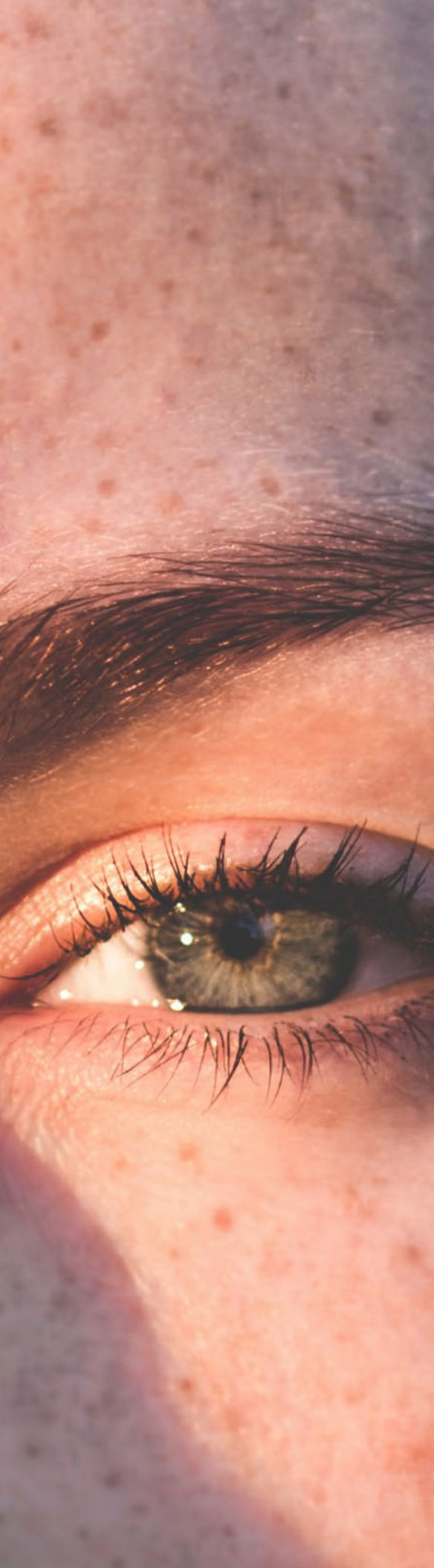
author note

We thank Shlomo Benartzi, Robert Cialdini, Richard Larrick, and David Schkade for sharing details of their case studies with us and Carsten Erner for assistance in preparing this article. We also thank Carol Graham, Jeffrey Pfeffer, Todd Rogers, Denise Rousseau, Cass Sunstein, and David Tannenbaum for helpful comments and suggestions.

References

1. Shepard, A. C. (2001, January/February). How they blew it. *American Journalism Review*. Retrieved from http://www.ajrarchive.org/
2. VanNatta, D., Jr., & Canedy, D. (2000, November 9). The 2000 elections: The Palm Beach ballot; Florida Democrats say ballot's design hurt Gore. *The New York Times.* Retrieved from http://www.nytimes.com
3. Fitts, P. M., & Seeger, C. M. (1953). S-R compatibility: Spatial characteristics of stimulus and response codes. *Journal of Experimental Psychology, 46,* 199–210.
4. Wickens, C. D. (1984). Processing resources in attention. In R. Parasuraman & D. R. Davies (Eds.), *Varieties of attention* (pp. 63–102). Orlando, FL: Academic Press.
5. Wand, J. N., Shotts, K. W., Sekhon, J. S., Mebane, W. R., Herron, M. C., & Brady, H. E. (2001). The butterfly did it: The aberrant vote for Buchanan in Palm Beach County, Florida. *American Political Science Review, 95,* 793–810.
6. Anderson, J. R. (2009). *Cognitive psychology and its implications* (7th ed.). New York, NY: Worth.
7. Abaluck, J., & Gruber, J. (2011). Choice inconsistencies among the elderly: Evidence from plan choice in the Medicare Part D program. *American Economic Review, 101,* 1180–1210.
8. Bhargava, S., Loewenstein, G., & Sydnor, J. (2015). *Do individuals make sensible health insurance decisions? Evidence from a menu with dominated options* (NBER Working Paper No. 21160). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w21160
9. Epley, N., & Gneezy, A. (2007). The framing of financial windfalls and implications for public policy. *Journal of Socio-Economics, 36,* 36–47.
10. Samuelson, W., & Zeckhauser, R. (1988). Status quo bias in decision making. *Journal of Risk and Uncertainty, 1,* 7–59.
11. Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty, 7,* 35–51.
12. Schweitzer, M., Hershey, J. C., & Asch, D. A. (1996). Individual choice in spending accounts: Can we rely on employees to choose well? *Medical Care, 34,* 583–593.
13. Halpern, S. D., Loewenstein, G., Volpp, K. G., Cooney, E., Vranas, K., Quill, C.M., . . . Bryce, C. (2013). Default options in advance directives influence how patients set goals for end-of-life care. *Health Affairs, 32,* 408–417.
14. Gelman, A., & King, G. (1990). Estimating incumbency advantage without bias. *American Journal of Political Science, 34,* 1142–1164.
15. Staw, B. M. (1976). Knee-deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance, 16,* 27–44.
16. Moshinsky, A., & Bar-Hillel, M. (2010). Loss aversion and status quo label bias. *Social Cognition, 28,* 191–204.
17. Hershfield, H. E., Goldstein, D. G., Sharpe, W. F., Fox, J., Yeykelis, L., Carstensen, L.L., & Bailenson, J. N. (2011). Increasing saving behavior through age-progressed renderings of the future self. *Journal of Marketing Research, 48*(SPL), 23–37.
18. Slovic, P. (2000). *The perception of risk.* London, United Kingdom: Routledge.
19. Slovic, P. (2010). *The feeling of risk: New perspectives on risk perception.* London, United Kingdom: Routledge.
20. Sunstein, C. R. (2012). If misfearing is the problem, is cost–benefit analysis the solution? In E. Shafir (Ed.), *The behavioral foundations of public policy* (pp. 231–244). Princeton, NJ: Princeton University Press.
21. Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.
22. Shafir, E. (Ed.). (2012). *The behavioral foundations of public policy*. Princeton, NJ: Princeton University Press.
23. Oliver, A. (Ed.). (2013). *Behavioural public policy*. Cambridge, United Kingdom: Cambridge University Press.
24. Rousseau, D. M. (Ed.). (2012). *The Oxford handbook of evidence-based management*. Oxford, United Kingdom: Oxford University Press.
25. Johnson, E. J., Shu, S. B., Dellaert, B. G. C., Fox, C. R., Goldstein, D. G., Häubl, G., . . . Weber, E. U. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters, 23,* 487–504.
26. Salas, E., Klein, C., King, H., Salisbury, N., Augenstein, J. S., Birnbach, D. J., . . . Upshaw, C. (2008). Debriefing medical teams: 12 evidence-based best practices and tips. *Joint Commission Journal on Quality and Patient Safety, 34,* 518–527.
27. Ellis, S., & Davidi, I. (2005). After-event reviews: Drawing lessons from successful and failed experience. *Journal of Applied Psychology, 90,* 857–871.
28. Stango, V., & Zinman, J. (2011). Fuzzy math, disclosure regulation, and market outcomes: Evidence from truth-in-lending reform. *Review of Financial Studies, 24,* 506–534.
29. Downs, J. S., Wisdom, J., & Loewenstein, G. (in press). Helping consumers use nutrition information: Effects of format and presentation. *American Journal of Health Economics.*
30. Gawande, A. (2009). *The checklist manifesto: How to get things right*. New York, NY: Metropolitan Books.
31. Hackmann, J. R. (2011). *Collaborative intelligence: Using teams to solve hard problems*. San Francisco, CA: Berrett-Koehler.
32. Weick, K. E., & Sutcliffe, K. M. (2001). *Managing the unexpected: Assuring high performance in an age of complexity*. San Francisco, CA: Jossey-Bass.
33. Miller, J. M., & Krosnick, J. A. (1998). The impact of candidate name order on election outcomes. *Public Opinion Quarterly, 62,* 291–330.
34. Tannenbaum, D., Doctor, J. N., Persell, S. D, Friedberg, M. W., Meeker, D., Friesema, E. M., . . . Fox, C. R. (2015). Nudging physician prescription decisions by partitioning the order set: Results of a vignette-based study. *Journal of General Internal Medicine, 30,* 298–304.
35. Fryer, R. G., Jr., Levitt, S. D., List, J., & Sadoff, S. (2012). *Enhancing the efficacy of teacher incentives through loss aversion: A field experiment* (NBER Working Paper No. 18237). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w18237
36. Gneezy, U., & Rustichini, A. (2000). A fine is a price. *Journal of Legal Studies, 29,* 1–17.
37. Meeker, D., Linder, J. A., Fox, C. R., Friedberg, M. W., Persell, S. D., Goldstein, N. J., . . . Doctor, J. N. (2015). *Behavioral interventions to curtail antibiotic overuse: A multisite randomized trial.* Unpublished manuscript, Leonard D. Schaeffer Center for Health Policy and Economics, University of Southern California, Los Angeles.
38. Meeker, D., Knight, T. K., Friedberg, M. W., Linder, J. A., Goldstein, N. J., Fox, C. R., . . . Doctor, J. N. (2014). Nudging guideline-concordant antibiotic prescribing: A randomized clinical trial. *JAMA Internal Medicine, 174,* 425–431.
39. Mullen, K. J., Frank, R. G., & Rosenthal, M. B. (2010). Can you get what you pay for? Pay-for-performance and the quality of healthcare providers. *Rand Journal of Economics, 41,* 64–91.
40. Johnson, E. J., & Goldstein, D. (2003, November 21). Do defaults save lives? *Science, 302,* 1338–1339.
41. Ross, L., & Nisbett, R. E. (2011). *The person and the situation: Perspectives of social psychology.* New York, NY: McGraw-Hill.
42. Bettinger, E. P., Long, B. T., Oreopoulos, P., & Sanbonmatsu, L. (2012). The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment. *Quarterly Journal of Economics, 127,* 1205–1242.

43. Wanskink, B., Painter, J. E., & Lee, Y. K. (2006). The office candy dish: Proximity's influence on estimated and actual consumption. *International Journal of Obesity, 30,* 871–875.
44. Dolan, P., Layard, R., & Metcalfe, R. (2011). *Measuring subjective wellbeing for public policy: Recommendations on measures* (Special Paper No. 23). London, United Kingdom: Office of National Statistics.
45. Kahneman, D., Diener, E., & Schwarz, N. (2003). *Well-being: The foundations of hedonic psychology.* New York, NY: Russell Sage Foundation.
46. Steblay, N. K., & Loftus, E. F. (2013). Eyewitness identification and the legal system. In E. Shafir (Ed.), *The behavioral foundations of public policy* (pp. 145–162). Princeton, NJ: Princeton University Press.
47. Epstein, J. L., & Sheldon, S. B. (2002). Present and accounted for: Improving student attendance through family and community involvement. *Journal of Education Research, 95,* 308–318.
48. Rogers, T., Fox, C. R., & Gerber, A. S. (2013). Rethinking why people vote: Voting as dynamic social expression. In E. Shafir (Ed.), *The behavioral foundations of public policy* (pp. 91–107). Princeton, NJ: Princeton University Press.
49. Bertrand, M., Mullainathan, S., & Shafir, E. (2004). A behavioral economics view of poverty. *American Economic Review, 94,* 419–423.
50. Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why having too little means so much.* New York, NY: Times Books.
51. Hausman, J. A. (1993). *Contingent valuation: A critical assessment.* Amsterdam, the Netherlands: Elsevier Science.
52. Kahneman, D., & Knetsch, J. L. (1992). Valuing public goods: The purchase of moral satisfaction. *Journal of Environmental Economics and Management, 22,* 57–70.
53. Haskins, R., & Margolis, G. (2014). *Show me the evidence: Obama's fight for rigor and results in social policy*. Washington, DC: Brookings Institution Press.
54. Pfeffer, J., & Sutton, R. I. (2006). *Hard facts, dangerous half-truths, and total nonsense.* Cambridge, MA: Harvard Business School Press.
55. Alter, A. L., & Oppenheimer, D. M. (2009). Uniting the tribes of fluency to form a metacognitive nation. *Personality and Social Psychology Review, 13,* 219–235.
56. Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T., & Olsen, T. (2012). *Active vs. passive decisions and crowdout in retirement savings accounts: Evidence from Denmark* (NBER Working Paper No. 18565). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w18565
57. Kahneman, D. (2013). Foreword. In E. Shafir (Ed.), *The behavioral foundations of public policy* (pp. vii–x). Princeton, NJ: Princeton University Press.
58. National Bureau of Economic Research. (n.d.). About the NBER. Retrieved May 15, 2015, from http://nber.org/info.html
59. Halpern, D. (2015). *Inside the Nudge Unit: How small changes can make a big difference.* London, United Kingdom: Allen.
60. Sunstein, C. R. (2013). *Simpler: The future of government.* New York, NY: Simon & Schuster.
61. Thaler, R. H., & Benartzi, S. (2004). Save More Tomorrow: Using behavioral economics to increase employee saving. *Journal of Political Economy*, *112*(S1), S164–S187.
62. Benartzi, S., & Thaler, R. H. (2013, March 8). Behavioral economics and the retirement savings crisis. *Science, 339,* 1152–1153.
63. Larrick, R. P., & Soll, J. B. (2008, June 20). The MPG illusion. *Science, 320,* 1593–1594.
64. Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science, 18,* 429–434.
65. Cuddy, A. J. C., Doherty, K. T., & Bos, M. W. (2012). *OPOWER: Increasing energy efficiency through normative influence. Part A* (Harvard Business Review Case Study No. 9-911-061). Cambridge, MA: Harvard University.
66. Allcott, H., & Rogers, T. (2014). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review, 104,* 3003–3037.
67. Opower. (2015). Opower surpasses 400 billion meter reads worldwide [Press release]. Retrieved from http://investor.opower.com/company/investors/press-releases/press-release-details/2015/Opower-Surpasses-400-Billion-Meter-Reads-Worldwide/default.aspx
68. Committee on Increasing Rates of Organ Donation, Childress, J. F., & Liverman, C. T. (Eds.). (2006). *Organ donation: Opportunities for action.* New York, NY: National Academies Press.
69. August, J. G. (2013). Modern models of organ donation: Challenging increases of federal power to save lives. *Hastings Constitutional Law Quarterly, 40,* 339–422.
70. U.S. Department of Health and Human Services. (2014). *OPTN/SRTR 2012 Annual Data Report*. Retrieved from http://srtr.transplant.hrsa.gov/annual_reports/2012/Default.aspx
71. Iyengar, S. S., & Lepper, M. R. (2000). When choice is demotivating: Can one desire too much of a good thing? *Journal of Personality and Social Psychology, 79,* 995–1006.
72. Shah, A. M., & Wolford, G. (2007). Buying behavior as a function of parametric variation of number of choices. *Psychological Science, 18,* 369–370.
73. Cialdini, R. B. (2009). We have to break up. *Perspectives on Psychological Science, 4,* 5–6.
74. Small, A. M. (1990). Foreword. In R. W. Proctor & T. G. Reeve (Eds.), *Stimulus–response compatibility: An integrated perspective* (pp. v–vi). Amsterdam, the Netherlands: Elsevier Science.
75. Campbell, D. T. (1969). Reforms as experiments. *American Psychologist, 24,* 409–429.



Review

Intuition is not evidence: Prescriptions for behavioral interventions from social psychology

Timothy D. Wilson & Lindsay P. Juarez

abstract. Many behavioral interventions are widely implemented before being adequately tested because they meet a commonsense criterion. Unfortunately, once these interventions are evaluated with randomized controlled trials (RCTs), many have been found to be ineffective or even to cause harm. Social psychologists take a different approach, using theories developed in the laboratory to design small-scale interventions that address a wide variety of behavioral and educational problems. Many of these interventions, tested with RCTs, have had large positive effects. The advantages of this approach are discussed, as are conditions necessary for scaling up any intervention to larger populations.

> Does anyone know if there's a scared straight program in Eagle Pass? My son is a total screw up and if he don't straighten out he's going to end up in jail or die from using drugs. Anyone please help!
>
> —Upset dad, Houston, TX[1]

It is no surprise that a concerned parent would want to enroll his or her misbehaving teenager in a so-called scared straight program. This type of dramatic intervention places at-risk youths in prisons where hardened inmates harangue them in an attempt to shock them out of a life of crime. An Academy Award–winning documentary film and a current television series on the A&E network celebrate this approach, adding to its popular appeal. It just makes sense: A parent might not be able to convince a wayward teen that his or her choices will have real consequences, but surely a prisoner serving a life sentence could. Who has more credibility than an inmate who experiences the horrors of prison on a daily basis? What harm could it do?

As it happens, a lot of harm. Scared straight programs not only don't work, they increase the likelihood that teenagers will commit crimes. Seven well-controlled studies that randomly assigned at-risk teens to participate in a scared straight program or a control group found that the kids who took part were, on average, 13% more likely to commit crimes in the following months.[2] Why scared straight programs increase criminal activity is not entirely clear. One possibility is that bringing at-risk kids together subjects them to negative peer influences;[3] another is that going to extreme lengths to convince kids to avoid criminal behavior conveys that there must be something attractive about those behaviors.[4] Whatever the reason, the data are clear: Scared straight programs increase criminal activity.

"Do No Harm"

The harmful effects of scared straight programs have been well documented, and many (although not all) states have eliminated such programs as a result. Unfortunately, this is but one example of a commonsense behavioral intervention that proved to be *iatrogenic*, a treatment that induces harm rather than healing.[5] Other examples include the Cambridge-Somerville Youth Study, a program designed to prevent at-risk youth from engaging in delinquent behaviors;[6] critical incident stress debriefing, an intervention designed to prevent posttraumatic stress in people who have experienced severe traumas; Dollar-a-Day programs, in which teen mothers receive money for each day they are not pregnant; and some diversity training programs (see reference 4 for a review of the evidence of these and other ineffective programs). At best, millions of dollars have been wasted on programs that have no effect. At worst, real harm has been done to thousands of unsuspecting people. For example, an estimated 6,500 teens in New Jersey alone have been induced to commit crimes as a result of a scared straight program.[4] Also, boys who were randomly assigned to take part in the Cambridge-Somerville Youth Study committed significantly more crimes and died an average of five years sooner than did boys assigned to the control group.[6]

Still another danger of these fiascos is that policymakers could lose faith in the abilities of social psychologists, whom they might assume helped create ineffective programs. "If that's the best they can do," a policymaker might conclude, "then the heck with them—let's turn it back over to the economists." To be fair, the aforementioned failures were designed and implemented not by research psychologists but by well-meaning practitioners who based their interventions on intuition and common sense. But common sense alone does not always translate to effective policy.

Psychological science does have tools needed to guide policymakers in this arena. For example, the field of social psychology, which involves the study of individuals' thoughts, feelings, and behaviors in a social context, can help policymakers address many important issues, including preventing child abuse, increasing voter turnout, and boosting educational achievement. This approach involves translating social psychological principles into real-world interventions and testing those interventions rigorously with small-scale randomized controlled trials (RCTs). As interventions are scaled up, they are tested experimentally to see when, where, and how they work. This approach, which has gathered considerable steam in recent years, has had some dramatic successes. Our goal here is to highlight the advantages and limits of this approach.

Social Psychological Interventions

Since its inception in the 1950s, the field of social psychology has investigated how social influence

Wilson, T. D., & Juarez, L. P. (2015). Intuition is not evidence: Prescriptions for behavioral interventions from social psychology. *Behavioral Science & Policy, 1*(1), pp. 13–20.

shapes human behavior and thought, primarily with the use of laboratory experiments. By examining people's behavior under carefully controlled conditions, social psychologists have learned a great deal about social cognition and social behavior. One of the most enduring lessons is the power of *construals,* the subjective ways individuals perceive and interpret the world around them. These subjective views often influence behavior more than objective facts do.[7–11] Hundreds of laboratory experiments, mostly with college student participants, have demonstrated the importance of this basic point, showing that people's behavior stems from their construals. Further, these construals sometimes go wrong, such that people adopt negative or pessimistic views that lead to maladaptive behaviors.

For example, Carol Dweck's studies of mindsets with elementary school, secondary school, and college students show that academic success often depends as much on people's theories about intelligence as on their actual intelligence.[12] People who view intelligence as a fixed trait are at a disadvantage, especially when they encounter obstacles. Poor grades can send them into a spiral of academic failure because they interpret those grades as a sign that they are not as smart as they thought they were, and so what is the point of trying? People who view intelligence as a set of skills that improves with practice often do better because they interpret setbacks as an indication that they need to try harder or seek help from others. By adopting these strategies, they do better.

Significantly, social psychologists have also found that construals can be changed, often with surprisingly subtle techniques, which we call *story-editing interventions.*[4] Increasingly, researchers are taking these principles out of the laboratory and transforming them into interventions to address a number of real-world problems, often with remarkable success.[4,13,14] Social scientists have long been concerned with addressing societal problems, of course, but the social psychological approach is distinctive in these ways:

- *The interventions are based on social psychological theory:* Rather than relying on common sense, social psychologists have developed interventions based on theoretical principles honed in decades of laboratory research. This has many advantages, not the least of which is that it has produced counterintuitive approaches that never otherwise would have been thought to work.[15]
- *Focus is on changing construals:* As noted, chief among these theoretical principles is that changing people's construals regarding themselves and their social world can have cascading effects that result in long-term changes in behavior.
- *The interventions start small and are tested with rigor:* Social psychologists begin by testing interventions in specific real-world contexts with tightly controlled experimental designs (RCTs), allowing for confident causal inference about the effects of the interventions. That is, rather than beginning by applying an intervention to large populations, they first test the intervention on a smaller scale to see if it works.

Editing Success Stories

The social psychological approach has been particularly successful in boosting academic achievement by helping students stay in school and improve their grades. In one study, researchers looked at whether a story-editing intervention could help first-year college students who were struggling academically. Often such students blame themselves, thinking that maybe they are not really "college material," and can be at risk of dropping out. These first-year participants were told that many students do poorly at first but then improve and were shown a video of third- and fourth-year students who reported that their grades had improved over time. Those who received this information (compared with a randomly assigned control group) achieved better grades over the next year and were less likely to drop out of college.[16,17] Other interventions, based on Dweck's work on growth mindsets, have improved academic performance in middle school, high school, and college students by communicating that intelligence is malleable rather than fixed.[18,19]

Social psychologists are taking aim at closing the academic achievement gap by overcoming *stereotype threat*, the widely observed fact that people are at risk of confirming negative stereotypes associated with groups they are associated with, including their ethnicity. Self-affirmation writing exercises can help. In one study, middle school students were asked to write about things they valued, such as their family and

friends or their faith. For low-performing African American students, this simple intervention produced better grades over the next two years.[20]

What about the fact that enrollment in high school science courses is declining in the United States? A recent study found that ninth-grade science students who wrote about the relevance of the science curriculum to their own lives increased their interest in science and improved their grades. This was especially true for students who had low expectations about how they would do in the course.[21] Another study that looked at test-taking anxiety in math and science courses found that high school and college students who spent 10 minutes writing about their fears right before taking an exam improved their performance.[22]

Education is not the only area to benefit from story-editing interventions. For example, this technique can dramatically reduce child abuse. Parents who abuse their children tend to blame the kids, with words such as "He's trying to provoke me" or "She's just being defiant." In one set of studies, home visitors helped to steer parents' interpretations away from such pejorative causes and toward more benign interpretations, such as the possibility that the baby was crying because he or she was hungry or tired. This simple intervention reduced child abuse by 85%.[23]

Story-editing interventions can make for happier marriages, too. Couples were asked to describe a recent major disagreement from the point of view of an impartial observer who had their best interests in mind. The couples who performed this writing exercise reported higher levels of marital satisfaction than did couples who did not do the exercise.[24]

These interventions can also increase voter turnout. When potential voters in California and New Jersey were contacted in a telephone survey, those who were asked how much they wanted to "be a voter" were more likely to vote than were those who were asked how much they wanted to "vote." The first wording led people to construe voting as a reflection of their self-image, motivating them to act in ways consistent with their image of engaged citizens.[25] Interventions that invoke social norms, namely, people's beliefs about what others are doing and what others approve of, have been shown to reduce home energy use[26] and reduce alcohol use on college campuses.[27] Simply informing people about where they stand in relation to what other people do and approve of helps them modify their behavior to conform to that norm.

Although these successful interventions used different approaches, they shared common features. Each targeted people's construals in a particular area, such as students' beliefs about why they were performing poorly academically. They each used a gentle push instead of a giant shove, with the assumption that this would lead to cascading changes in behavior over time. That is, rather than attempting to solve problems with massive, expensive, long-term programs, they changed people's construals with small, cheap, and short-term interventions. Each intervention was tested rigorously with an experimental design in one specific context, which gave researchers a good idea of how and why it worked. This is often not the case with massive "kitchen sink" interventions such as the Cambridge-Somerville Youth Study, which combined many treatments into one program. Even when these programs work, why they create positive change is not clear.

When we say that interventions should be tested with small samples, we do not mean underpowered samples. There is a healthy debate among methodologists as to the proper sample size in psychological research, with some arguing that many studies are underpowered.[28,29] We agree that intervention researchers should be concerned with statistical power and choose their sample sizes accordingly. But this can still be done while starting small, in the sense that an intervention is tested locally with one sample before being scaled up to a large population.

Scaling up and the Importance of Context

We do not mean to imply that the social psychological approach will solve every problem or will work in every context. Indeed, it would be naive to argue that every societal issue can be traced to people's construals—that it is all in people's heads—and that the crushing impact of societal factors such as poverty and racism can be ignored. Obviously, we should do all that we can to improve people's objective environments by addressing societal problems.

But there is often some latitude in how people interpret even dire situations, and the power of targeting these construals should be recognized. As an anecdotal

example, after asserting in a recent book[4] that "no one would argue that the cure for homelessness is to get homeless people to interpret their problem differently," one of us received an e-mail from a formerly homeless person, Becky Blanton. Ms. Blanton wrote,

> In 2006 I was living in the back of a 1975 Chevy van with a Rottweiler and a house cat in a Walmart Parking lot. Three years later, in 2009, I was the guest of Daniel Pink and was speaking at TED Global at Oxford University in the UK. . . . It was reframing and redirecting that got me off the streets. . . . Certainly having some benefits, financial, emotional, family, skill etc. matters, but where does the DRIVE to overcome come from?

As Ms. Blanton has described it, her drive came from learning that the late Tim Russert, who hosted NBC's *Meet the Press,* used an essay she wrote in his book about fathers. The news convinced her that she was a skilled writer despite her circumstances. Although there is a pressing need to improve people's objective circumstances, Ms. Blanton's e-mail is a poignant reminder that even for people in dire circumstances, construals matter.

And yet helping people change in positive ways by reshaping their construals can be complicated. It is vital to understand the interplay between people's construals and their environments. Social psychologists start small because they are keenly aware that the success of their interventions is often tied to the particular setting in which they are developed. As a result, interventions depend not only on changing people's construals but also on variables in their environments that support and nurture positive changes. These moderator variables are often unknown, and there is no guarantee that an intervention that worked in one setting, for example, a supportive school, will be as effective in another setting, such as a school with indifferent teachers. For example, consider the study[20] that found that African American middle school students earned better grades after writing essays about what they personally valued. This study took place in a supportive middle school with responsive teachers, and the same intervention might prove to be useless in an overcrowded school with a less supportive climate.

At this point, policymakers might again throw up their hands and say, "Are you saying that just because an intervention works in one school or community means that I can't use it elsewhere? Of what use are these studies to me if I can't implement their findings in other settings?" This is an excellent question to which we suggest two answers. First, we hope it is clear why it is dangerous to start big by applying a program broadly without testing it or understanding when and how it works. Doing so has led to massive failures that damaged people's lives, such as in the case of scared straight programs. Second, even if it is not certain that the findings from one study will generalize to a different setting, they provide a place to start. The key is to continue to test interventions as they are scaled up to new settings, with randomly assigned control groups, rather than assuming that they will work everywhere. That is the way to discover both how to effectively generalize an intervention and which variables moderate its success. In short, policymakers should partner with researchers who embrace the motto "Our work is never done" when it comes to testing and refining interventions (see references 30 and 31 for excellent discussion of the issues with scaling up).

There are exciting efforts in this direction. For example, researchers at Stanford University have developed a website that can be used to test self-affirmation and mindset interventions in any school or university in the United States (http://www.perts.net). Students sign on to the website at individual computers and are randomly assigned to receive treatment or control interventions; the schools agree to give the researchers anonymized data on the students' subsequent academic performance. Thousands of high school and college students have participated in studies through this website, and as a result, several effective ways of improving student performance have been discovered.[19]

Unfortunately, these lessons about continuing to test interventions when scaling up have not been learned in all quarters. Consider the Comprehensive Soldier Fitness program (now known as CSF2). After years of multiple deployments to Iraq and Afghanistan, U.S. troops have been experiencing record numbers of suicides, members succumbing to alcohol and drug abuse, and cases of posttraumatic stress disorder, among other signs of psychological stress. In

response, the U.S. Army rolled out a program intended to increase psychological resilience in soldiers and their families.[32] Unfortunately, the program was implemented as a mandatory program for all troops, with no control groups. The positive psychology studies on which the intervention was based were conducted with college students and school children. It is quite a leap to assume that the intervention would operate in the same way in a quite different population that has experienced much more severe life stressors, such as combat. By failing to include a randomly assigned control group, the U.S. Army and the researchers involved in this project missed a golden opportunity to find out whether the intervention works in this important setting, has no effect, or does harm.[33–35]

It is tempting when faced with an urgent large-scale need to forgo the approach we recommend here. Some rightly argue that millions of people are suffering every day from hunger, homelessness, and discrimination and they need to be helped today, not after academics in ivory towers conduct lengthy studies. We sympathize with this point of view. Many people need immediate help, and we are certainly not recommending that all aid be suspended until RCTs are conducted.

In many cases, however, it is possible to intervene and to test an intervention at the same time. People could be randomly assigned to different treatments to see which ones work best, or researchers could deliver a treatment to a relatively large group of people while designating a smaller, randomly chosen group of people to a no-treatment control condition.

This raises obvious ethical issues: Do we as researchers have the right to withhold treatment from some people on the basis of a coin toss? This is unethical only if we know for sure that the treatment is effective. One could make an equally compelling argument that it is unethical to deliver a treatment that has not been evaluated and might do more harm than good (for example, scared straight programs). Ethicists have no problem with withholding experimental treatments in the medical domain; it is standard practice to test a new cancer treatment, for example, by randomly assigning some patients to get it and others to a control group that does not. There is no reason to have different standards with behavioral treatments that have unknown effects.

One way to maintain research protocols while serving as many people as possible is to use a wait-list design. Imagine, for example, that a new after-school mentoring and tutoring program has been developed to help teens at risk of dropping out of school. Suppose further that there are 400 students in the school district who are eligible for the program but that there is funding to accommodate only 200. Many administrators would solve this by picking the 200 neediest kids. A better approach would be to randomly assign half to the program and the other half to a wait list and track the academic achievement of both groups.[36] If the program works—if those in the program do better than those on the wait list—then the program can be expanded to include the others. If the program doesn't work, then a valuable lesson has been learned, and its designers can try something new.

Some may argue that the gold standard of scientific tests of interventions—an RCT—is not always workable in the field. Educators designing a new charter school, for example, might find it difficult to randomly assign students to attend the school. Our sense, however, is that researchers and policymakers often give up too readily and that, with persistence and cleverness, experiments often can be conducted. In the case in which a school system uses a lottery to assign students to charter schools, researchers can compare the enrolled students with those who lost the lottery.[37,38] Another example of creativity in designating control groups in the field comes from studies designed to test whether radio soap operas could alleviate prejudice and conflict in Rwanda and the Democratic Republic of the Congo. The researchers created control groups by broadcasting the programs to randomly chosen areas of the countries or randomly chosen villages.[39,40]

There is no denying that many RCTs can be difficult, expensive, and time-consuming. But the costs of not vetting interventions with experimental tests must be considered, including the millions of dollars wasted on ineffective programs and the human cost of doing more harm than good. Understanding the importance of testing interventions with RCTs and then continuing to test their effectiveness when scaling up will, we hope, produce more discerning consumers and, crucially, more effective policymakers.

Recommendations for Policymakers

We close with a simple recommendation for increased partnerships between social psychological researchers and policymakers. Many social psychologists are keen on testing their theoretical ideas in real-world settings, but because there are practical barriers to gaining the trust and cooperation of practitioners, they often lack entry into those settings. Further, because they were trained in the ivory tower, social psychologists may lack a full understanding of the nuances of applied problems and the difficulties practitioners face in addressing them. Each would benefit greatly from the expertise of the other. We hope that practitioners and policymakers will come to appreciate the power and potential of the social psychological approach and be open to collaborations with researchers who bring to the table theoretical expertise and methodological rigor. Together, they can form a powerful team with the potential to make giant strides in solving a broad range of social and behavioral problems.

author affiliation

Wilson and Juarez, Department of Psychology, University of Virginia. Corresponding author's e-mail: tdw@virginia.edu

author note

The writing of this article was supported in part by National Science Foundation Grant SES-0951779.

References

1. Upset dad. (2013, January 5). Does anyone know if there's a scared straight program in Eagle Pass? [Online forum comment]. Retrieved from http://www.topix.com/forum/city/eagle-pass-tx/T6U00R1BNDTRB746V
2. Petrosino, A., Turpin-Petrosino, C., & Finckenauer, J. O. (2000). Well-meaning programs can have harmful effects! Lessons from experiments of programs such as scared straight. *Crime & Delinquency, 46,* 354–379. http://dx.doi.org/10.1177/0011128700046003006
3. Dishion, T. J., McCord, J., & Poulin, F. (1999). When interventions harm: Peer groups and problem behavior. *American Psychologist, 54,* 755–764. http://dx.doi.org/10.1037/0003-066X.54.9.755
4. Wilson, T. D. (2011). *Redirect: The surprising new science of psychological change.* New York, NY: Little, Brown.
5. Lilienfeld, S. O. (2007). Psychological treatments that cause harm. *Perspectives on Psychological Science, 2,* 53–69. http://dx.doi.org/10.1111/j.1745-6916.2007.00029.x
6. McCord, J. (2003). Cures that harm: Unanticipated outcomes of crime prevention programs. *Annals of the American Academy of Political and Social Science, 587,* 16–30. http://dx.doi.org/10.1177/0002716202250781
7. Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 6, pp. 1–62). New York, NY: Academic Press.
8. Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: The attribution process in social psychology. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 2, pp. 219–266). New York, NY: Academic Press.
9. Heider, F. (1958). *The psychology of interpersonal relations.* New York, NY: Wiley.
10. Kelley, H. H. (1967). Attribution theory in social psychology. In D. Levine (Ed.), *Nebraska Symposium on Motivation* (Vol. 15, pp. 192–238). Lincoln: University of Nebraska Press.
11. Ross, L. (1977). The intuitive psychologist and his shortcomings: Distortions in the attribution process. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 10, pp. 173–220). Orlando, FL: Academic Press.
12. Dweck, C. S. (2006). *Mindset: The new psychology of success*. New York, NY: Random House.
13. Walton, G. M. (2014). The new science of wise interventions. *Current Directions in Psychological Science, 23,* 73–82. http://dx.doi.org/10.1177/0963721413512856
14. Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81,* 267–301. http://dx.doi.org/10.3102/0034654311405999
15. Deaton, A. (2010). Instruments, randomization, and learning about development. *Journal of Economic Literature, 48,* 424–455. http://dx.doi.org/10.1257/jel.48.2.424
16. Wilson, T. D., & Linville, P. W. (1982). Improving the academic performance of college freshmen: Attribution therapy revisited. *Journal of Personality and Social Psychology, 42,* 367–376.
17. Wilson, T. D., Damiani, M., & Shelton, N. (2002). Improving the academic performance of college students with brief attributional interventions. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 88–108). San Diego, CA: Academic Press.
18. Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development, 78,* 246–263.
19. Yeager, D. S., Paunesku, D., Walton, G. M., & Dweck, C. S. (2013). *How can we instill productive mindsets at scale? A review of the evidence and an initial R&D agenda.* Unpublished

manuscript, Stanford University, Stanford, CA.
20. Cohen, G. L., Garcia, J., Purdie-Vaughns, V., Apfel, N., & Brzustoski, P. (2009, April 17). Recursive processes in self-affirmation: Intervening to close the achievement gap. *Science, 324,* 400–403. http://dx.doi.org/10.1126/science.1170769
21. Hulleman, C. S., & Harackiewicz, J. M. (2009, December 4). Promoting interest and performance in high school science classes. *Science, 326,* 1410–1412. http://dx.doi.org/10.1126/science.1177067
22. Ramirez, G., & Beilock, S. L. (2011, January 14). Writing about testing worries boosts exam performance in the classroom. *Science, 331,* 211–213. http://dx.doi.org/10.1126/science.1199427
23. Bugental, D. B., Beaulieu, D. A., & Silbert-Geiger, A. (2010). Increases in parental investment and child health as a result of an early intervention. *Journal of Experimental Child Psychology, 106,* 30–40. http://dx.doi.org/10.1016/j.jecp.2009.10.004
24. Finkel, E. J., Slotter, E. B., Luchies, L. B., Walton, G. M., & Gross, J. J. (2013). A brief intervention to promote conflict reappraisal preserves marital quality over time. *Psychological Science, 24,* 1595–1601.
25. Bryan, C. J., Walton, G. M, Rogers, T., & Dweck, C. S. (2011). Motivating voter turnout by invoking the self. *PNAS: Proceedings of the National Academy of Sciences, USA, 108,* 12653–12656. http://dx.doi.org/10.1073/pnas.1103343108
26. Cialdini, R. B. (2012). The focus theory of normative conduct. In P. van Lange, A. Kruglanski, & T. Higgins (Eds.), *Handbook of theories of social psychology* (pp. 295–312). London, United Kingdom: Sage.
27. DeJong, W., Schneider, S. K., Towvim, L. G., Murphy, M. J., Doerr, E. E., Simonsen, N. R., . . . Scribner, R. (2006). A multisite randomized trial of social norms marketing campaigns to reduce college student drinking. *Journal of Studies on Alcohol, 67,* 868–879.
28. Button, K. S., Ioannidis, J. P. A., Mokrysz, C., Nosek, B. A., Flint, J., Robinson, E. S. J., & Munafò, M. R. (2013). Power failure: Why small sample size undermines the reliability of neuroscience. *Nature Reviews Neuroscience, 14,* 365–376. http://dx.doi.org/10.1038/nrn3475
29. Jager, L. R., & Leek, J. T. (2014). An estimate of the science-wise false discovery rate and application to the top medical literature. *Biostatistics, 15,* 1–12. http://dx.doi.org/10.1093/biostatistics/kxt007
30. Cohen, G. L. (2011, October 14). Social psychology and social change. *Science, 334,* 178–179. http://dx.doi.org/10.1126/science.1212887
31. Evans, S. H., & Clarke, P. (2011). Disseminating orphan innovations. *Stanford Social Innovation Review, 9*(1), 42–47.
32. Reivich, K. J., Seligman, M. E. P., & McBride, S. (2011). Master resilience training in the U.S. Army. *American Psychologist, 66,* 25–34. http://dx.doi.org/10.1037/a0021897
33. Eidelson, R., Pilisuk, M., & Soldz, S. (2011). The dark side of Comprehensive Soldier Fitness. *American Psychologist, 66,* 643–644. http://dx.doi.org/10.1037/a0025272
34. Smith, S. L. (2013). Could Comprehensive Soldier Fitness have iatrogenic consequences? A commentary. *Journal of Behavioral Health Services & Research, 40,* 242–246. http://dx.doi.org/10.1007/s11414-012-9302-2
35. Steenkamp, M. M., Nash, W. P., & Litz, B. T. (2013). Post-traumatic stress disorder: Review of the Comprehensive Soldier Fitness program. *American Journal of Preventive Medicine, 44,* 507–512. http://dx.doi.org/10.1016/j.amepre.2013.01.013
36. Schreier, H. M. C., Schonert-Reichl, K. A., & Chen, E. (2013). Effect of volunteering on risk factors for cardiovascular disease in adolescents: A randomized controlled trial. *JAMA: Pediatrics, 167,* 327–332. http://dx.doi.org/10.1001/jamapediatrics.2013.1100
37. Dobbie, W., & Fryer, R. G., Jr. (2010). *Are high-quality schools enough to increase achievement among the poor? Evidence from the Harlem Children's Zone.* Unpublished manuscript. Retrieved February 18, 2014, from http://scholar.harvard.edu/files/fryer/files/hcz_nov_2010.pdf
38. Tuttle, C. C., Gill, B., Gleason, P., Knechtel, V., Nichols-Barrer, I., & Resch, A. (2013). *KIPP middle schools: Impacts on achievement and other outcomes* (Mathematica Policy Research No. 06441.910). Retrieved February 17, 2014, from KIPP Foundation website: http://www.kipp.org/files/dmfile/KIPP_Middle_Schools_Impact_on_Achievement_and_Other_Outcomes1.pdf
39. Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology, 96,* 574–587. http://dx.doi.org/10.1037/a0011989
40. Paluck, E. L. (2010). Is it better not to talk? Group polarization, extended contact, and perspective-taking in eastern Democratic Republic of Congo. *Personality and Social Psychology Bulletin, 36,* 1170–1185.


mt
55

Review

Small behavioral science–informed changes can produce large policy-relevant effects

Robert B. Cialdini, Steve J. Martin, & Noah J. Goldstein

abstract. Policymakers traditionally have relied upon education, economic incentives, and legal sanctions to influence behavior and effect change for the public good. But recent research in the behavioral sciences points to an exciting new approach that is highly effective and cost-efficient. By leveraging one or more of three simple yet powerful human motivations, small changes in reframing motivational context can lead to significant and policy-relevant changes in behaviors.

There is a story the late Lord Grade of Elstree often told about a young man who once entered his office seeking employ. Puffing on his fifth Havana of the morning, the British television impresario stared intently at the applicant for a few minutes before picking up a large jug of water and placing it on the desk that divided them. "Young man, I have been told that you are quite the persuader. So, sell me that jug of water."

Undaunted, the man rose from his chair, reached for the overflowing wastepaper basket beside Lord Grade's desk, and placed it next to the jug of water. He calmly lit a match, dropped it into the basket of discarded papers, and waited for the flames to build to an impressive (and no doubt anxiety-raising) level. He then turned to his potential employer and asked, "How much will you give me for this jug of water?"

The story is not only entertaining. It is also instructive, particularly for policymakers and public officials, whose success depends on influencing and changing behaviors. To make the sale, the young man persuaded his prospective employer not by changing a specific feature of the jug or by introducing a monetary incentive but by changing the psychological environment in which the jug of water was viewed. It was this shift in *motivational context* that caused Lord Grade's desire to purchase the jug of water to mushroom, rather like the flames spewing from the basket.

Small Shifts in Motivational Context

Traditionally, policymakers and leaders have relied upon education, economic incentives, and legal sanctions to influence behavior and effect change for the public good. Today, they have at hand a number of relatively new tools, developed and tested by behavioral scientists. For example, researchers have demonstrated the power of appeals to strong emotions such as fear, disgust, and sadness.[1–3] Likewise, behavioral scientists now know how to harness the enormous power of defaults, in which people are automatically included in a program unless they opt out. For example, simply setting participation as the default can increase the

Cialdini, R. B., Martin, S. J., & Goldstein, N. J. (2015). Small behavioral science–informed changes can produce large policy-relevant effects. *Behavioral Science & Policy, 1*(1), pp. 21–27.

number of people who become organ donors or the amount of money saved for retirement.[4–6]

In this review, we focus on another set of potent tools for policymakers that leverage certain fundamental human motivations: the desires to make accurate decisions, to affiliate with and gain the approval of others, and to see oneself in a positive light.[7,8] We look at these three fundamental motivations in particular because they underlie a large portion of the approaches, strategies, and tactics that have been scientifically demonstrated to change behaviors. Because these motivations are so deeply ingrained, policymakers can trigger them easily, often through small, costless changes in appeals.

As a team of behavioral scientists who study both the theory and the practice of persuasion-driven change,[9,10] we have been fascinated by how breathtakingly slight the changes in a message can be to engage one of these basic motivations and generate big behavioral effects. Equally remarkable to us is how people can be largely unaware about the extent to which these basic motivations affect their choices. For example, in one set of studies,[11] homeowners were asked how much four different potential reasons for conserving energy would motivate them to reduce their own overall home energy consumption: Conserving energy helps the environment, conserving energy protects future generations, conserving energy saves you money, or many of your neighbors are already conserving energy. The homeowners resoundingly rated the last of these reasons—the actions of their neighbors—as having the least influence on their own behavior. Yet when the homeowners later received one of these four messages urging them to conserve energy, only the one describing neighbors' conservation efforts significantly reduced power usage. Thus, a small shift in messaging to activate the motive of aligning one's conduct with that of one's peers had a potent but underappreciated impact. The message that most people reported would have the greatest motivational effect on them to conserve energy—conserving energy helps the environment—had hardly any effect at all.

Policymakers have two additional reasons to use small shifts in persuasive messaging beyond the outsized effects from some small changes. First, such shifts are likely to be cost-effective. Very often, they require only slight changes in the wording of an appeal. No additional program resources, procedures, or personnel are needed. Second, precisely because the adjustments are small, they are more likely to be embraced by program staff and implemented as planned.

Accuracy Motivation

The first motivation we examine is what we call the *accuracy motivation*. Put simply, people are motivated to be accurate in their perceptions, decisions, and behaviors.[7,12–15] To respond correctly (and therefore advantageously) to opportunities and potential threats in their environments, people must have an accurate perception of reality. Otherwise, they risk wasting their time, effort, or other important resources.

The accuracy motivation is perhaps most psychologically prominent in times of uncertainty, when individuals are struggling to understand the context, make the right decision, and travel down the best behavioral path.[16,17] Much research has documented the potent force of *social proof*[18]—the idea that if many similar others are acting or have been acting in a particular way within a situation, it is likely to represent a good choice.[19–21]

Indeed, not only humans are influenced by the pulling power of the crowd. So fundamental is the tendency to do what others are doing that even organisms with little to no brain cortex are subject to its force. Birds flock, cattle herd, fish school, and social insects swarm—behaviors that produce both individual and collective benefits.[22]

How might a policymaker leverage such a potent influence? One example comes from the United Kingdom. Like tax collectors in a lot of countries, Her Majesty's Revenue & Customs (HMRC) had a problem: Too many citizens weren't submitting their tax returns and paying what they owed on time. Over the years, officials at HMRC created a variety of letters and communications targeted at late payers. The majority of these approaches focused on traditional consequence-based inducements such as interest charges, late penalties, and the threat of legal action for those who failed to pay on time. For some, the traditional approaches worked well, but for many others, they did not. So, in early 2009, in consultation with Steve J. Martin, one of the present authors, HMRC piloted an alternative approach that was strikingly subtle. A

single extra sentence was added to the standard letters, truthfully stating the large number of UK citizens (the vast majority) who do pay their taxes on time. This one sentence communicated what similar others believe to be the correct course of action.

This small change was remarkable not only for its simplicity but also for the big difference it made in response rates. For the segment of outstanding debt that was the focus of the initial pilot, the new letters resulted in the collection of £560 million out of £650 million owed, representing a clearance rate of 86%. To put this into perspective, in the previous year, HMRC had collected £290 million of a possible £510 million—a clearance rate of just 57%.[23]

Because the behavior of the British taxpayers was completely private, this suggests the change was induced through what social psychologists call *informational influence,* rather than a concern about gaining the approval of their friends, neighbors, and peers. We contend that the addition of a social proof message to the tax letters triggered the fundamental motivation to make the "correct" choice. That is, in the context of a busy, information-overloaded life, doing what most others are doing can be a highly efficient shortcut to a good decision, whether that decision concerns which movie to watch; what restaurant to frequent; or, in the case of the UK's HMRC, whether or when to pay one's taxes.

Peer opinions and behaviors are not the only powerful levers of social influence. When uncertainty or ambiguity makes choosing accurately more difficult, individuals look to the guidance of experts, whom they see as more knowledgeable.[24–26] Policymakers, therefore, should aim to establish their own expertise—and/or the credibility of the experts they cite—in their influence campaigns. A number of strategies can be used to enhance one's expert standing. Using third parties to present one's credentials has proven effective in elevating one's perceived worth without creating the appearance of self-aggrandizement that undermines one's public image.[27] When it comes to establishing the credibility of cited experts, policymakers can do so by using a version of social proof: Audiences are powerfully influenced by the combined judgments of multiple experts, much more so than by the judgment of a single authority.[28] The implication for policymakers: Marshall the support of multiple experts, as they lend credibility to one another, advancing your case more forcefully in the process.

Another subtle way that communicators can establish their credibility is to use specific rather than round numbers in their proposals. Mason, Lee, Wiley, and Ames examined this idea in the context of negotiations.[29] They found that in a variety of types of negotiations, first offers that used precise-sounding numbers such as $1,865 or $2,135 were more effective than those that used round numbers like $2,000. A precise number conveys the message that the parties involved have carefully researched the situation and therefore have very good data to support that number. The policy implications of this phenomenon are clear. Anyone engaged in a budget negotiation should avoid using round estimates in favor of precise numbers that reflect actual needs—for example, "We believe that an expenditure of $12.03 million will be necessary." Not only do such offers appear more authoritative, they are more likely to soften any counteroffers in response.[29]

Affiliation and Approval

Humans are fundamentally motivated to create and maintain positive social relationships.[30] Affiliating with others helps fulfill two other powerful motivations: Others afford a basis for social comparison so that an individual can make an accurate assessment of the self,[31] and they provide opportunities to experience a sense of self-esteem and self-worth.[32] Social psychologists have demonstrated that the need to affiliate with others is so powerful that even seemingly trivial similarities among individuals can create meaningful social bonds. Likewise, a lack of shared similarities can spur competition.[33–36] For instance, observers are more likely to lend their assistance to a person in need if that person shares a general interest in football with observers, unless the person in need supports a rival team.[37]

Because social relations are so important to human survival, people are strongly motivated to gain the approval of others—and, crucially, to avoid the pain and isolation of being disapproved of or rejected.[12,38,39] This desire for social approval—and avoidance of social disapproval—can manifest itself in a number of ways. For example, in most cultures, there is a norm for keeping the environment clean, especially in public settings. Consequently, people refrain from littering so as to maximize the social approval and minimize the

social disapproval associated with such behavior.

What behavioral scientists have found is that minimizing social disapproval can be a stronger motivator than maximizing social approval. Let us return to the example of social norms for keeping public spaces clean. In one study, visitors to a city library found a handbill on the windshields of their cars when they returned to the public parking lot. On average, 33% of this control group tossed the handbill to the ground. A second group of visitors, while on the way to their cars, passed a man who disposed of a fast-food restaurant bag he was carrying by placing it in a trash receptacle; in these cases, a smaller proportion of these visitors (26%) subsequently littered with the handbill. Finally, a third set of visitors passed a man who disapprovingly picked up a fast-food bag from the ground; in this condition, only 6% of those observers improperly disposed of the handbill they found on their cars.[40] These data suggest that the most effective way to communicate behavioral norms is to express disapproval of norm breakers.

Furthermore, expressions of social disapproval in one area can induce desirable behavior beyond the specifically targeted domain. In one study, pedestrians walking alone encountered an individual who "accidentally" spilled a bag of oranges on a city sidewalk; 40% of them stopped to help pick the oranges up. Another set of pedestrians witnessed an individual who dropped an empty soft drink can immediately pick it up, thereby demonstrating normatively approved behavior; when this set of pedestrians encountered the stranger with the spilled oranges, 64% stopped to help. In a final condition, the pedestrians passed an individual who was sweeping up other people's litter, this time providing clear disapproval of socially undesirable behavior. Under these circumstances, 84% of the pedestrians subsequently stopped to help with the spilled oranges. Here is another example of the power of witnessed social disapproval to promote desired conduct. But in this instance, observed disapproval of littering led to greater helping in general.[41]

This phenomenon has significance for policymakers. Such findings suggest that programs should go beyond merely discouraging undesirable actions. Programs that depict people publically reversing those undesirable actions can be more effective.

Municipalities could allocate resources for the formation and/or support of citizens groups that want to demonstrate their disapproval of disordered environments by cleaning debris from lakes and beaches, graffiti from buildings, and litter from streets. Moreover, city governments would be well advised to then publicize those citizens' efforts and the manifest disapproval of disorder they reflect.

Another phenomenon arising from the primal need for affiliation and approval is the *norm of reciprocity*. This norm, which obliges people to repay others for what they have been given, is one of the strongest and most pervasive social forces across human cultures.[42] The norm of reciprocity tends to operate most reliably and powerfully in public domains.[8] Nonetheless, it is so deeply ingrained in human society that it directs behavior in private settings as well[43] and can be a powerful tool for policymakers for influencing others.

Numerous organizations use this technique under the banner of *cause-related marketing*. They offer to donate to causes that people consider important if, in return, those people will take actions that align with the organizations' goals. However, such tit-for-tat appeals are less effective if they fail to engage the norm of reciprocity properly.

The optimal activation of the norm requires a small but crucial adjustment in the sequencing of the exchange.[44] That is, benefits should be provided first in an unconditional manner, thereby increasing the extent to which individuals feel socially obligated to return the favor. For instance, a message promising a monetary donation to an environmental cause if hotel guests reused their towels (the typical cause-related marketing strategy) was no more effective than a standard control message simply requesting that the guests reuse their towels for the sake of the environment. However, consistent with the obligating force of reciprocity, a message that the hotel had already donated on behalf of its guests significantly increased subsequent towel reuse. This study has clear implications for governments and organizations that wish to encourage citizens to protect the environment: Be the first to contribute to such campaigns on behalf of those citizens and ask for congruent behavior after the fact.

To See Oneself Positively

Social psychologists have well documented people's desire to think favorably of themselves[45–50] and to take actions that maintain this positive self-view.[51,52] One

central way in which people maintain and enhance their positive self-concepts is by behaving consistently with their actions, statements, commitments, beliefs, and self-ascribed traits.[53,54] This powerful motivation can be harnessed by policymakers and practitioners to address all sorts of large-scale behavioral challenges. A couple of studies in the field of health care demonstrate how to do so.

Health care practitioners such as physicians, dentists, psychologists, and physical therapists face a common predicament: People often fail to appear for their scheduled appointments. Such episodes are more than an inconvenience; they are costly for practitioners. Recent research demonstrates how a small and no-cost change can solve this vexing problem. Usually, when a patient makes a future appointment after an office visit, the receptionist writes the appointment's time and date on a card and gives it to the patient. A recent study showed that if receptionists instead asked patients to fill in the time and date on the card, the subsequent no-show rate in their health care settings dropped from an average of 385 missed appointments per month (12.1%) to 314 missed appointments per month (9.8%).[55] Why? One way that people can think of themselves in a positive light is to stay true to commitments they personally and actively made.[56] Accordingly, the simple act of committing by writing down the appointment time and date was the small change that sparked a measurable difference.

Staying within the important domain of health care, whenever we consult with health management groups and ask who in the system is most difficult to influence, the answer is invariably "physicians." This can raise significant challenges, especially when procedural safeguards, such as hand washing before patient examinations, are being ignored.

In a study at a U.S. hospital, researchers varied the signs next to soap and sanitizing-gel dispensers in examination rooms.[57] One sign (the control condition) said, "Gel in, Wash out"; it had no effect on handwashing frequency. A second sign raised the possibility of adverse personal consequences to the practitioners. It said, "Hand hygiene prevents *you* from catching diseases"; it also had no measurable effect. But a third sign that said, "Hand hygiene prevents *patients* from catching diseases," increased hand washing from 37% to 54%. Reminding doctors of their professional commitment to their patients appeared to activate the motivation to be consistent with that commitment. Notice too that this small change did not even require an active commitment (as in the appointment no-show study). All that was necessary, with the change of a single word, was to remind physicians of a strong commitment they had made at the outset of their careers.

Potent Policy Tools

How can such small changes in procedure spawn such significant outcomes in behavior, and how can they be used to address longstanding policy concerns? It is useful to think of a triggering or releasing model in which relatively minor pressure—like pressing a button or flipping a switch—can launch potent forces that are stored within a system. In the particular system of factors that affect social influence, the potent forces that generate persuasive success often are associated with the three basic motivations we have described. Once these stored forces are discharged by even small triggering events, such as a remarkably minor messaging shift, they have the power to effect profound changes in behavior.

Of course, the power of these motivation-triggering strategies is affected by the context in which people dwell. For example, strategies that attempt to harness the motivation for accuracy are likely to be most effective when people believe the stakes are high,[16,58] such as in the choice between presidential candidates. Approaches that aim to harness the motivation for affiliation tend to be most effective in situations where people's actions are visible to a group that will hold them accountable,[59] such as a vote by show of hands at a neighborhood association meeting. The motivation for positive self-regard tends to be especially effective in situations possessing a potential threat to self-worth,[51,60] such as in circumstances of financial hardship brought on by an economic downturn. Therefore, policymakers, communicators, and change agents should carefully consider the context when choosing which of the three motivations to leverage.

Finally, it is heartening to recognize that behavioral science is able to offer guidance on how to significantly improve social outcomes with methods that are not costly, are entirely ethical, and are empirically grounded. None of the effective changes described in this piece had emerged naturally as best practices

within government tax offices, hotel sustainability programs, medical offices, or hospital examination rooms. Partnerships with behavioral science led to the conception and successful testing of these strategies. Therefore, the prospect of a larger policymaking role for such partnerships is exciting.

At the same time, it is reasonable to ask how such partnerships can be best established and fostered. We are pleased to note that several national governments—the United Kingdom, first, but now the United States and Australia as well—are creating teams designed to generate and disseminate behavioral science–grounded evidence regarding wise policymaking choices. Nonetheless, we think that policymakers would be well advised to create internal teams as well. A small cadre of individuals knowledgeable about current behavioral science thinking and research could be highly beneficial to an organization. First, they could serve as an immediately accessible source of behavioral science–informed advice concerning the unit's specific policymaking challenges. Second, they could serve as a source of new data regarding specific challenges; that is, they could be called upon to conduct small studies and collect relevant evidence if that evidence was not present in the behavioral science literature. We are convinced that such teams would promote more vibrant and productive partnerships between behavioral scientists and policymakers well into the future.

author affiliation

Cialdini, Department of Psychology, Arizona State University; Martin, Influence At Work UK; Goldstein, Anderson School of Management, UCLA. Corresponding author's e-mail: Robert.Cialdini@asu.edu

References

1. Kogut, T., & Ritov, I. (2005). The singularity effect of identified victims in separate and joint evaluations. *Organizational Behavior and Human Decision Processes, 97,* 106–116.
2. Leshner, G., Bolls, P., & Thomas, E. (2009). Scare 'em or disgust 'em: The effects of graphic health promotion messages. *Health Communication, 24,* 447–458.
3. Small, D. A., & Loewenstein, G. (2003). Helping a victim or helping the victim: Altruism and identifiability. *Journal of Risk and Uncertainty, 26,* 5–16.
4. Johnson, E. J., & Goldstein, D. G. (2003, November 21). Do defaults save lives? *Science, 302,* 1338–1339.
5. Madrian, B., & Shea, D. (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics, 66,* 1149–1188.
6. Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.
7. Cialdini, R. B., & Trost, M. R. (1998). Social influence: Social norms, conformity, and compliance. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The handbook of social psychology* (Vol. 2, pp. 151–192). Boston, MA: McGraw-Hill.
8. Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology, 55,* 591–621.
9. Goldstein, N. J., Martin, S. J., & Cialdini, R. B. (2008). *Yes! 50 scientifically proven ways to be persuasive.* New York, NY: Free Press.
10. Martin, S. J, Goldstein, N. J., & Cialdini, R. B. (2014). *The small big: Small changes that spark big influence.* New York, NY: Hachette.
11. Nolan, J. P., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2008). Normative social influence is underdetected. *Personality and Social Psychology Bulletin, 34,* 913–923.
12. Deutsch, M., & Gerard, H. B. (1955). A study of normative and informational social influences upon individual judgment. *Journal of Abnormal and Social Psychology, 51,* 629–636.
13. Jones, E. E., & Gerard, H. (1967). *Foundations of social psychology*. New York, NY: Wiley.
14. Sherif, M. (1936). *The psychology of social norms*. New York, NY: Harper.
15. White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66,* 297–333.
16. Baron, R. S., Vandello, J. A., & Brunsman, B. (1996). The forgotten variable in conformity research: Impact of task importance on social influence. *Journal of Personality and Social Psychology, 71,* 915–927.
17. Wooten, D. B., & Reed, A. (1998). Informational influence and the ambiguity of product experience: Order effects in the weighting of evidence. *Journal of Consumer Psychology, 7,* 79–99.
18. Cialdini, R. B. (2009). *Influence: Science and practice*. Boston, MA: Pearson Education.
19. Hastie, R., & Kameda, T. (2005). The robust beauty of majority rules in group decisions. *Psychological Review, 112,* 494–508.
20. Hill, G. W. (1982). Group versus individual performance: Are *N* + 1 heads better than one? *Psychological Bulletin, 91,* 517–539.
21. Surowiecki, J. (2005). *The wisdom of crowds*. New York, NY: Anchor.
22. Claidière, N., & Whiten, A. (2012). Integrating the study of conformity and culture in humans and nonhuman animals. *Psychological Bulletin, 138,* 126–145.
23. Martin, S. (2012, October). 98% of HBR readers love this article. *Harvard Business Review, 90*(10), 23–25.

24. Hovland, C. I., Janis, I. L., & Kelley, H. H. (1953). *Communication and persuasion: Psychological studies of opinion and change*. New Haven, CT: Yale University Press.
25. Kelman, H. C. (1961). Processes of opinion change. *Public Opinion Quarterly, 25,* 57–78.
26. McGuire, W. J. (1969). The nature of attitudes and attitude change. In G. Lindzey & E. Aronson (Eds.), *The handbook of social psychology* (2nd ed., Vol. 3, pp. 136–314). Reading, MA: Addison-Wesley.
27. Pfeffer, J., Fong, C. T., Cialdini, R. B., & Portnoy, R. R. (2006). Why use an agent in transactions? *Personality and Social Psychology Bulletin, 32,* 1362–1374.
28. Mannes, A. E., Soll, J. B., & Larrick, R. P. (2014). The wisdom of select crowds. *Journal of Personality and Social Psychology, 107,* 276–299.
29. Mason, M. F., Lee, A. J., Wiley, E. A., & Ames, D. R. (2013). Precise offers are potent anchors: Conciliatory counteroffers and attributions of knowledge in negotiations. *Journal of Experimental Social Psychology, 49,* 759–763.
30. Baumeister, R. F., & Leary, M. R. (1995). The need to belong: Desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin, 117,* 497–529.
31. Festinger, L. (1954). A theory of social comparison processes. *Human Relations, 7,* 117–140.
32. Crocker, J., & Wolfe, C. T. (2001). Contingencies of self-worth. *Psychological Review, 108,* 593–623.
33. Brewer, M. B. (1979). In-group bias in the minimal intergroup situation: A cognitive-motivational analysis. *Psychological Bulletin, 86,* 307–324.
34. Sherif, M., Harvey, O. J., White, B. J., Hood, W. R., & Sherif, C. W. (1961). *Intergroup conflict and cooperation: The Robbers Cave experiment*. Norman, OK: University Book Exchange.
35. Tajfel, H. (1970, November). Experiments in intergroup discrimination. *Scientific American*, *223*(5), 96–102.
36. Turner, J. C. (1991). *Social influence*. Pacific Grove, CA: Brooks/Cole.
37. Levine, M., Prosser, A., & Evans, D. (2005). Identity and emergency intervention: How social group membership and inclusiveness of group boundaries shape helping behavior. *Personality and Social Psychology Bulletin, 31,* 443–453.
38. Eisenberger, N. I., Lieberman, M. D., & Williams, K. D. (2003, October 10). Does rejection hurt? An fMRI study of social exclusion. *Science, 302,* 290–292.
39. Williams, K. D. (2007). Ostracism. *Annual Review of Psychology, 58,* 425–452.
40. Reno, R. R., Cialdini, R. B., & Kallgren, C. A. (1993). The trans-situational influence of social norms. *Journal of Personality and Social Psychology, 64,* 104–112.
41. Keizer, K., Lindenberg, S., & Steg, L. (2013). The importance of demonstratively restoring order. *PLoS One, 8*(6), Article e65137.
42. Gouldner, A. W. (1960). The norm of reciprocity: A preliminary statement. *American Sociological Review, 25,* 161–178.
43. Whatley, M. A., Webster, J. M., Smith, R. H., & Rhodes, A. (1999). The effect of a favor on public and private compliance: How internalized is the norm of reciprocity? *Basic and Applied Social Psychology, 21,* 251–259.
44. Goldstein, N. J., Griskevicius, V., & Cialdini, R. B. (2011). Reciprocity by proxy: A novel influence strategy for stimulating cooperation. *Administrative Science Quarterly, 56,* 441–473.
45. Kleine, R. E., III, Kleine, S. S., & Kernan, J. B. (1993). Mundane consumption and the self: A social-identity perspective. *Journal of Consumer Psychology, 2,* 209–235.
46. Taylor, S. E., & Brown, J. D. (1988). Illusion and well-being: A social psychological perspective on mental health. *Psychological Bulletin, 103,* 193–210.
47. Blaine, B., & Crocker, J. (1993). Self-esteem and self-serving biases in reactions to positive and negative events: An integrative review. In R. Baumeister (Ed.), *Self-esteem: The puzzle of low self-regard* (pp. 55–85). New York, NY: Springer.
48. Greenwald, A. G. (1980). The totalitarian ego: Fabrication and revision of personal history. *American Psychologist, 35,* 603–618.
49. Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology, 77,* 1121–1134.
50. Ross, M., & Sicoly, F. (1979). Egocentric biases in availability and attribution. *Journal of Personality and Social Psychology, 37,* 322–336.
51. Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. *Advances in Experimental Social Psychology, 21,* 261–302.
52. Tesser, A. (1988). Toward a self-evaluation maintenance model of social behavior. *Advances in Experimental Social Psychology, 21,* 181–227.
53. Cialdini, R. B., Trost, M. R., & Newsom, J. T. (1995). Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications. *Journal of Personality and Social Psychology, 69,* 318–328.
54. Heider, F. (1958). *The psychology of interpersonal relations*. New York, NY: Wiley.
55. Martin, S. J., Bassi, S., & Dunbar-Rees, R. (2012). Commitments, norms and custard creams—A social influence approach to reducing did not attends (DNAs). *Journal of the Royal Society of Medicine, 105,* 101–104.
56. Cioffi, D., & Garner, R. (1996). On doing the decision: Effects of active versus passive choice on commitment and self-perception. *Personality and Social Psychology Bulletin, 22,* 133–147.
57. Grant, A. M., & Hofmann, D. A. (2011). It's not all about me: Motivating hand hygiene among health care professionals by focusing on patients. *Psychological Science, 22,* 1494–1499.
58. Marsh, K. L., & Webb, W. M. (1996). Mood uncertainty and social comparison: Implications for mood management. *Journal of Social Behavior and Personality, 11,* 1–26.
59. Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin, 125,* 255–275.
60. Steele, C. M. (1997). A threat in the air: How stereotypes shape intellectual identity and performance. *American Psychologist, 52,* 613–629.

Essay

Active choosing or default rules? The policymaker's dilemma

Cass R. Sunstein

abstract. It is important for people to make good choices about important matters, such as health insurance or retirement plans. Sometimes it is best to ask people to make active choices. But in some contexts, people are busy or aware of their own lack of knowledge, and providing default options is best for choosers. If people elect not to choose or would do so if allowed, they should have that alternative. A simple framework, which assesses the costs of decisions and the costs of errors, can help policymakers decide whether active choosing or default options are more appropriate.

Consider the following problems:

- Public officials are deciding whether to require people, as a condition for obtaining a driver's license, to choose whether to become organ donors. The alternatives are to continue with the existing *opt-in* system, in which people become organ donors only if they affirmatively indicate their consent, or to switch to an *opt-out* system, in which consent is presumed.
- A public university is weighing three options: to enroll people automatically in a health insurance plan; to make them opt in if they want to enroll; or, as a condition for starting work, to require them to indicate whether they want health insurance and, if so, which plan they want.
- A utility company is deciding which is best: a "green default," with a somewhat more expensive but environmentally favorable energy source, or a "gray default," with a somewhat less expensive but environmentally less favorable energy source. Or should the utility ask consumers which energy source they prefer?
- A social media site is deciding whether to adopt a system of default settings for privacy or to require first-time users to identify, as a condition for access, what privacy settings they want. Public officials are monitoring the decision and are considering regulatory intervention if the decision does not serve users' interests.

In these cases and countless others, policymakers are evaluating whether to use or promote a *default rule,* meaning a rule that establishes what happens if people do not actively choose a different option. A great deal of research has shown that for identifiable reasons, default rules have significant effects on outcomes; they

Sunstein, C. R. (2015). Active choosing or default rules? The policymaker's dilemma. *Behavioral Science & Policy, 1*(1), pp. 29–33.

tend to "stick" or persist over time.[1] For those who prize freedom of choice, active choosing might seem far preferable to any kind of default rule.

My goal here is to defend two claims. The first is that in many contexts, an insistence on active choosing is a form of paternalism, not an alternative to it. The reason is that people often choose not to choose, for excellent reasons. In general, policymakers should not force people to choose when they prefer not to do so (or would express that preference if asked).

The second claim is that when policymakers decide between active choosing and a default rule, they should focus on two factors. The first is the costs of making decisions. If active choosing is required, are people forced to incur large costs or small ones? The second is the costs of errors: Would the number and magnitude of mistakes be higher or lower with active choosing than with default rules?

These questions lead to some simple rules of thumb. When the situation is complex, technical, and unfamiliar, active choosing may impose high costs on choosers, and they might ultimately err. In such cases, there is a strong argument for a default rule rather than for active choosing. But if the area is one that choosers understand well, if their situations (and needs) are diverse, and if policymakers lack the means to devise accurate defaults, then active choosing would be best.

This framework can help orient a wide range of policy questions. In the future, it may be feasible to personalize default rules and tailor them to particular groups or people. This may avoid current problems associated with both active choosing and defaults designed for very large groups of people.[2]

Active Choosing Can Be Paternalistic

With the help of modern technologies, policymakers are in an unprecedented position to ask people this question: What do you choose? Whether the issue involves organ donation, health insurance, retirement plans, energy, privacy, or nearly anything else, it is simple to pose that question (and, in fact, to do so repeatedly and in real time, thus allowing people to signal new tastes and values). Those who reject paternalism and want to allow people more autonomy tend to favor active choosing. Indeed, there is empirical evidence that in some contexts, ordinary people will pay a premium to be able to choose as they wish.[3,4] (Compare the related phenomenon of *reactance,* which suggests a negative reaction to coercive efforts, produced in part by the desire to assert autonomy.[5]) In other cases, people will pay a premium to be relieved of that very obligation.

There are several reasons why people might choose not to choose. They might fear that they will err. They might not enjoy choosing. They might be too busy. They might lack sufficient information or *bandwidth*.[6] They might not want to take responsibility for potentially bad outcomes for themselves (and at least indirectly for others).[7,8] They might find the underlying questions confusing, difficult, painful, and troublesome—empirically, morally, or otherwise. They might anticipate their own regret and seek to avoid it. They might be keenly aware of their own lack of information or perhaps even of their own behavioral biases (such as unrealistic optimism or *present bias*, understood as an undue focus on the near term). In the area of retirement savings or health insurance, many employees might welcome a default option, especially if they trust the person or institution selecting the default.

It is true that default rules tend to stick, and some people distrust them for that reason. The concern is that people do not change default options out of inertia (and thus reduce the costs of effort). With an opt-in design (by which the chooser has to act to participate), there will be far less participation than with an opt-out design (by which the chooser has to act to avoid participation).[1] Internet shopping sites often use an opt-out default for future e-mail correspondence: The consumer must uncheck a box to avoid being put on a mailing list. It is well established that social outcomes are decisively influenced by the choice of default in areas that include organ donation, retirement savings, environmental protection, and privacy. Policymakers who are averse to any kind of paternalism might want to avoid the appearance of influencing choice and require active choosing.[9]

When policymakers promote active choosing on the ground that it is good for people to choose, they are acting paternalistically. *Choice-requiring paternalism* might appear to be an oxymoron, but it is a form of paternalism nonetheless.

Respecting Freedom of Choice

Those who favor paternalism tend to focus on the

quality of outcomes.[10] They ask, "What promotes human welfare?" Those who favor libertarianism tend to focus instead on process. They ask, "Did people choose for themselves?" Some people think that libertarian paternalism is feasible and seek approaches that will promote people's welfare while also preserving freedom of choice.[11] But many committed libertarians are deeply skeptical of the attempted synthesis: They want to ensure that people actually choose.[9]

It is worth distinguishing between the two kinds of libertarians. For some, freedom of choice is a means. They believe that such freedom should be preserved, because choosers usually know what is best for them. At the very least, choosers know better than outsiders (especially those outsiders employed by the government) what works in their situation. Those who endorse this view might be called *epistemic libertarians*, because they are motivated by a judgment about who is likely to have the most knowledge. Other libertarians believe that freedom of choice is an end in itself. They think that people have a right to choose even if they will choose poorly. People who endorse this view might be called a*utonomy libertarians*.

When people choose not to choose, both types of libertarians should be in fundamental agreement. Suppose, for example, that Jones believes that he is not likely to make a good choice about his retirement plan and that he would therefore prefer a default option, chosen by a financial planner. Or suppose that Smith is exceedingly busy and wants to focus on her most important or immediate concerns, not on which health insurance plan or computer privacy setting best suits her. Epistemic libertarians think that people are uniquely situated to know what is best for them. If so, then that very argument should support respect for people when they freely choose not to choose. Autonomy libertarians insist that it is important to respect people's autonomy. If so, then it is also important to respect people's decisions about whether and when to choose.

If people are required to choose even when they would prefer not to do so, active choosing becomes a form of paternalism. If, by contrast, people are asked whether they want to choose and can opt out of active choosing (in favor of, say, a default option), active choosing counts as a form of libertarian paternalism. In some cases, it is an especially attractive form. A private or public institution might ask people whether they want to choose the privacy settings on their computer or instead rely on the default, or whether they want to choose their electricity supplier or instead rely on the default.

With such an approach, people are being asked to make an active choice between the default and their own preference: In that sense, their liberty is fully preserved. Call this *simplified active choosing*. This approach has evident appeal, and in the future, it is likely to prove attractive to a large number of institutions, both public and private.

It is important to acknowledge that choosers' best interests may not be served by the choice not to choose. Perhaps a person lacks important information, which would reveal that the default rule might be harmful. Or perhaps a person is myopic, being excessively influenced by the short-term costs of choosing while underestimating the long-term benefits, which might be very large. A form of present bias might infect the decision not to choose.

For those who favor freedom of choice, these kinds of concerns are usually a motivation for providing more and better information or for some kind of nudge—not for blocking people's choices, including their choices not to choose. In light of people's occasional tendency to be overconfident, the choice not to choose might, in fact, be the best action. That would be an argument against choice-requiring paternalism. Consider in this regard behavioral evidence that people spend too much time pursuing precisely the right choice. In many situations, people underestimate the temporal costs of choosing and exaggerate the benefits, producing "systematic mistakes in predicting the effect of having more, vs. less, choice freedom on task performance and task-induced affect."[12]

If people prefer not to choose, they might favor either an opt-in or an opt-out design. In the context of both retirement plans and health insurance, for example, many people prefer opt-out options on the grounds that automatic enrollment overcomes inertia and procrastination and produces sensible outcomes for most employees. Indeed, the Affordable Care Act calls for automatic enrollment by large employers, starting in 2015. For benefits programs that are either required by law or generally in people's interests, automatic enrollment has considerable appeal.

In the context of organ donation, by contrast, many people prefer an opt-in design on moral grounds,

even though more lives would be saved with opt-out designs. If you have to opt out to avoid being an organ donor, maybe you'll stay in the system and not bother to opt out, even if you do not really want to be an organ donor. That might seem objectionable. As the experience in several states suggests, a system of active choosing can avoid the moral objections to the opt-out design while also saving significant numbers of lives.

Are people genuinely bothered by the existence of default rules, or would they be bothered if they were made aware that such rules had been chosen for them? A full answer is not available for this question: The setting and the level of trust undoubtedly matter. In the context of end-of-life care, when it is disclosed that a default rule is in place, there is essentially no effect on what people do. (*Editor's note: See the article "Warning: You Are about to Be Nudged" in this issue*.) This finding suggests that people may not be uncomfortable with defaults, even when they are made aware that choice architects have selected them to influence outcomes.[13] More research on this question is highly desirable.

Weighing Decision Costs and Error Costs

The choice between active choosing and default rules cannot be made in the abstract. If welfare is the guide, policymakers need to investigate two factors: the costs of decisions and the costs of errors. In some cases, active choosing imposes high costs, because it is time-consuming and difficult to choose. For example, it can be hard to select the right health insurance plan or the right retirement plan. In other cases, the decision is relatively easy, and the associated costs are low. For most people, it is easy, to choose among ice cream flavors. Sometimes people actually enjoy making decisions, in which case decision costs turn out to be benefits.

The available information plays a role here as well. In some cases, active choosing reduces the number and magnitude of errors, because choosers have far better information about what is good for them than policymakers do. Ice cream choices are one example; choices among books and movies are another. In other cases, active choosing can increase the number and magnitude of errors, because policymakers have more relevant information than choosers do. Health insurance plans might well be an example.

With these points in mind, two propositions are clear, and they can help orient this inquiry in diverse settings. First, policymakers should prefer default rules to active choosing when the context is confusing and unfamiliar; when people would prefer not to choose; and when the population is diverse with respect to wants, values, and needs. The last point is especially important. Suppose that with respect to some benefit, such as retirement plans, one size fits all or most, in the sense that it promotes the welfare of a large percentage of the affected population. If so, active choosing might be unhelpful or unnecessary.

Second, policymakers should generally prefer active choosing to default rules when choice architects lack relevant information, when the context is familiar, when people would actually prefer to choose (and hence choosing is a benefit rather than a cost), when learning matters, and when there is relevant heterogeneity. Suppose, for example, that with respect to health insurance, people's situations are highly diverse with regard to age, preexisting conditions, and risks for future illness, so any default rule will be ill suited to most or many. If so, there is a strong argument for active choosing.

To be sure, the development of personalized default rules, designed to fit individual circumstances, might solve or reduce the problems posed by heterogeneity.[14,15] As data accumulate about what informed people choose or even about what particular individuals choose, it will become more feasible to devise default rules that fit diverse situations. With retirement plans, for example, demographic information is now used to produce different initial allocations, and travel websites are able to incorporate information about past choices to select personalized defaults (and thus offer advice on future destinations).[2,14] For policymakers, the rise of personalization promises to reduce the costs of uniform defaults and to reduce the need for active choosing. At the same time, however, personalization also raises serious questions about both feasibility and privacy.

A further point is that active choosing has the advantage of promoting learning and thus the development of preferences and values. In some cases, policymakers might know that a certain outcome is in the interest of most people. But they might also believe that it is important for people to learn about underlying issues, so they can apply what was gained to future choices. In the context of decisions that involve health and

retirement, the more understanding people develop, the more they will be able to choose well for themselves. Those who favor active choosing tend to emphasize this point and see it as a powerful objection to default rules. They might be right, but the context greatly matters. People's time and attention are limited, and the question is whether it makes a great deal of sense to force them to get educated in one area when they would prefer to focus on others.

Suppose that an investigation into decision and error costs suggests that a default rule is far better than active choosing. If so, epistemic libertarians should be satisfied. Their fundamental question is whether choice architects know as much as choosers do, and the idea of error costs puts a spotlight on the question that most troubles them. If a default rule reduces those costs, they should not object.

It is true that in thinking about active choosing and default rules, autonomy libertarians have valid and distinctive concerns. Because they think that choice is important in itself, they might insist that people should be choosing even if they might err. The question is whether their concerns might be alleviated or even eliminated so long as freedom of choice is fully preserved by offering a default option. If coercion is avoided and people are allowed to go their own way, people's autonomy is maintained.

In many contexts, the apparent opposition between active choosing and paternalism is illusory and can be considered a logical error. The reason is that some people choose not to choose, or they would do so if they were asked. If policymakers are overriding that particular choice, they may well be acting paternalistically. With certain rules of thumb, based largely on the costs of decisions and the costs of errors, policymakers can choose among active choosing and default rules in a way that best serves choosers.

author affiliation

Sunstein, Harvard University Law School. Corresponding author's e-mail: csunstei@law.harvard.edu

author note

The author, Harvard's Robert Walmsley University Professor, is grateful to Eric Johnson and three anonymous referees for valuable suggestions. This article draws on longer treatments of related topics, including Cass R. Sunstein, *Choosing Not to Choose* (Oxford University Press, 2015).

References

1. Johnson, E. J., & Goldstein, D. G. (2012). Decisions by default. In E. Shafir (Ed.), *The behavioral foundations of policy* (pp. 417–418). Princeton, NJ: Princeton University Press.
2. Goldstein, D. G., Johnson, E. J., Herrmann, A., & Heitmann, M. (2008). Nudge your customers toward better choices. *Harvard Business Review, 86,* 99–105.
3. Fehr, E., Herz, H., & Wilkening, T. (2013). The lure of authority: Motivation and incentive effects of power. *American Economic Review, 103,* 1325–1359.
4. Bartling, B., Fehr, E., & Herz, H. (2014). *The intrinsic value of decision rights* (Working Paper No. 120). Zurich, Switzerland: University of Zurich, Department of Economics.
5. Pavey, L., & Sparks, P. (2009). Reactance, autonomy and paths to persuasion: Examining perceptions of threats to freedom and informational value. *Motivation and Emotion, 33,* 277–290.
6. Mullainathan, S., & Shafir, E. (2013). *Scarcity: Why having too little means so much*. New York, NY: Times Books.
7. Bartling, B., & Fischbacher, U. (2012). Shifting the blame: On delegation and responsibility. *Review of Economic Studies, 79,* 67–87.
8. Dwengler, N., Kübler, D., & Weizsäcker, G. (2013). *Flipping a coin: Theory and evidence.* Unpublished manuscript. Retrieved from http://www.wiwi.hu-berlin.de/professuren/vwl/mt-anwendungen/team/flipping-a-coin
9. Rebonato, R. (2012). *Taking liberties: A critique of libertarian paternalism*. London, United Kingdom: Palgrave Macmillan.
10. Conly, S. (2013). *Against autonomy: Justifying coercive paternalism.* Cambridge, United Kingdom: Cambridge University Press.
11. Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about wealth, health, and happiness*. New York, NY: Penguin.
12. Botti, S., & Hsee, C. (2010). Dazed and confused by choice: How the temporal costs of choice freedom lead to undesirable outcomes. *Organizational Behavior and Human Decision Processes, 112,* 161–171.
13. Loewenstein, G., Bryce, C., Hagman, D., & Rajpal, S. (2015). Warning: You are about to be nudged. *Behavioral Science & Policy, 1,* 35–42.
14. Smith, N. C., Goldstein, D. G., & Johnson, E. J. (2013). Choice without awareness: Ethical and policy implications of defaults. *Journal of Public Policy & Marketing, 32,* 159–172.
15. Sunstein, C. R. (2013). Deciding by default. *University of Pennsylvania Law Review, 162,* 1–57.


SWITZE

Finding

Warning: You are about to be nudged

George Loewenstein, Cindy Bryce, David Hagmann, & Sachin Rajpal

abstract. Presenting a default option is known to influence important decisions. That includes decisions regarding advance medical directives, documents people prepare to convey which medical treatments they favor in the event that they are too ill to make their wishes clear. Some observers have argued that defaults are unethical because people are typically unaware that they are being nudged toward a decision. We informed people of the presence of default options before they completed a hypothetical advance directive, or after, then gave them the opportunity to revise their decisions. The effect of the defaults persisted, despite the disclosure, suggesting that their effectiveness may not depend on deceit. These findings may help address concerns that behavioral interventions are necessarily duplicitous or manipulative.

Nudging people toward particular decisions by presenting one option as the default can influence important life choices. If a form enrolls employees in retirement savings plans by default unless they opt out, people are much more likely to contribute to the plan.[1] Likewise, making organ donation the default option rather than just an opt-in choice dramatically increases rates of donation.[2] The same principle holds for other major decisions, including choices about purchasing insurance and taking steps to protect personal data.[3,4]

Decisions about end-of-life medical care are similarly susceptible to the effects of defaults. Two studies found that default options had powerful effects on the end-of-life choices of participants preparing hypothetical advance directives. One involved student respondents, and the other involved elderly outpatients.[5,6] In a more recent study, defaults also proved robust when seriously ill patients completed real advance directives.[7]

The use of such defaults or other behavioral *nudges*[8] has raised serious ethical concerns, however. The House of Lords *Behaviour Change* report produced in the United Kingdom in 2011 contains one of the most significant critiques.[9] It argued that the "extent to which an intervention is covert" should be one of the main criteria for judging if a nudge is defensible. The report considered two ways to disclose default interventions: directly or by ensuring that a perceptive person could discern a nudge is in play. While acknowledging that the former would be preferable from a purely ethical perspective, the report concluded that the latter should be adequate, "especially as this fuller sort of transparency might limit the effectiveness of the intervention."

Philosopher Luc Bovens in "The Ethics of Nudge" noted that default options "typically work best in the dark."[10] Bovens observed the lack of disclosure in a study in which healthy foods were introduced at a school cafeteria with no explanation, prompting students to eat fewer unhealthy foods. The same lack of transparency existed during the rollout of the Save More Tomorrow program, which gave workers the option of precommitting themselves to increase their savings rate as their income rose in the future. Bovens noted,

---

Loewenstein, G., Bryce, C., Hagmann, D., & Rajpal, S. (2015). Warning: You are about to be nudged. *Behavioral Science & Policy, 1*(1), pp. 35–42.

> If we tell students that the order of the food in the Cafeteria is rearranged for dietary purposes, then the intervention may be less successful. If we explain the endowment effect [the tendency for people to value amenities more when giving them up than when acquiring them] to employees, they may be less inclined to Save More Tomorrow.

When we embarked on our research into the impact of disclosing nudges, we understood that alerting people about defaults could make them feel that they were being manipulated. Social psychology research has found that people tend to resist threats to their freedom to choose, a phenomenon known as *psychological reactance*.[11] Thus, it is reasonable to think, as both the House of Lords report and Bovens asserted, that people would deliberately resist the influence of defaults (if informed ahead of time, or *preinformed*) or try to undo their influence (if told after the fact, or *postinformed*). Such a reaction to disclosure might well reduce or even eliminate the influence of nudges.

But our findings challenge the idea that fuller transparency substantially harms the effectiveness of defaults. If what we found is confirmed in broader contexts, fuller disclosure of a nudge could potentially be achieved with little or no negative impact on the effectiveness of the intervention. That could have significant practical applications for policymakers trying to help people make choices that are in their and society's long-term interests while disclosing the presence of nudges.

Testing Effects from Disclosing Defaults

We explored the impact of disclosing nudges in a study of individual choices on hypothetical advance directives, documents that enable people to express their preferences for medical treatment for times when they are near death and too ill to express their wishes. Participants completed hypothetical advance directives by stating their overall goals for end-of-life care and their preferences for specific life-prolonging measures such as cardiopulmonary resuscitation and feeding tube insertion. Participants were randomly assigned to receive a version of an advance directive form on which the default options favored either prolonging life or minimizing discomfort. For both defaults, participants were further randomly assigned to be informed about the defaults either before or after completing the form. Next, they were allowed to change their decisions using forms with no defaults included. The design of the study enabled us to assess the effects of participants' awareness of defaults on end-of-life decisionmaking.

We recognize that the hypothetical nature of the advance directive in our study may raise questions about how a similar process would play out in the real world. However, recent research by two of the current authors and their colleagues examined the impact of defaults on real advance directives[7] and obtained results similar to prior work on the topic examining hypothetical choices.[5,6] All of these studies found that the defaults provided on advance directive forms had a major impact on the final choices reached by respondents. Just as the question of whether defaults could influence the choices made in advance directives was initially tested in hypothetical tasks, we test first in a hypothetical setting whether alerting participants to the default diminishes its impact.

To examine the effects of disclosing the presence of defaults, we recruited via e-mail 758 participants (out of 4,872 people contacted) who were either alumni of Carnegie Mellon University or *New York Times* readers who had consented to be contacted for research. Respondents were not paid for participating. Although not a representative sample of the general population, the 1,027 people who participated included a large proportion of older individuals for whom the issues posed by the study are salient. The mean age for both samples was about 50 years, an age when end-of-life care tends to become more relevant. (Detailed descriptions of the methods and analysis used in this research are published online in the Supplemental Material.)

Our sample populations are more educated than the U.S. population as a whole, which reduces the extent to which we can generalize the results to the wider population. However, the study provides information about whether the decisions of a highly educated and presumably commensurately deliberative group are changed by their awareness of being defaulted, that is, having the default options selected for them should they not take action to change them. Prior research has documented larger default effects for individuals of lower socioeconomic status,[1,12] which suggests that the default effects we observe would likely be larger in a

less educated population.

Obtaining End-of-Life Preferences Participants completed an online hypothetical advance directive form. First, they were asked to indicate their broad goals for end-of-life care by selecting one of the following options:

- I want my health care providers and agent to pursue treatments that help me to live as long as possible, even if that means I might have more pain or suffering.
- I want my health care providers and agent to pursue treatments that help relieve my pain and suffering, even if that means I might not live as long.
- I do not want to specify one of the above goals. My health care providers and agent may direct the overall goals of my care.

Next, participants expressed their preferences regarding five specific medical life-prolonging interventions. For each question, participants expressed a preference for pursuing the treatment (the prolong option), declining it (the comfort option), or leaving the decision to a family member or other designated person (the no-choice option). The specific interventions included the following:

- cardiopulmonary resuscitation, described as "manual chest compressions performed to restore blood circulation and breathing";
- dialysis (kidney filtration by machine);
- feeding tube insertion, described as "devices used to provide nutrition to patients who cannot swallow, inserted either through the nose and esophagus into the stomach or directly into the stomach through the belly";
- intensive care unit admission, described as a "hospital unit that provides specialized equipment, services, and monitoring for critically ill patients, such as higher staffing-to-patient ratios and ventilator support"; and
- mechanical ventilator use, described as "machines that assist spontaneous breathing, often using either a mask or a breathing tube."

The advance directive forms that participants completed randomly defaulted them into either accepting or rejecting each of the life-prolonging treatments. Those preinformed about the use of defaults were told before filling out the form; those postinformed learned after completing the form.

One reason that defaults can have an effect is that they are sometimes interpreted as implicit recommendations.[2,13–15] This is unlikely in our study, because both groups were informed that other study participants had been provided with forms populated with an alternative default. This disclosure also rules out the possibility that respondents attached different meanings to opting into or out of the life-extending measures (for example, donating organs is seen as more altruistic in countries in which citizens must opt in to donate than in countries in which citizens must opt out of donation)[16] or the possibility that the default would be perceived as a social norm (that is, a standard of desirable or common behavior).

After completing the advance directive a first time (either with or without being informed about the default at the outset), both groups were then asked to complete the advance directive again, this time with no defaults. Responses to this second elicitation provide a conservative test of the impact of defaults. Defaults can influence choices if people do not wish to exert effort or are otherwise unmotivated to change their responses. Requiring people to complete a second advance directive substantially reduces marginal switching costs (that is, the additional effort required to switch) when compared with a traditional default structure in which people only have to respond if they want to reject the default. In our two-stage setup, participants have already engaged in the fixed cost (that is, expended the initial effort) of entering a new response, so the marginal cost of changing their response should be lower. The fact that the second advance directive did not include any defaults means that the only effect we captured is a carryover from the defaults participants were given in the first version they completed.

In sum, the experiment required participants to make a first set of advance directive decisions in which a default had been indicated and then a second set of decisions in which no default had been indicated. Participants were randomly assigned into one of four groups in which they were either preinformed or postinformed that they had been assigned either a prolong default or a comfort default for their first choice, as depicted in Table 1.

The disclosure on defaults for the preinformed group

Table 1. Experimental design

Group 1: Comfort preinformed	Group 2: Comfort postinformed	Group 3: Prolong preinformed	Group 4: Prolong postinfomed
Disclosure		Disclosure	
Choice 1 Comfort default	Choice 1 Comfort default	Choice 1 Prolong default	Choice 1 Prolong default
	Disclosure		Disclosure
Choice 2 No default	Choice 2 No default	Choice 2 No default	Choice 2 No default

read as follows:

> The specific focus of this research is on "defaults"—decisions that go into effect if people don't take actions to do something different. Participants in this research project have been divided into two experimental groups.
>
> If you have been assigned to one group, the Advance Directive you complete will have answers to questions checked that will direct health care providers to help relieve pain and suffering even it means not living as long. If you want to choose different options, you will be asked to check off a different option and place your initials beside the different option you select.
>
> If you have been assigned to the other group, the Advance Directive you complete will have answers to questions checked that will direct health care providers to prolong your life as much as possible, even if it means you may experience greater pain and suffering.

The disclosure for the postinformed group was the same, except that participants in this group were told that that they had been defaulted rather than would be defaulted.

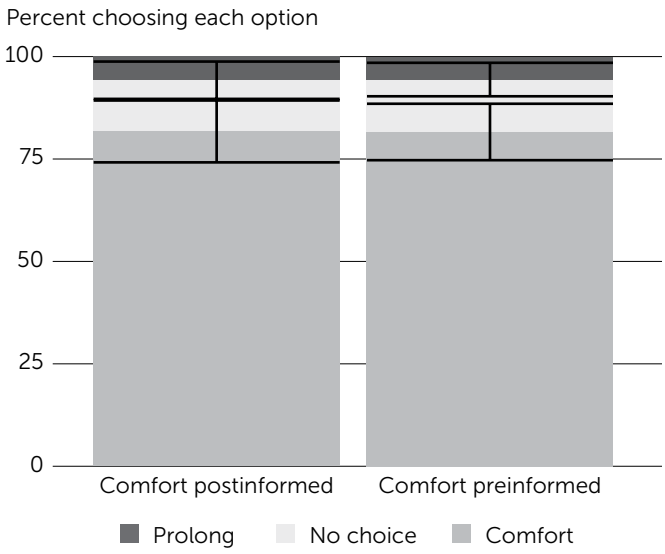
Capturing Effects from Disclosing Nudges

A detailed description of the results and our analyses of those data are available online in this article's Supplemental Material. Here we summarize our most pertinent findings, which are presented numerically in Table 2 and depicted visually in Figures 1 and 2.

Participants showed an overwhelming preference for minimizing discomfort at the end of life rather than prolonging life, especially for the general directives (see Figure 1). When the question was posed in general terms, more than 75% of responses reflected this general goal in all experimental conditions and both choice stages. By comparison, less than 15% of responses selected the goal of prolonging life, with the remaining participants leaving that decision to someone else.

Preferences for comfort in the general directive were so fixed that they were not affected by defaults or disclosure of defaults (that is, choices did not differ by condition in Figure 1). We note that these results

Figure 1. The impact of defaults on overall goal for care



Error bars are included to indicate 95% confidence intervals. The bars display how much variation exists among data from each group. If two error bars overlap by less than a quarter of their total length (or do not overlap), the probability that the differences were observed by chance is less than 5% (i.e., statistical significance at $p < .05$).

Table 2. Percentage choosing goal and treatment options by stage, default, and condition

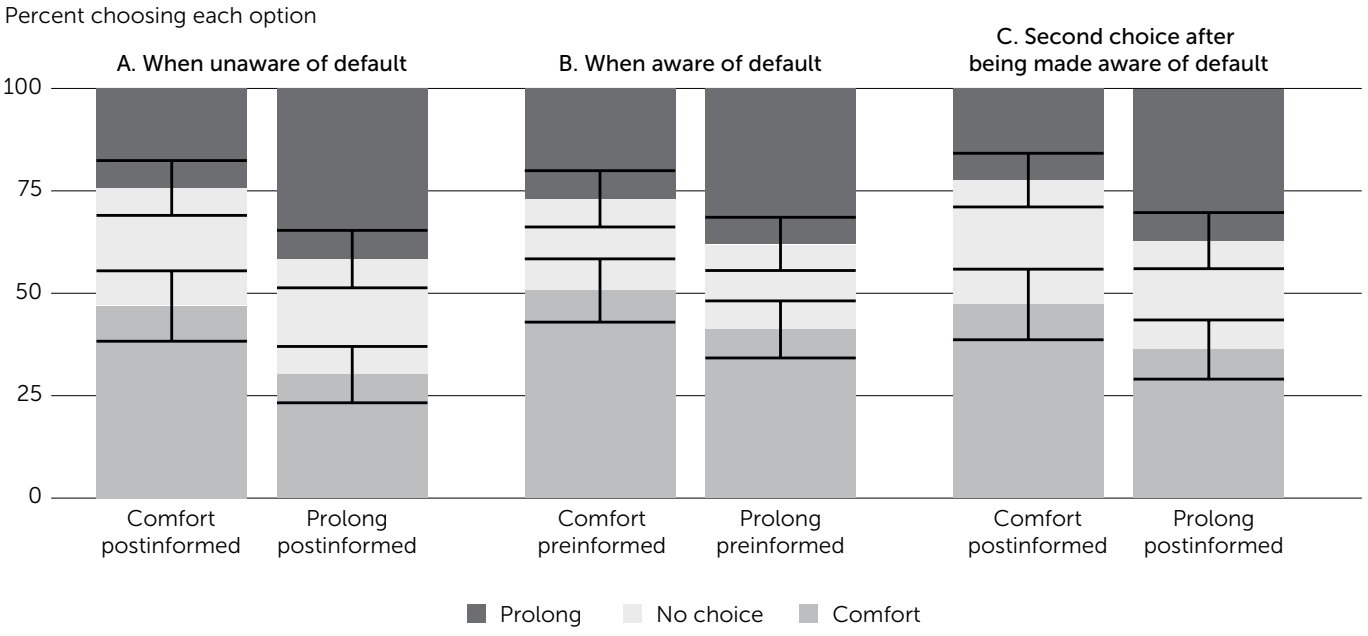
		Choice 1				Choice 2			
		Comfort default		Prolong default		Comfort default		Prolong default	
Question	Choice	Pre-informed	Post-informed	Pre-informed	Post-informed	Pre-informed	Post-informed	Pre-informed	Post-informed
Overall goal	Choose comfort	81.6%	81.7%	80.5%	78.2%	76.0%	76.9%	79.7%	79.8%
	Do not choose	12.8%	12.5%	7.5%	16.1%	12.8%	15.4%	7.5%	14.5%
	Choose prolong	5.6%	5.8%	12.0%	5.6%	11.2%	7.7%	12.8%	5.6%
Average of 5 specific treatments	Choose comfort	50.7%	46.9%	41.2%	30.2%	53.8%	47.3%	45.4%	36.3%
	Do not choose	22.4%	28.8%	20.9%	28.2%	24.6%	30.4%	22.1%	26.6%
	Choose prolong	26.9%	24.2%	37.9%	41.6%	21.6%	22.3%	32.5%	37.1%

differ from recent work using real advance directives[7] in which defaults had a large impact on participants' general goals. One possible explanation is that the highly educated respondents in our study had more definitive preferences about end-of-life care than did the less educated population from the earlier article.

Unlike the results for general directives, defaults for specific treatments, when the participant is only informed after the fact, are effective (see Figure 2A in Figure 2). We could observe this after averaging across the five specific interventions that participants considered: On this combined measure, 46.9% of participants who were given the comfort default (but not informed about it in advance) expressed a preference for comfort. By comparison, only 30.2% of those given the prolong default (again with no warning about defaults) expressed a preference for comfort (a difference of 17 percentage points, or 36% [17/46.9]).

The main purpose of the study was to examine the impact on nudge effectiveness of informing people

Figure 2. The impact of default on responses to specific treatments



Error bars are included to indicate 95% confidence intervals. The bars display how much variation exists among data from each group. If two error bars overlap by less than a quarter of their total length (or do not overlap), the probability that the differences were observed bychance is less than 5% (i.e., statistical significance at $p < .05$).

that they were being nudged, a question that is best addressed by analyzing the effects of preinforming people about directive choices. Figure 2B presents the impact of the default when people were preinformed. As can be seen in the figure, preinforming people about defaults weakened but did not wipe out their effectiveness (see Figure 2B). When participants completed the advance directive after being informed about the impact of the defaults, 50.7% of participants given the comfort default expressed a preference for comfort, compared with only 41.2% of those given the prolong life default (a difference of 10 percentage points, or 19%). Although all specific treatment choices were affected by the default in the predicted direction, the effect is statistically significant only for a single item (dialysis) and for the average of all five items (see the Supplemental Material). Preinforming participants about the default may have weakened its impact, but did not eliminate the default's effect.

Postinforming people that they have been defaulted and then asking them to choose again in a neutral way, with no further nudge, produces a substantial default effect that is not much smaller than the standard default effect, as seen in Figure 2C. When participants completed the advance directive a second time (this time without a default), having been informed after the fact that they had been defaulted, 47.3% of participants given the comfort default expressed a preference for comfort, compared with only 36.3% of those given the prolong life default (a difference of 11 percentage points, or 23%). Again, postinforming participants about the default and allowing them to change their decision may have weakened its impact, but did not eliminate the default's effect.

These results are important because they suggest that either a preinforming or a postinforming strategy can be effective in both disclosing the presence of a nudge and preserving its effectiveness. In addition, the results provide a conservative estimate of the power of defaults because all respondents who were informed at either stage had, by the second stage, been informed both that they had been randomly selected to be defaulted and that others had been randomly selected to receive alternative defaults. In addition, the second-stage advance directives did not include defaults, so any effect of defaults reflects a carryover effect from the first-stage choice. (More detailed analysis of our results and more information listed by specific treatments are available in the online Supplemental Material.)

Defaults Survive Transparency

Despite extensive research questioning whether advance directives have the intended effect of improving quality of end-of-life care,[17,18] they continue to be one of the few and major tools that exist to promote this goal. Combining advance directives with default options could steer people toward the types of comfort options for end-of-life care that many experts recommend and that many people desire for themselves. This study suggests such defaults can be transparently implemented, addressing the concerns of many ethicists without losing defaults' effectiveness.

More broadly, our findings demonstrate that default options are a category of nudges that can have an effect even when people are aware that they are in play. Our results are conservative in two ways. First, not only were respondents informed that they were about to be or had been defaulted, but they also learned that other participants received different defaults, thereby eliminating any implicit recommendation in the default. Given that the nudge continued to have an impact, we can only conjecture that the default effect would have been even more persistent if the warning informed them that they had been defaulted deliberately to the choice that policymakers believe is the best option.

Second, our results are conservative in the sense that the second advance directive that participants completed contained no defaults, so the effect of the initial default had to carry over to the second choice. Our experimental design minimized the added cost of switching: Regardless of whether they wanted to switch, respondents had to provide a second set of responses. Presumably, the impact of the initial default would have been even stronger if switching had required more effort for respondents than sticking with their original response.

What exactly produced the carryover effect remains uncertain. It is possible, and perhaps most interesting, that the prior default led respondents to think about the choice in a different way, specifically in a way that reinforced the rationality of the default they were presented with (consistent with reference 16). It is, however, also possible that the respondents were mentally lazy and declined to exert effort to reconsider

their previous decisions.

Although the switching costs in our study design were small, such costs may explain why we observed default effects for the specific items but not for the overall goal for care. If respondents were sufficiently concerned about representing their preferences accurately for their overall goal item, they may have been willing to engage in the mental effort to overcome the effect of the default. Finally, it is possible that the carryover from the defaults of stage 1 to the (default-free) responses in stage 2 reflected a desire for consistency.[19] If so, then carryover effects would be weaker in real-world contexts involving important decisions. If the practice of informing people that they were being defaulted became widespread, moreover, it is unlikely that either of these default-weakening features would be common. That is because defaults would not be chosen at random and advance directives would be filled out only once, with a disclosed default.

Despite our results, it would be premature to conclude that the impact of nudges will always persist when people are aware of them. Our findings are based on hypothetical advance directives—an appropriate first step in research given both the ethical issues involved and the potential repercussions for choices made regarding preferences for medical care at the end of life. Before embracing the general conclusion that warnings do not eliminate the impact of defaults, further research should examine different types of alerts across different settings. Given how weakly defaults affected overall goals for care in this study, it would especially be fruitful to examine the impact of pre- or postinforming participants in areas in which defaults are observed to have robust impact in the absence of transparency. Those areas include decisionmaking regarding retirement savings and organ donation.

Most generally, our findings suggest that the effectiveness of nudges may not depend on deceiving those who are being nudged. This is good news, because policymakers can satisfy the call for transparency advocated in the House of Lords report[9] with little diminution in the impact of positive interventions. This could help ease concerns that behavioral interventions are manipulative or involve trickery.

author affiliation

Loewenstein and Hagmann, Department of Social and Decision Sciences, Carnegie Mellon University; Bryce, Graduate School of Public Health, University of Pittsburgh; Rajpal, Bethesda, Maryland. Corresponding author's e-mail: gl20@andrew.cmu.edu

supplemental material

- http://behavioralpolicy.org/supplemental-material
- Methods & Analysis

References

1. Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics, 116,* 1149–1187.
2. Johnson, E. J., & Goldstein, D. G. (2003, November 21). Do defaults save lives? *Science, 302,* 1338–1339.
3. Johnson, E. J., Hershey, J., Meszaros, J., & Kunreuther, H. (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk & Uncertainty, 7,* 35–53.
4. Acquisti, A., John, L., & Loewenstein, G. (2013). What is privacy worth? *Journal of Legal Studies, 42,* 249–274.
5. Kressel, L. M., & Chapman, G. B. (2007). The default effect in end-of-life medical treatment preferences. *Medical Decision Making, 27,* 299–310.
6. Kressel, L. M., Chapman, G. B., & Leventhal, E. (2007). The influence of default options on the expression of end-of-life treatment preferences in advance directives. *Journal of General Internal Medicine, 22,* 1007–1010.
7. Halpern, S. D., Loewenstein, G., Volpp, K. G., Cooney, E., Vranas, K., Quill, C. M., . . . Bryce, C. (2013). Default options in advance directives influence how patients set goals for end-of-life care. *Health Affairs, 32,* 408–417.
8. Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.
9. House of Lords, Science and Technology Select Committee. (2011). *Behaviour change* (Second report). London, United Kingdom: Author.
10. Bovens, L. (2008). The ethics of nudge. In T. Grüne-Yanoff & S. O. Hansson (Eds.), *Preference change: Approaches from philosophy, economics and psychology* (pp. 207–220). Berlin, Germany: Springer.
11. Wortman, C. B., & Brehm, J. W. (1975). Responses to uncontrollable outcomes: An integration of reactance theory and the learned helplessness model. *Advances in Experimental Social Psychology, 8,* 277–336.
12. Haisley, E., Volpp, K., Pellathy, T., & Loewenstein, G. (2012). The impact of alternative incentive schemes on completion of health risk assessments. *American Journal of Health Promotion, 26,* 184–188.
13. Halpern, S. D., Ubel, P. A., & Asch, D. A. (2007). Harnessing the power of default options to improve health care. *New England Journal of Medicine, 357,* 1340–1344.
14. Johnson, E. J., & Goldstein, D. (2004). Default donation decisions. *Transplantation, 78,* 1713–1716.
15. McKenzie, C. R., Liersch, M. J., & Finkelstein, S. K. (2006). Recommendations implicit in policy defaults. *Psychological Science, 17,* 414–420.
16. Davidai, S., Gilovich, T., & Ross, L. D. (2012). The meaning of default options for potential organ donors. *PNAS: Proceedings of the National Academy of Sciences, USA, 109,* 15201–15205.
17. Writing Group for the SUPPORT Investigators. (1995, November 22). A controlled trial to improve care for seriously ill hospitalized patients: The Study to Understand Prognoses and Preferences for Outcomes and Risks of Treatments (SUPPORT). *JAMA, 274,* 1591–1598.
18. Fagerlin, A., & Schneider, C. E. (2004). Enough: The failure of the living will. *Hastings Center Report, 34*(2), 30–42.
19. Falk, A., & Zimmerman, F. (2013). A taste for consistency and survey response behavior. *CESifo Economic Studies, 59,* 181–193.

Finding

Workplace stressors & health outcomes: Health policy for the workplace

Joel Goh, Jeffrey Pfeffer, & Stefanos A. Zenios

abstract. Extensive research focuses on the causes of workplace-induced stress. However, policy efforts to tackle the ever-increasing health costs and poor health outcomes in the United States have largely ignored the health effects of psychosocial workplace stressors such as high job demands, economic insecurity, and long work hours. Using meta-analysis, we summarize 228 studies assessing the effects of ten workplace stressors on four health outcomes. We find that job insecurity increases the odds of reporting poor health by about 50%, high job demands raise the odds of having a physician-diagnosed illness by 35%, and long work hours increase mortality by almost 20%. Therefore, policies designed to reduce health costs and improve health outcomes should account for the health effects of the workplace environment.

Confronting ever-rising health benefits costs, Stanford University in 2007 began a sustained effort to slow the growth of its medical bills. Seeking particularly to help its workforce prevent or better control lifestyle-related diseases such as type 2 diabetes, the university created an employee wellness program. The program included modest financial incentives for participation (approximately $500 per participant in 2014); annual health screenings; a health assessment and behavior questionnaire; and opportunities to participate in exercise, nutrition, and stress-reduction classes.

Although wellness programs are a common policy response to employee health issues, evidence for their effectiveness is mixed. One recent meta-analysis reported health care cost savings of more than $3 for every $1 invested,[1] but an analysis at the University of Minnesota found no evidence that a lifestyle management program reduced health care costs.[2] According to a 2013 RAND Corporation report,[3] about half of all U.S. employers with 50 or more employees now offer some form of wellness promotion program. Although the RAND report, consistent with other empirical evidence,[4,5] noted some effects of these programs on lifestyle choices such as diet and exercise, the study reported that fewer than half of employees in workplaces offering wellness programs participated in them, in part because of rigid work schedules. The RAND report also contained separate case studies of five large U.S. employers. Using the data from these case studies,

Goh, J., Pfeffer, J., & Zenios, S. A. (2015). Workplace stressors & health outcomes: Health policy for the workplace. *Behavioral Science & Policy, 1*(1), pp. 43–52.

the authors of the report found that the average difference in health care costs between people who participated in such programs and those who did not was just $157 annually, an amount that is neither substantively nor statistically significant.

Why might such policy interventions not consistently show better results? One answer could be variation in services. Some programs include financial incentives to achieve specific biometric goals, whereas others do not. Some programs include health-related activities such as exercise and yoga classes, whereas others include only the assessments. There are also important differences in the workplace cultures in which such programs are implemented. For example, some companies emphasize employee well-being as a source of competitive advantage, whereas others push employee cost reduction. These different cultures and program elements could produce different health outcomes.[6]

But another possibility is that with their focus on individual behavior, wellness interventions miss an important factor affecting people's health: the work environment. Management practices in the workplace can either produce or mitigate stress related to long working hours, heavy job demands, an absence of job control, a lack of social support, and pervasive work–family conflict. More than 30% of respondents to a Stanford survey, for instance, reported that they experienced stress at work of sufficient severity to adversely affect their health.[7]

It is scarcely news that stress negatively affects health both directly[8,9] and indirectly through its influence on individual behaviors such as alcohol abuse, smoking, and drug consumption.[10–14] There is also recognition that stress produced in the workplace is related to numerous health outcomes, including increased risks of cardiovascular disease, depression, and anxiety. The physiological pathways through which some of these effects operate have been demonstrated.[15] Work contexts matter for health.[16]

Nonetheless, U.S. employers and policymakers have paid scant attention to the connections between workplace conditions and health. There has been somewhat more policy attention in Europe. Many European countries have laws that seek to more stringently regulate work hours, promote employment stability, and reduce work–family conflict.[17]

In the United States, the role of the work environment in workers' health has gained some attention through research sponsored by the National Institute for Occupational Safety and Health.[18] Nevertheless, most policy discussions and resources remain devoted to the relatively narrow objectives of promoting physical workplace safety (for example, reducing exposure to harmful chemicals) and offering health-promotion activities. Although both focuses are important, employers and policymakers have not sufficiently considered broader dimensions of the workplace environment that are affected by employer decisions and that impact the psychological and social well-being of employees—choices concerning layoffs, work hours, flexibility, and medical insurance benefits, for example.

Sustained policy attention to such issues will almost certainly require (a) assessing the relative size and importance of the health effects of various workplace conditions, (b) collecting data to enable regular analysis of the relationship between workplace conditions and health, and (c) reporting the incidence of exposure to unhealthy workplace conditions. It is almost impossible to overstate how the detailed reporting of job-related physical injury and death rates stimulated both policy attention and consistent improvement in physical working conditions over time.

In this article, we quantitatively review the extensive evidence on the connections between workplace stressors and health outcomes. Our results suggest that many workplace conditions profoundly affect human health. In fact, the effect of workplace stress is about as large as that of secondhand tobacco smoke, an exposure that has generated much policy attention and efforts to prevent or remediate its effects.

Why Health and Health Costs Are Important

The United States spends a higher proportion of its gross domestic product on health care than do other advanced industrialized economies and about twice as much per capita as 15 other rich industrialized nations. The United States has also experienced a higher growth rate in health care spending than other countries.[19] But despite higher U.S. health care spending, life expectancy is lower and infant mortality is higher than in countries that spend far less on health care, including Japan, Sweden, and Switzerland. According to 2013 data, the United States ranks 26th in life expectancy, below the average of member countries that make up the Organisation for Economic Co-operation and

Development, which are mostly high-income, developed nations.[20]

Health matters to individuals, to their employers, and to governments. Poor health takes a heavy toll on sick individuals and their families in many ways, including financially. One study reported that in 2001, almost half of all bankruptcies were related to medical bills; by 2007, that proportion had grown to 62%.[21] Other studies have found that even people with health insurance face increasing financial stress from health care costs.[22]

Employers care about health costs. They pay a significant portion of Medicare and Medicaid taxes and more than half of private health insurance premiums.[23] Ever-growing health care bills constrain employers' ability to offer raises, hire additional people, and make the capital investments necessary for long-term growth.

Governments likewise worry about the ever-increasing share of their budgets that is diverted away from other public purposes and toward health costs for both active employees and retirees.[24] Still, many people reasonably believe that a healthy and long life is a fundamental human right.[25]

The Health Effects of Workplace Stressors

Analyzing Workplace Stressors

We examined the effect of workplace stressors on health through an analytical procedure known as *meta-analysis,* which statistically summarizes the results of multiple studies. We identified these studies by what is known as a *systematic literature review,* in which we searched public scientific databases for research articles that contained keywords such as *work hours, overtime, job control, job security,* and *layoff,* among others (details are provided in the Supplemental Material). We used predefined criteria to winnow the list of studies down to a smaller set of relevant studies. This procedure is widely accepted as a way of minimizing researchers' biases in searching for the studies to include in a review.

Authors of numerous reviews and meta-analyses have examined the health effects of individual workplace stressors such as job insecurity,[26–28] long work hours,[29,30] lack of social support in the workplace,[31] and psychological demands and job discretion.[32–34] Narrative reviews (that is, reviews that do not use systematic procedures of study selection) have revealed consistent evidence in the literature that work stress is associated with a variety of negative health outcomes, including cardiovascular disease, clinical depression, and death.[15] However, to our knowledge, no researcher has used common meta-analysis methods and criteria to investigate the health effects of a fairly comprehensive set of workplace stressors, something that is necessary to estimate the relative importance of various workplace conditions for health. We perform such a meta-analysis by analyzing the effects of 10 different stressors on four health outcomes, thus allowing policymakers to weigh the magnitude of each stressor's effects.

Our objective was to analyze work stressors that affect people's psychological and physical health and that can be reasonably addressed by either public policy or managerial interventions. We focused our analysis on single stressors rather than on composites because it is usually easier for employers or policymakers to address workplace problems individually than to tackle many at once. Also, minimizing individual stressors should naturally lessen the impact of any broader composite that includes those individual stressors.

We examined numerous workplace conditions presumed to undermine health: long working hours[35] and shift work;[36] work–family conflict;[37,38] job control, which refers to the level of discretion that employees have over their work;[39,40] and job demands.[41,42] The combination of these latter two stressors is referred to as *job strain*.[43] We also examined workplace conditions that might mitigate the negative effects of job stressors. These included social support and social networking opportunities;[44,45] organizational justice, which refers to the perceived level of fairness in the workplace;[46] and availability of health insurance, which affects access to health care and preventive screenings and, therefore, mortality.[47] Finally, we assessed what may be the most important factor of all: whether a person is employed at all. Research consistently finds that layoffs, job loss, and unemployment all have important effects on health,[48,49] as does economic insecurity.[50] Although macroeconomic conditions that are beyond the control of an employer undoubtedly influence this last stressor, the ultimate decision to lay off employees and thereby increase not only that individual's economic insecurity but the insecurity of others, including people who retain their jobs but see those jobs as being at risk, resides with the employer.

Our next step was to identify important health

outcomes. We focused on four outcomes typically used in studies of the health effects of the work environment: the presence of a diagnosed medical condition, a person's perception of being in poor physical health, a person's perception of having poor mental health, and death. Regardless of how these outcomes are measured, researchers usually classify them in an either–or way—for example, a person's health is either "poor" or "good." Studies repeatedly have shown that people's perception of their own health status—even when measured by a single survey question such as "How would you say your health in general is?"—significantly predicts the likelihood of subsequent illness and risk of death. That is true even when other health-relevant predictors such as marital status and age are taken into account.[51,52] Moreover, the predictive value of single-item measures of self-reported health holds across various ethnicities[53] and age groups.[54]

Our initial search yielded 741 studies that examined health effects of workplace conditions in some way. However, about two-thirds of those did not meet our criteria for inclusion in the meta-analysis—for example, because they were review articles or had too small a study sample. Our final sample included 228 studies. All 228 studies had sample sizes larger than 1,000, and 115 of them followed subjects over a period of time, so that researchers could relate workplace stressors to later health outcomes. (We furnish further details of our study selection criteria, meta-analytic methods, and statistical techniques in the online Supplemental Material, including a description of the analyses we conducted to ensure that our results were robust and that our estimates of effect sizes were not unduly inflated because of publication bias, the phenomenon in which positive and statistically significant results are more likely to get published.)

Increased Odds of Poor Health Outcomes

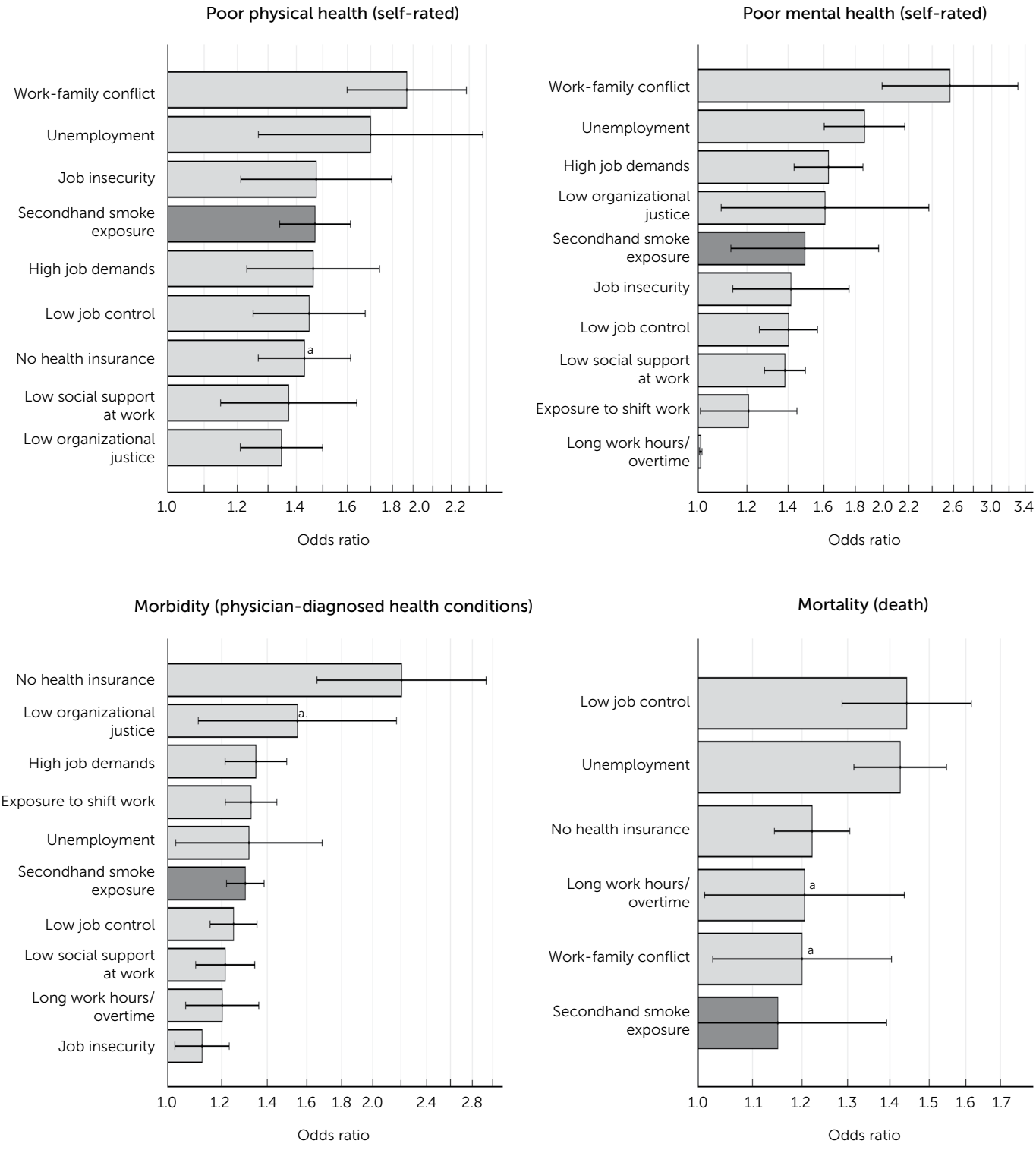
The four panels of Figure 1 show the statistically significant effects that work stressors had on the four categories of health outcomes: self-rated poor health, self-rated poor mental health, physician-diagnosed health conditions, and death. The sizes of these effects are presented as odds ratios, a statistical concept that may be new to some readers. An odds ratio conveys how the presence of one factor increases the odds of another factor being present. More concretely, the odds ratios in our study capture the extent to which individual workplace stressors increased the odds of having negative health outcomes. Knowing the scale helps make sense of these ratios. An odds ratio of 1 means an exposure produces no change in the odds of a negative health outcome occurring. An odds ratio of 2 means a stressor doubles the odds of a negative health outcome.

Odds ratios offered in isolation can be difficult to interpret. Therefore, to better convey the sizes of the effects we calculated, we compare them with something familiar to many: negative health outcomes from exposure to secondhand tobacco smoke. The odds ratios we found in the research literature on the effects of secondhand smoke were 1.47 for self-reported poor health.[55] In other words, exposure to secondhand tobacco smoke increases the odds that a person rates his or her general health as poor by almost 50%. In addition, odds ratios on the effects of exposure to secondhand smoke were 1.49 for self-reported mental health problems,[56] 1.30 for physician-diagnosed medical conditions,[57] and 1.15 for mortality.[58,59] (Although the biological pathway for the effect of secondhand smoke on mental health is less well established than it is for the other outcomes, some animal studies suggest that tobacco smoke can directly induce negative mood.[60])

The health effects of secondhand smoke exposure are widely viewed as sufficiently large to warrant regulatory intervention. For example, secondhand smoke is recognized as a carcinogen,[61] and smoking in enclosed public places, including workspaces, is banned in many states in the United States and in many other countries. The results of our meta-analysis show that workplace stressors generally increased the odds of poor health outcomes to approximately the same extent as exposure to secondhand smoke. These results support several conclusions:

- Unemployment and low job control have significant associations with all of the health outcomes, as does an absence of health insurance for those outcomes for which there are sufficient numbers of studies. With the exception of work–family conflict, all of the work stressors we examined are significantly associated with an increased likelihood of developing a medical condition, as diagnosed by a doctor.

Figure 1. Comparing health effects from work stressors to secondhand smoke exposure



Odds ratios higher than 1 indicate that the exposures listed here increased the odds of negative health outcomes. No health insurance, for instance, increased the odds of a physician-diagnosed health condition by more than 100%. Odds ratios for exposures marked with "a" were calculated with two or fewer studies and may be less reliable. Error bars are included to indicate standard errors. These bars indicate how much variation exists among data from each group. If two error bars are separated by at least half the width of the bars, this indicates less than a 5% probability that a difference was observed by chance (i.e., statistical significance at $p < .05$).

- Psychological and social aspects of the work environment, such as a lack of perceived fairness in the organization, low social support, work–family conflict, and low job control, are associated with health as strongly as more concrete aspects of the workplace, such as exposure to shift work, long work hours, and overtime.
- The association between workplace stressors and health is strong in many instances. For example, work–family conflict increases the odds of self-reported poor physical health by about 90%, and low organizational justice increases the odds of having a physician-diagnosed condition by about 50%.

Similar to the health effects of secondhand tobacco smoke, the effects of workplace practices are larger for self-reported physical and mental health and for physician-diagnosed illness than for mortality. This finding is not unexpected. Group differences in mortality rates typically take longer than other health outcomes to emerge, and therefore, other intervening factors that contribute to the hazard of mortality can dilute the effect of workplace stressors. Also, because of the longer time periods over which mortality effects occur, they are especially prone to bias because people who are sicker are more likely to drop out of the workforce (and therefore also out of the data set) during the research. Once individuals are out of the workforce, people also face a lower cumulative exposure to workplace stressors. Both of these factors could lead to an underestimation of effect sizes for mortality.

Policy Implications

Our primary conclusion that psychosocial work stressors are important determinants of health suggests several policy recommendations. First, if initiatives to improve employee health are to be effective, they cannot simply address health behaviors, such as reducing smoking and promoting exercise, but should also include efforts to redesign jobs and reduce or eliminate the workplace practices that contribute to workplace-induced stress.[62] For example, possible job redesigns could involve limiting working hours, reducing shift work and unpredictable working hours, and encouraging flexible work arrangements that help employees to achieve a better balance between their work life and their family life. A detailed discussion of interventions to prevent and remediate workplace stressors is beyond the scope of this article. We refer interested readers to a recent review[63] or RAND Europe report[64] for discussions of specific workplace intervention strategies.

We also recommend that greater effort be put forth to gather data on these workplace stressors and their health effects at both the national and the organizational levels of analysis. Despite the long-recognized and important health effects of workplace conditions, we are not aware of any nationally representative longitudinal data set in the United States that contains individual-level data on both workplace stressors and health outcomes. Such an effort would likely require (and benefit from) the involvement of government agencies that have interests in promoting worker or population health, such as the Occupational Safety and Health Administration or the Agency for Healthcare Research and Quality. In constructing such a data set, care should be taken to assess the exposures to these stressors at different points in time so that the cumulative exposure to stressors can be measured.

Organizations seeking to improve the health of their employees (and thereby reduce their health costs) need to have a complete picture of the work environment by assessing the prevalence of workplace stressors. Therefore, employers should measure both management practices and the workplace environment as well as employee health over time. This would permit employers to assess the effectiveness of any interventions, which they can do easily through self-rated health measures that are known to be effective proxies for actual health.

Because resources are limited and policymakers have to be selective about which stressors to target, our results can be used to identify where to focus attention. A simple way to do this would be to look at the effect sizes (odds ratios) from our analysis. Clearly, all else being equal, stressors with larger effect sizes contribute more toward poorer health outcomes. However, a more complete analysis should also incorporate two other pieces of information that are specific to the population in question: the rate of occurrence for each exposure and the baseline prevalence of each health outcome within that population.

To understand why these other two rates are important, consider a hypothetical example in which an exposure almost never occurs in a target population. Also consider another example in which the health outcome itself is so rare that any proportionate increase in its prevalence is insignificant in terms of raw numbers. In either case, even if the exposure has a large effect size on the outcome, the overall health impact of the exposure would be minimal in the study population as a whole. Therefore, in general, a stressor would have a large health impact in a population (and therefore represent a good candidate for policy attention) if (a) it has a high occurrence rate, (b) it has a large effect size on some health outcome, and (c) that health outcome also occurs with high baseline prevalence.

In another article,[65] we detailed how these pieces of information can be combined to generate new policy insights. In particular, we used data from the General Social Survey and the Current Population Survey to estimate the prevalence of workplace stressors in the United States and data from the Medical Expenditure Panel Survey and Vital Statistics Reports to estimate the prevalence of the negative health outcomes and their associated costs. We then combined these data through a mathematical model to estimate the annual excess mortality and costs that can be attributed to workplace stressors in the United States. Our analysis suggests that measures of workplace stressors can provide valuable information for insurers or employers who wish to perform more accurate risk adjustment and risk assessment. Of course, for this to be feasible, employers or insurers must first collect data on these aspects of the work environment.

Finally, given the pernicious health effects of workplace stressors, we recommend that policymakers consider increasing regulatory oversight of work conditions. Although some stressors—such as long work hours and shift work (through wage and hour laws and overtime rules)—are already subject to regulation (although there is some debate about the extent of the enforcement of these rules), other stressors could be fruitful avenues for attention. For example, employers could receive tax incentives if they offer work arrangements that support work–family balance and thereby minimize work–family conflict or, as in many European countries, incentives that would encourage more employment continuity and fewer layoffs. Any intervention in the labor market entails trade-offs, and we are not advocating a simplistic approach that focuses on health effects at the expense of other considerations. However, the lack of policy attention to psychological and social aspects of the workplace environment leaves many avenues for addressing health and health care costs untouched.

Furthermore, a host of nonregulatory actions can be taken to combat workplace stress. For example, policymakers could publish guidelines or best practices that could help raise awareness among employers and workers about the links between work stressors and health. Agencies or industry associations could encourage employers to take actions to help mitigate workplace stress and its causes. Similar actions have already been taken in the European Union,[17] where the European Framework Agreement on Work-Related Stress has led to concrete actions including "training, stress barometers, assessment tools for establishments . . . or general surveys to gather data and raise awareness."[66]

Limitations and Future Research

Our study's primary limitation is that all of the studies in our meta-analysis were observational (and not randomized controlled trials), which prevents us from making a strong causal inference linking workplace stressors to poor health outcomes. Furthermore, about half of the studies used cross-sectional designs, which are prone to biases from reverse causality. That is, these studies measured stressors in the same time window during which outcomes were measured, and the strength of associations could potentially be driven by poor health causing work stressors instead of work stressors causing poor health. Therefore, our results do not conclusively establish that these stressors cause poor health. Instead, they show that work stressors are strongly associated with poor health and suggest that these stressors could be fruitful targets for policy attention.

A second limitation is that our results represent averaged effect sizes. People will inevitably differ with respect to how each stressor affects each health outcome because they have different coping mechanisms and also differ in how they respond to workplace stress—for example, whether they believe that stress

has fundamentally positive or negative consequences.[67] The studies in our sample did not survey subjects on their attitudes toward stress, so we were not able to estimate the effects that different stress attitudes have on the results. Future researchers should assess how differential psychological beliefs about workplace stress affect the health effects of work stressors.

A final limitation of our study is that we focused exclusively on simple stressors that can be reasonably addressed by interventions. Consequently, we omitted work stressors such as effort–reward imbalance and job strain even though some studies suggest both of these stressors may have significant health effects,[43,68,69] perhaps with even larger odds ratios than we found in the studies we examined in this article. This limitation underscores a broader question that future researchers should address: Because many different and (at least partially) overlapping factors contribute to work stress, how do researchers assess the health effects of the totality of the work experience and design appropriate policies to cost-effectively increase employee health and productivity and reduce health care costs?

More than 100 years ago, after Upton Sinclair's book *The Jungle*[70] exposed dangerous conditions in meat-packing plants, public policy and voluntary company behavior began focusing on reducing occupational injuries and deaths, to great success. Although the dangers emanating from the psychological and social conditions of work are not as visible, they can also be quite harmful to health. Unless and until companies and governments more rigorously measure and intervene to reduce harmful workplace stressors, efforts to improve people's health—and their lives—and reduce health care costs will be limited in their effectiveness.

author affiliation

Goh, Harvard Business School; Pfeffer & Zenios, Graduate School of Business, Stanford University. Corresponding author's e-mail: jgoh@hbs.edu

author note

We are grateful to Ed Kaplan and Scott Wallace for their feedback on an earlier version of this article. We also thank the senior disciplinary editor, Adam Grant; the associate disciplinary editor and associate policy editor; and three referees for their helpful comments and suggestions. The collective feedback helped us improve this article significantly.

supplemental material

- http://behavioralpolicy.org/supplemental-material
- Data, Analyses & Results
- Additional Figures
- Additional References

References

1. Baicker, K., Cutler, D., & Song, Z. (2010). Workplace wellness programs can generate savings. *Health Affairs, 29,* 304–311.
2. Nyman, J. A., Abraham, J. M., Jeffery, M. M., & Barleen, N. A. (2012). The effectiveness of a health promotion program after 3 years: Evidence from the University of Minnesota. *Medical Care, 50,* 772–778. http://dx.doi.org/10.1097/MLR.0b013e31825a8b1f
3. Mattke, S., Liu, H., Caloyeras, J. P., Huang, C. Y., Van Busum, K. R., Khodyakov, D., & Shier, V. (2013). *Workplace wellness programs study: Final report* (No. RR-254-DOL). Santa Monica, CA: RAND Corporation.
4. Hochart, C., & Lang, M. (2011). Impact of a comprehensive worksite wellness program on health risk, utilization, and health care costs. *Population Health Management, 14,* 111–116. http://dx.doi.org/10.1089/pop.2010.0009
5. Milani, R. V., & Lavie, C. J. (2009). Impact of worksite wellness intervention on cardiac risk factors and one-year health care costs. *American Journal of Cardiology, 104,* 1389–1392. http://dx.doi.org/10.1016/j.amjcard.2009.07.007
6. Caloyeras, J. P., Liu, H., Exum, E., Broderick, M., & Mattke, S. (2014). Managing manifest diseases, but not health risks, saved PepsiCo money over seven years. *Health Affairs, 33,* 124–131. http://dx.doi.org/10.1377/hlthaff.2013.0625
7. Stanford University, BeWell Program. (2011). *BeWell@Stanford 2011 annual report.* Retrieved from https://bewell.stanford.edu/sites/default/files/2011BeWellAnnualReport_0.pdf
8. Chandola, T., Brunner, E., & Marmot, M. (2006, March 4). Chronic stress at work and the metabolic syndrome: Prospective study. *British Medical Journal, 332,* 521–525. http://dx.doi.org/10.1136/bmj.38693.435301.80
9. Kivimäki, M., Leino-Arjas, P., Luukkonen, R., Riihimäi, H., Vahtera, J., & Kirjonen, J. (2002, October 19). Work stress and risk of cardiovascular mortality: Prospective cohort study of industrial employees. *British Medical Journal, 325,* 857–861. http://dx.doi.org/10.1136/bmj.325.7369.857
10. Harris, M., & Fennell, M. (1988). Perceptions of an employee assistance program and employees' willingness to participate. *Journal of Applied Behavioral Science*, *24,* 423–438.
11. Kouvonen, A., Kivimäki, M., Virtanen, M., Pentti, J., & Vahtera, J. (2005). Work stress, smoking status, and smoking intensity: An observational study of 46,190 employees. *Journal of Epidemiology and Community Health, 59,* 63–69. http://dx.doi.org/10.1136/jech.2004.019752
12. Nishitani, N., & Sakakibara, H. (2006). Relationship of obesity to job stress and eating behavior in male Japanese workers. *International Journal of Obesity, 30,* 528–533. http://dx.doi.org/10.1038/sj.ijo.0803153
13. Piazza, P. V., & Le Moal, M. (1998). The role of stress in drug self-administration. *Trends in Pharmacological Sciences, 19,* 67–74. http://dx.doi.org/10.1016/S0165-6147(97)01115-2
14. Wardle, J., Steptoe, A., Oliver, G., & Lipsey, Z. (2000). Stress, dietary restraint and food intake. *Journal of Psychosomatic Research, 48,* 195–202. http://dx.doi.org/10.1016/S0022-3999(00)00076-3
15. Ganster, D. C., & Rosen, C. C. (2013). Work stress and employee health: A multidisciplinary review. *Journal of Management, 39,* 1085–1122. http://dx.doi.org/10.1177/0149206313475815
16. Heaphy, E. D., & Dutton, J. E. (2008). Positive social interactions and the human body at work: Linking organizations and physiology. *Academy of Management Review, 33,* 137–162. http://dx.doi.org/10.5465/AMR.2008.27749365
17. Monks, J., de Buck, P., Benassi, A., & Plassmann, R. (2008). *Implementation of the European autonomous framework agreement on work-related stress.* Brussels, Belgium: European Social Partners.
18. National Institute for Occupational Safety and Health. (2012). *The research compendium: The NIOSH Total Worker Health Program: Seminal research papers 2012* (DHHS Publication No. 2012-146). Washington, DC: U.S. Department of Health and Human Services, Center for Disease Control and Prevention, National Institute for Occupational Safety and Health.
19. Kaiser Family Foundation. (2011). *Snapshots: Health care spending in the United States & selected OECD countries.* Retrieved from http://kff.org/health-costs/issue-brief/snapshots-health-care-spending-in-the-united-states-selected-oecd-countries/
20. Organisation for Economic Co-operation and Development. (2013). *Health at a glance 2013: OECD indicators.* Retrieved from http://dx.doi.org/10.1787/health_glance-2013-en
21. Himmelstein, D. U., Thorne, D., Warren, E., & Woolhandler, S. (2009). Medical bankruptcy in the United States, 2007: Results of a national study. *American Journal of Medicine, 122,* 741–746. http://dx.doi.org/10.1016/j.amjmed.2009.04.012
22. Banthin, J. S., Cunningham, P., & Bernard, D. M. (2008). Financial burden of health care, 2001–2004. *Health Affairs, 27,* 188–195. http://dx.doi.org/10.1377/hlthaff.27.1.188
23. Bureau of Labor Statistics. (2013). Employee benefits in the United States—March 2013 [News Release USDL-13-1344]. Retrieved from http://www.bls.gov/news.release/archives/ebs2_07172013.pdf
24. Fox, B. J., Taylor, L. L., & Yucel, M. K. (1993, Third Quarter). America's health care problem: An economic perspective. *Federal Reserve Bank of Dallas Economic Review*. Retrieved from http://www.dallasfed.org/assets/documents/research/er/1993/er9303b.pdf
25. Marmot, M., Allen, T., Bell, R., & Goldblatt, P. (2012, January 14). Building the global movement for health equity: From Santiago to Rio and beyond. *Lancet, 379,* 181–188. http://dx.doi.org/10.1016/S0140-6736(11)61506-7
26. Sverke, M., Hellgren, J., & Näswall, K. (2002). No security: A meta-analysis and review of job insecurity and its consequences. *Journal of Occupational Health Psychology, 7,* 242–264. http://dx.doi.org/10.1037/1076-8998.7.3.242
27. Virtanen, M., Kivimäki, M., Joensuu, M., Virtanen, P., Elovainio, M., & Vahtera, J. (2005). Temporary employment and health: A review. *International Journal of Epidemiology, 34,* 610–622. http://dx.doi.org/10.1093/ije/dyi024
28. Virtanen, M., Nyberg, S. T., Batty, G. D., Jokela, M., Heikkilä, K., Fransson, E. I., . . . Kivimäki, M. (2013). Perceived job insecurity as a risk factor for incident coronary heart disease: Systematic review and meta-analysis. *British Medical Journal*, *347*, Article f4746. http://dx.doi.org/10.1136/bmj.f4746
29. Sparks, K., Cooper, C., Fried, Y., & Shirom, A. (1997). The effects of hours of work on health: A meta-analytic review. *Journal of Occupational and Organizational Psychology, 70,* 391–408. http://dx.doi.org/10.1111/j.2044-8325.1997.tb00656.x
30. Bannai, A., & Tamakoshi, A. (2014). The association between long working hours and health: A systematic review of epidemiological evidence. *Scandinavian Journal of Work and Environmental Health, 40,* 5–18. http://dx.doi.org/10.5271/sjweh.3388
31. Viswesvaran, C., Sanchez, J. I., & Fisher, J. (1999). The role of social support in the process of work stress: A meta-analysis. *Journal of Vocational Behavior, 54,* 314–334. http://dx.doi.org/10.1006/jvbe.1998.1661
32. Pieper, C., Lacroix, A. Z., & Karasek, R. A. (1989). The relation of psychosocial dimensions of work with coronary heart disease risk factors: A meta-analysis of five United States data bases. *American Journal of Epidemiology, 129,* 483–494.
33. Bonde, J. P. E. (2008). Psychosocial factors at work and risk

of depression: A systematic review of the epidemiological evidence. *Occupational and Environmental Medicine, 65,* 438–445. http://dx.doi.org/10.1136/oem.2007.038430
34. Kivimäki, M., Nyberg, S. T., Batty, G. D., Fransson, E. I., Heikkilä, K., Alfredsson, L., . . . Theorell, T. (2012, October 27). Job strain as a risk factor for coronary heart disease: A collaborative meta-analysis of individual participant data. *Lancet, 380,* 1491–1497. http://dx.doi.org/10.1016/S0140-6736(12)60994-5
35. Yang, H., Schnall, P. L., Jauregul, M., Su, T.-C., & Baker, D. (2006). Work hours and self-reported hypertension among working people in California. *Hypertension, 48,* 744–750. http://dx.doi.org/10.1161/01.HYP.0000238327.41911.52
36. Virkkunen, H., Härma, J., Kauppinene, T., & Tenkanen, L. (2006). The triad of shift work, occupational noise, and physical workload and risk of coronary heart disease. *Occupational and Environmental Medicine, 63,* 378–386. http://dx.doi.org/10.1136/oem.2005.022558
37. Frone, M. R. (2000). Work–family conflict and employee psychiatric disorder: The National Comorbidity Survey. *Journal of Applied Psychology, 85,* 888–895. http://dx.doi.org/10.1037/0021-9010.85.6.888
38. Frone, M. R., Russell, M., & Barnes, G. M. (1996). Work–family conflict, gender, and health-related outcomes: A study of employed parents in two community samples. *Journal of Occupational Health Psychology, 1,* 57–69. http://dx.doi.org/10.1037/1076-8998.1.1.57
39. Marmot, M. G., Rose, G., Shipley, M., & Hamilton, P. J. (1978). Employment grade and coronary heart disease in British civil servants. *Journal of Epidemiology and Community Health, 32,* 244–249. http://dx.doi.org/10.1136/jech.32.4.244
40. Marmot, M. G., Bosma, H., Hemingway, H., Brunner, E., & Stansfeld, S. (1997, July 26). Contribution of job control and other risk factors to social variations in coronary heart disease incidence. *Lancet, 350,* 235–239. http://dx.doi.org/10.1016/S0140-6736(97)04244-X
41. Shields, M. (2006). Stress and depression in the employed population. *Health Reports, 17*(4), 11–29.
42. Tsutsumi, A., Kayaba, K., Kario, K., & Ishikawa, S. (2009, January 12). Prospective study on occupational stress and risk of stroke. *Archives of Internal Medicine, 169,* 56–61. http://dx.doi.org/10.1001/archinternmed.2008.503
43. Karasek, R. A., Jr. (1979). Job demands, job decision latitude, and mental strain: Implications for job redesign. *Administrative Science Quarterly, 24,* 285–308.
44. Broadhead, W., Kaplan, B., James, S., Wagner, E., Schoenbach, V., Grimson, R., . . . Gehlbach, S. (1983). The epidemiological evidence for a relationship between social support and health. *American Journal of Epidemiology, 117,* 521–537.
45. Cohen, S., & Wills, T. A. (1985). Stress, social support, and the buffering hypothesis. *Psychological Bulletin, 98,* 310–357. http://dx.doi.org/10.1037/0033-2909.98.2.310
46. Robbins, J. M., Ford, M. T., & Tetrick, L. E. (2012). Perceived unfairness and employee health: A meta-analytic integration. *Journal of Applied Psychology, 97,* 235–272. http://dx.doi.org/10.1037/a0025408
47. Wilper, A. P., Woolhandler, S., Lasser, K. E., McCormick, D., Bor, D. H., & Himmelstein, D. U. (2009). Health insurance and mortality in US adults. *American Journal of Public Health*, *99,* 2289–2295. http://dx.doi.org/10.2105/AJPH.2008.157685
48. Eliason, M., & Storrie, D. (2009). Does job loss shorten life? *Journal of Human Resources, 44,* 277–302. http://dx.doi.org/10.3368/jhr.44.2.277
49. Strully, K. W. (2009). Job loss and health in the U.S. labor market. *Demography, 46,* 221–246. http://dx.doi.org/10.1353/dem.0.0050
50. Lee, S., Colditz, G. A., Berkman, L. F., & Kawachi, I. (2004). Prospective study of job insecurity and coronary heart disease in US women. *Annals of Epidemiology, 14,* 24–30. http://dx.doi.org/10.1016/S1047-2797(03)00074-7
51. Idler, E. L., & Benyamini, Y. (1997). Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior, 38,* 21–37.
52. Miilunpalo, S., Vuori, I., Oja, P., Pasanen, M. & Urponen, H. (1997). Self-rated health status as a health measure: The predictive value of self-reported health status on the use of physician services and on mortality in the working-age population. *Journal of Clinical Epidemiology, 50,* 517–528. http://dx.doi.org/10.1016/S0895-4356(97)00045-0
53. McGee, D. L., Liao, Y., Cao, G., & Cooper, R. S. (1999). Self-reported health status and mortality in a multiethnic US cohort. *American Journal of Epidemiology, 149,* 41–46.
54. Grant, M. D., Piotrowski, Z. H., & Chappell, R. (1995). Self-reported health and survival in the Longitudinal Study of Aging, 1984–1986. *Journal of Clinical Epidemiology, 48,* 375–387. http://dx.doi.org/10.1016/0895-4356(94)00143-E
55. Mannino, D. M., Siegel, M., Rose, D., Nkuchia, J., & Etzel, R. (1997). Environmental tobacco smoke exposure in the home and worksite and health effects in adults: Results from the 1991 National Health Interview Survey. *Tobacco Control, 6,* 296–305. http://dx.doi.org/10.1136/tc.6.4.296
56. Hamer, M., Stamatakis, E., & Batty, G. D. (2010). Objectively assessed secondhand smoke exposure and mental health in adults: Cross-sectional and prospective evidence from the Scottish Health Survey. *Archives of General Psychiatry, 67,* 850–855. http://dx.doi.org/10.1001/archgenpsychiatry.2010.76
57. Law, M. R., Morris, J. K., & Wald, N. J. (1997, October 18). Environmental tobacco smoke exposure and ischaemic heart disease: An evaluation of the evidence. *British Medical Journal, 315,* 973–980. http://dx.doi.org/10.1136/bmj.315.7114.973
58. Hill, S., Blakely, T., Kawachi, I., & Woodward, A. (2004, April 22). Mortality among "never smokers" living with smokers: Two cohort studies, 1981–4 and 1996–9. *British Medical Journal, 328,* 988–989. http://dx.doi.org/10.1136/bmj.38070.503009
59. Wen, W., Shu, X. O., Gao, Y.-T., Yang, G., Li, Q., Li, H., & Zheng, W. (2006, August 17). Environmental tobacco smoke and mortality in Chinese women who have never smoked: Prospective cohort study. *British Medical Journal, 333,* 376–379. http://dx.doi.org/10.1136/bmj.38834.522894.2F
60. Iñiguez, S. D., Warren, B. L., Parise, E. M., Alcantara, L. F., Schuh, B., Maffeo, M. L., . . . Bolaños-Guzmán, C. A. (2009). Nicotine exposure during adolescence induces a depression-like state in adulthood. *Neuropsychopharmacology, 34,* 1609–1624. http://dx.doi.org/10.1038/npp.2008.220
61. U.S. Environmental Protection Agency. (1992). *Respiratory health effects of passive smoking: Lung cancer and other disorders* (Report EPA/600/6-90/006F). Washington, DC: Author.
62. LaMontagne, A. D., Keegel, T., Louie, A. M., Ostry, A., & Landsbergis, P. A. (2007). A systematic review of the job-stress intervention evaluation literature, 1990–2005. *International Journal of Occupational and Environmental Health, 13,* 268–280. http://dx.doi.org/10.1179/oeh.2007.13.3.268
63. Landsbergis, P. A. (2009). Interventions to reduce job stress and improve work organization and worker health. In P. L. Schnall, M. Dobson, & E. Rosskam (Eds.), *Unhealthy work: Causes, consequences, cures* (pp. 193–209). Amityville, NY: Baywood.
64. van Stolk, C., Staetsky, L., Hassan, E., & Kim, C. W. (2012). *Management of psychosocial risks at work: An analysis of the findings of the European Survey of Enterprises on New and Emerging Risks (ESENER).* Luxembourg, Grand Duchy of Luxembourg: Publications Office of the European Union.
65. Goh, J., Pfeffer, J., & Zenios, S.A. (2015). The relationship

between workplace stressors and mortality and health costs in the United States. *Management Science*. Advance online publication. http://dx.doi.org/10.1287/mnsc.2014.2115
66. European Commission. (2011). *Report on the implementation of the European social partners' framework agreement on work-related stress* (SEC[2011] 241 Final). Brussels, Belgium: Author.
67. Crum, A. J., Salovey, P., & Achor, S. (2013). Rethinking stress: The role of mindsets in determining the stress response. *Journal of Personality and Social Psychology, 104,* 716–733. http://dx.doi.org/10.1037/a0031201
68. Siegrist, J. (1996). Adverse health effects of high-effort/low-reward conditions. *Journal of Occupational Health Psychology, 1,* 27–41. http://dx.doi.org/10.1037/1076-8998.1.1.27
69. Tsutsumi, A., & Kawakami, N. (2004). A review of empirical studies on the model of effort–reward imbalance at work: Reducing occupational stress by implementing a new theory. *Social Science & Medicine, 59,* 2335–2359. http://dx.doi.org/10.1016/j.socscimed.2004.03.030
70. Sinclair, U. (1906). *The jungle.* New York, NY: Doubleday.

Finding

Time to retire: Why Americans claim benefits early & how to encourage delay

Melissa A. Z. Knoll, Kirstin C. Appelt, Eric J. Johnson, & Jonathan E. Westfall

Summary. Because they are retiring earlier, living longer, and not saving enough for retirement, many Americans would benefit financially if they delayed claiming Social Security retirement benefits. However, almost half of Americans claim benefits as soon as possible. Responding to the Simpson–Bowles Commission's 2010 recommendation that behavioral economics approaches be used to encourage delayed claiming, we analyzed this decision using query theory, which describes how the order in which people consider their options influences their choices. After confirming that people consider early claiming before and more often than they consider later claiming, we designed interventions intended to encourage later claiming. Changing how information was presented did not produce significant shifts, but asking people to focus on the future first significantly delayed preferred claiming ages. Policymakers can apply these insights.

Tom has worked hard since his teen years and has contributed to the Social Security program for more than 40 years. A week before he turns 62 years old, friends at work point out that he will finally be able to start collecting Social Security retirement benefits. This seems tempting to Tom—after all, he thinks he deserves to start his retirement after so many years in the workforce. He would love to take the trips he has always dreamed about. But claiming now might be a mistake for Tom. If he's like many Baby Boomers in America, he has about $150,000 saved,[1] which will only give him about $500 a month in retirement income (using the standard rates provided in reference 2).

Tom logs on to the Social Security website and sees that if he claims his benefits now, he will get $1,098 each month (this is the average monthly Social Security retirement benefit for 62-year-old claimants in 2014).[3] He learns that if he waits until he is 66 years old to claim his benefits, he will get $1,464 a month, and if he waits until he is 70, he will get even more: $1,932 a month.[3] Like the majority of Americans,[4,5] Tom will have to rely on his Social Security benefits for most of his expenses, such as housing, food, transportation, and maybe even a vacation or two. Suddenly, Tom realizes he may have a lot to think about: Should he take the smaller benefit now or the significantly larger benefit later?

Thirty-one million Americans are projected to retire within the next decade.[6] Many, if not all, will face decisions like Tom's—whether about Social Security retirement benefits specifically or about other similarly structured public benefits or employer program benefits.[7–9] Because people are living longer and retiring earlier,[10,11] the average American now spends about 19 years in retirement—about 60% longer than in the 1950s.[12] The decision of when to claim benefits significantly affects retirees' financial well-being during this time of life. This is especially true for the many Americans who have little or no money saved by the time they retire.[4,13,14]

Additionally, recent changes in the retirement savings landscape have put the responsibility of savings and decisionmaking on the shoulders of employees rather than employers.[15] For example, the majority of employees with employer-sponsored retirement plans used to be covered by defined benefit plans, in which the employer provided a retirement benefit guaranteeing monthly payments for life. Now, most are covered by defined contributions plans, in which workers receive a lump sum at retirement and then must make their own decisions about how to manage that money. This means that getting the Social Security benefit claiming decision right is more important than ever. However, many Americans could be making a suboptimal choice: Claiming benefits early significantly and permanently decreases the size of the monthly benefit, yet almost half of all Social Security recipients claim their benefits as early as possible.[16,17] Why are people claiming their benefits early? How can they be encouraged to delay claiming?

The Claiming Decision

Like Tom, people thinking of claiming benefits have many factors to consider when making this important decision. On the one hand, as people get closer to Social Security's early eligibility age of 62 years, the notion of leaving the workforce and/or tapping into the Social Security funds they have contributed to for years is tempting. Tom could be like the large proportion of Americans who claim benefits as early as possible.[16,17] On the other hand, waiting to claim benefits provides retirees with more monthly income for the rest of their lives—the longer someone waits to claim benefits (up to age 70 years), the larger the monthly benefit. This extra money could mean the difference between enjoying retirement and struggling to make ends meet, especially in later years when health care costs may rise and retirement savings may have dried up. Indeed, research suggests that delaying claiming is the wiser economic decision for many.[10,18,19]

Prospective retirees must weigh the pros and cons of the claiming decision. Given the importance of the retirement decision to their future financial well-being, one might expect that prospective retirees put a lot of thought into this decision well in advance of actually retiring. Unfortunately, surveys show that 22% of people first think about when to start claiming Social Security benefits only a year before they retire. Another 22% first think about it only six months before retirement.[20] Research also shows that the retirement decision is malleable and affected by the way the decision is presented.[21,22]

Knoll, M. A. Z., Appelt, K. C., Johnson, E. J., & Westfall, J. E. (2015). Time to retire: Why Americans claim benefits early & how to encourage delay. *Behavioral Science & Policy, 1*(1), pp. 53–62.

Not all early claiming is caused by poor health or health-related work limitations.[4,23,24] Instead, there may be behavioral or psychological reasons why many individuals claim their benefits early (for a discussion, see reference 25). The National Commission on Fiscal Responsibility and Reform, also known as the Simpson–Bowles Commission, advocated in 2010 that the Social Security Administration (SSA) consider behavioral economics approaches "with an eye toward encouraging delayed retirement" (p. 52).[26] The commission did this with good reason: Insights from behavioral economics and psychology can help explain why people claim when they do and what can be done to help them make better decisions.

Why Do People Claim Early?

Tom's choice about when to claim benefits is what behavioral economists and psychologists call a classic *intertemporal choice* problem—a choice between getting something smaller now and getting something larger later. In the case of the Social Security benefit claiming decision, choosing to claim sooner means that Tom will have a smaller monthly benefit for the rest of his life, but he gets the benefit starting now. Choosing to claim later means Tom will have a larger monthly benefit for the rest of his life, but he must wait to get it (for an analysis of Social Security retirement benefits, see reference 25; for more general reviews of intertemporal choice, see references 27 and 28).

It is important to note that people faced with intertemporal choices often emphasize receiving the reward right away.[29] For Social Security benefits, this may explain why so many people want to claim benefits as soon as possible, a pattern observed in surveys and in administrative data.[16–18,29,30] We suspect that many people claim their benefits early because, like Tom, they become impatient as the opportunity to claim benefits finally approaches. If this is the case, then interventions that have helped people make more patient decisions in other financial contexts, such as saving for retirement,[14,31–33] may also affect Social Security benefit claiming.

To explore how people make this intertemporal choice, we applied a psychological theory of decision making called *query theory*, which offers insight into how people make decisions in many contexts.[34–38] Query theory suggests that many people are just like Tom: When they think about the claiming decision, the first thoughts that come to mind have to do with claiming right away. Thoughts about reasons to wait to claim often only come after thoughts in favor of claiming early. This sequence of thoughts generally leads people to have more thoughts supporting early claiming and to choose to claim benefits early. According to query theory, if people reverse the order in which they consider the choice options, they will change their choice:[37,39,40] What would happen, we asked, if we altered the order in which people considered the consequences of claiming at different ages?

Can Later Claiming Be Encouraged?

To answer this question, we used query theory to develop and test interventions that encourage people to wait to claim Social Security benefits. First, we tested what we called a *representation intervention,* which passively alters how the options within a choice are presented but does not explicitly encourage people to change how they think about the decision (for examples of representation interventions, see references 41–43). A representation intervention can be as simple as reframing a choice, such as asking employees to contribute to their savings account from a future raise rather than from a current paycheck.[14] In the case of Social Security benefits, later claiming is often framed as a gain (a larger monthly benefit compared with what is received if one claims early). Here, early claiming acts as a reference point or status quo option. One representation intervention that has had mild success in influencing claiming age reframes the choice options so that early claiming is framed as a loss (a smaller monthly benefit compared with what is received if one claims later).[21] We developed a representation intervention that communicated this reframing graphically, but it did not encourage participants to change the order in which they considered their options.

We next tested a *process intervention,* an active intervention that changes how people approach a decision. A process intervention for an intertemporal choice problem may simply ask people to focus on the future first (rather than following the common inclination to focus on the present first).[37,39] We applied this to the Social Security benefit claiming decision by asking people to list their thoughts in favor of later claiming before listing their thoughts in favor of early claiming.

This process intervention successfully reversed the order in which participants considered their options and led them to prefer later claiming.

Studying the Claiming Decision

Interventions to change people's behavior must be tested before they are implemented, especially when the stakes are high, which is certainly the case with Social Security claiming decisions. We used a series of three framed field studies[44] to explore why people claim benefits early and to test how to encourage them to delay claiming. Framed field studies sample from the population that makes the real-world decision and use forms and materials similar to those used in the actual setting. Unlike a randomized control trial, framed field studies do not involve the actual decision and are usually less expensive and time-consuming to conduct. In our case, although participants made hypothetical, nonbinding decisions about their Social Security benefits, the participants were drawn from the relevant target population: older Americans who are eligible or soon to be eligible for benefits. Further, they were presented with realistic decision materials modeled after actual SSA materials. This combination of features offers insight into the decisionmaking process that would otherwise be unavailable and also increases the chances that our results will generalize to the target population. In each study, we asked participants a series of questions through an online survey. (Detailed methods and results for each of our three studies are available in the Supplemental Material posted online.) Participants ranged in age from 45 to 70 years and were either eligible for Social Security retirement benefits or approaching eligibility.

Study 1: Exploring Impatience

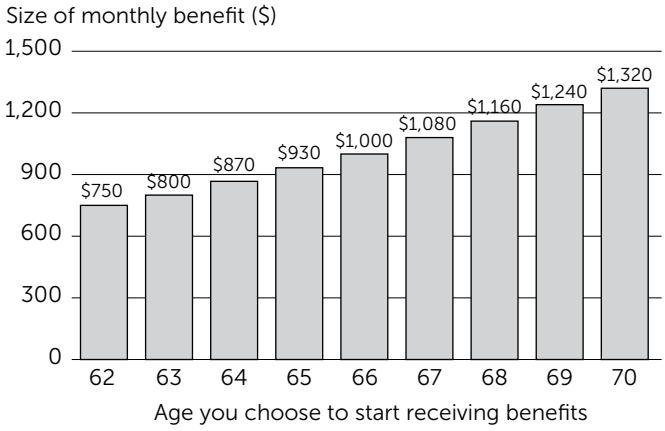
In Study 1, with 1,292 participants, we tested the assumption that prospective retirees tend to be impatient and prefer to claim their benefits as early as possible. We used information modeled after SSA's own materials to explain to participants how benefit claiming works (that is, how the size of the monthly benefit varies as a function of the age at which an individual claims benefits; see Figure 1A). We then asked participants to indicate at what age they would prefer to claim benefits. We found that nearly half of participants preferred to claim before their full retirement age (the age at which people become eligible for their full monthly benefit) and a third preferred the earliest possible benefit claiming age of 62 years (see Figure 2). This mirrors previous survey results as well as observed choices in the real world.[16–18,29,30]

We found it interesting that participants' decisions depended upon whether they were already eligible for benefits. Those who were eligible to collect benefits were much more likely to prefer claiming early compared with those who were not yet eligible (see Table S2 in the Supplemental Material). This suggests that people may have good intentions to delay claiming, but when the opportunity to claim finally presents itself, the temptation to claim right away can become too strong to resist. This strong preference for immediate rewards is what behavioral economists and psychologists call *present bias*, and it can explain why people make decisions that seem shortsighted.[45–47] Because present bias applies to immediate rewards and not future rewards, we expected it to contribute to early claiming when individuals were eligible to claim, not beforehand. Indeed, we found that before people become eligible for benefits, factors that are traditionally used in rational economic models of claiming, such as perceived health, predict claiming preferences. (Healthier individuals expect to live longer and spend more time in retirement and thus benefit more from claiming larger benefits later.) In contrast, present bias predicts claiming for already-eligible participants (see Table S3 in the Supplemental Material). These results are particularly striking given the hypothetical nature of the task: Even though participants were asked to imagine that they were approaching retirement and eligible for benefits, their actual eligibility status influenced their claiming preferences.

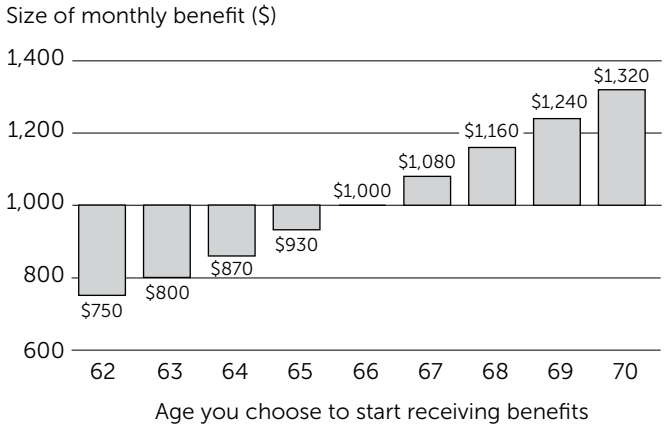
Because we successfully replicated real-world trends in claiming behavior, such as a preference for early claiming, we explored the claiming decision further to understand how people make their choice. We predicted that, like Tom, many participants would consider more reasons to claim their benefits early than reasons to claim later. We tested this hypothesis using a previously developed *type-aloud* protocol, often used in query theory studies, which asks participants to type every thought they have as they make a decision.[36,37] An analysis of these typed-aloud thoughts confirmed

Figure 1. Monthly benefit amount as a function of claiming age, assuming full benefit of $1,000 at full retirement age of 66 years

A: Standard graph used in Studies 2A, 2B, and 3*



B: Shifted x-axis graph used in Study 2A



C: Redesigned graph used in Study 2B

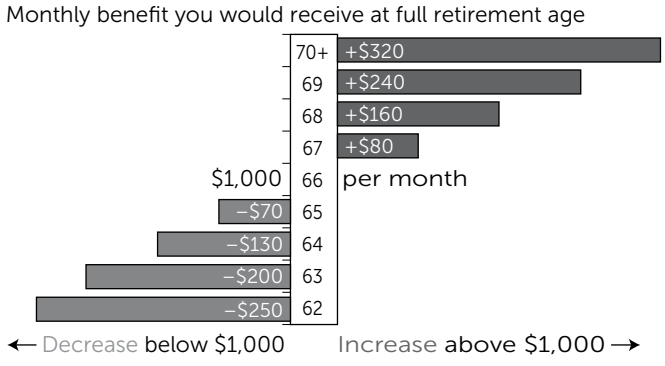
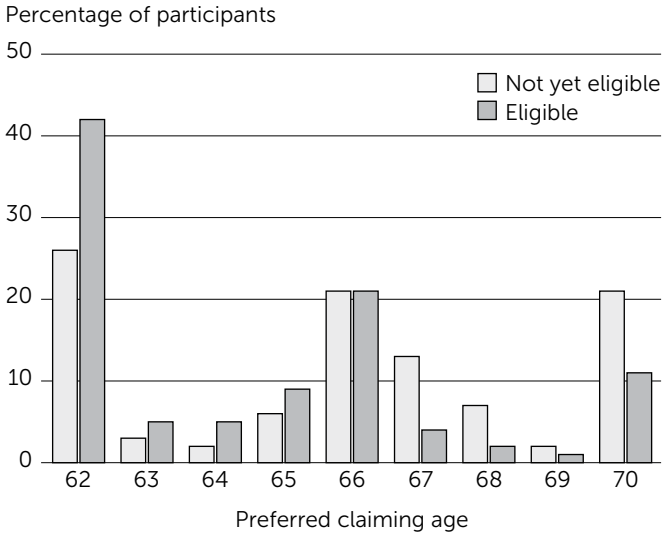


Figure adapted from *When to Start Receiving Retirement Benefits* (SSA Publication No. 05-10147, p. 1), Social Security Administration, 2014. Retrieved from http://www.socialsecurity.gov/pubs/EN-05-10147.pdf. (See the Supplemental Material for color versions of figures and detailed methods and results.) *In Study 1, the graph showed the monthly benefit as a percentage of full benefits.

Figure 2. Percentage of participants preferring to claim retirement benefits at each age from 62 to 70 years, by eligibility status, Study 1



that more participants thought predominately about early claiming (42%) than full claiming (18%) or delayed claiming (24%; see Table S4 in the Supplemental Material).

Next, we tested whether query theory—which highlights how the content and the order of thoughts predict preferences—can explain claiming preferences. We predicted that, like Tom, many participants would not only think more about claiming early than claiming later but would also think about claiming early before they thought about claiming later; this greater prominence (that is, greater number and earlier occurrence) of early-claiming thoughts would then lead participants to prefer to claim early. Using participants' typed-aloud thoughts, we found that the earlier and more participants thought about the benefits of claiming at early ages, the earlier they preferred to claim benefits. The participants with the most prominent early-claiming thoughts (that is, participants scoring in the top 25% on prominence of early-claiming thoughts) preferred to claim benefits over 4.5 years earlier than did the participants with the least prominent early-claiming thoughts (that is, participants scoring in the bottom 25%). Indeed, the content and order of participants' claiming-related thoughts are strong predictors of preferred claiming age even when controlling for benefit eligibility and traditional rational economic factors, such as education, wealth, and perceived health (see Table S2

in the Supplemental Material).

Study 1 showed that when people are shown typical information about benefit claiming, many of them think sooner and more often about reasons to claim their benefits early than about waiting to claim their benefits. This is associated with a preference for early claiming in a hypothetical claiming decision.

Study 2: Shifting the Focus

Using insights from Study 1 as guidance, in Studies 2A and 2B, we tested a representation intervention intended to encourage later claiming. Specifically, we made a number of changes to the standard graph to highlight the economic benefits of claiming later. We expected that these new graphs would make participants think more and earlier about reasons to delay claiming and this, in turn, would lead people to prefer later claiming ages.

We showed 785 participants one of three graphs depicting how the monthly benefit size varies as a function of the age at which one claims benefits: the standard graph depicting benefits as a series of increasing gains relative to $0 (see Figure 1A), a graph in which we shifted the *x*-axis from $0 to the full benefit amount (see Figure 1B), or a graph with an even stronger manipulation that highlighted losses in red and gains in green and rotated the figure to put later claiming at the top of the display (see Figure 1C; a color version of this figure is available in the Supplemental Material). We expected that making later claiming a visually prominent reference point would emphasize the later claiming option and reframe early claiming as a loss relative to full benefit claiming. This should increase the prominence of later claiming in participants' thoughts and shift participants' preferences to later claiming.

Our results, however, showed that neither representation intervention significantly influenced how participants thought about the claiming decision: Neither modified graph caused participants to think more or earlier about later claiming, and neither graph encouraged participants to prefer later claiming ages. Even though we believe that the graphs clearly make later claiming a visually prominent reference point, it is possible that the specific changes we made to the graphs were not strong or obvious enough to influence participants' thoughts. It is also possible, however, that graphical representations in general may not be an effective way to communicate retirement benefits information. This may be a particularly valuable finding because the SSA currently uses a graph to show how claiming age affects monthly benefits.

Study 3: Active Guidance

Query theory suggests that a process intervention that actively encourages people to change the order in which they think about the choice options can change the choice they make.[36] Previous research has shown that asking people faced with an intertemporal choice to focus on the future first encourages them to be more patient and choose a larger, later option over a smaller, sooner option.[38–40] In Study 3, we applied this query theory–based process intervention to the claiming decision. We expected that asking participants to reverse the order in which they considered early and later claiming (that is, to think about later claiming first) would increase the prominence of later claiming thoughts and this, in turn, would get people to prefer later claiming ages.

We asked 418 participants either to consider reasons favoring early claiming first and reasons favoring later claiming second (that is, the order in which participants consider the options given the standard presentation of benefits information in Study 1) or to consider reasons favoring later claiming first and reasons favoring early claiming second (that is, the reverse order). We found, as predicted by query theory, that participants who were prompted to consider claiming later before they considered claiming early thought more about claiming later and actually preferred later claiming ages, compared with participants who were prompted to think about claiming in the typical order of early claiming first and later claiming second. In other words, participants focusing on the future first have more prominent thoughts about later claiming, and this leads to a preference for claiming benefits later.

The different types of interventions we tested did not influence choices equally. Our process intervention was more successful than either of our representation interventions. The process intervention led to an average delay in preferred claiming age of 9.4 months, which is substantial when compared with the effects of various demographic and economic variables (for a discussion, see reference 21). Study 3 suggests that process interventions directing people to focus on the future first

are a promising approach for nudging older Americans toward later claiming.

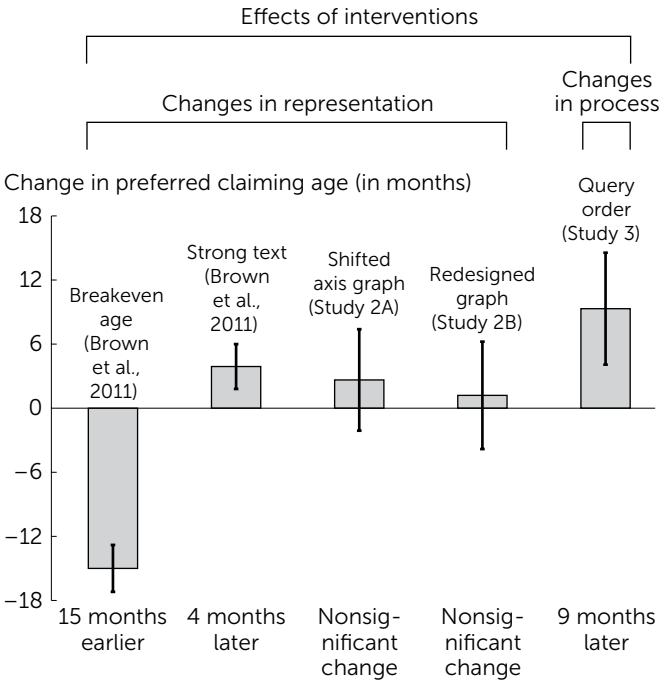
Policy Implications

As we described above, our research into consumers' decisions about when to claim Social Security benefits led us to test two types of interventions. In Study 2, representation interventions that changed the graphical depiction of monthly benefits produced nonsignificant delays in preferred claiming age of, at best, 2.6 months. In Study 3, however, a process intervention that encouraged people to focus on the future first resulted in a significant delay in preferred claiming age of, on average, 9.4 months.

Although this may seem like a modest change, it is sizeable when compared with the results of other interventions (see Figure 3). The accompanying permanent increase in monthly retirement benefits translates to substantially more money in the pockets of older Americans. For example, if Tom waited just nine months beyond his 62nd birthday to claim benefits, he would receive an extra $55 per month (a 5% increase) for life (these calculations are based on models provided by the SSA at http://www.socialsecurity.gov/OACT/quickcalc/early_late.html). If Tom lived to 85 years of age, about the average for his cohort (average life expectancy is averaged across genders and based on results from SSA's Life Expectancy Calculator, found at http://www.socialsecurity.gov/oact/population/longevity.html), this would add up to $4,776 in additional benefits. If Tom lived to 100 years of age, this would grow to $14,658 in additional benefits. The impact of seemingly modest delays is further magnified in aggregate, because more than 38 million Americans receive Social Security retirement benefits each month.[48]

Figure 3 makes another point as well. Choice architecture (that is, the way decision information is presented) is never neutral. Until a few years ago, SSA personnel computed prospective beneficiaries' *breakeven ages,* the age when the sum of the increase in monthly benefits from delaying claiming offsets the total benefits forgone during the delay period. This computation was intended to help potential retirees with their claiming decisions. However, as shown in Figure 3, this information accelerates preferred claiming age by 15 months,[21] which was not SSA's intention. SSA revised its description of benefits (see Figure 1A for a similar description), but Study 1 suggests the new description still leads many people to focus on early claiming.

Figure 3. Change in preferred claiming age relative to control (in months), by intervention



Error bars are included to indicate 95% confidence intervals. These bars indicate how much variation exists among data from each group. If two error bars overlap by up to a quarter of their total length, this indicates less than a 5% probability that the difference was observed by chance (that is, statistical significance at $p < .05$).

Given that all presentations of benefits information will influence choices in one direction or another, it is imperative that interventions be well informed by research. Framed field studies, such as those we have described here, can be extremely useful in designing and testing interventions for important real-world choices. Although this methodology has some constraints (for example, the dependence on hypothetical scenarios), it is a powerful complement to traditional lab and field studies because of its many strengths: sampling from relevant populations (that is, people for whom the retirement decision is real and, in many cases, imminent), presenting participants with realistic stimuli (that is, benefits information modeled on actual materials provided by SSA) to approximate how people normally encounter information, and discovering valuable process understanding insights

that lead directly to interventions that may be effective in changing behavior.

We recommend that full randomized control trials be pursued to further evaluate the interventions examined here and explore their effectiveness when the claiming decision is made with real consequences. Such research will likely require collaboration with SSA to expose retirees to interventions and provide access to data on retirees' actual claiming ages. With their new "*my* Social Security" website (http://www.ssa.gov/myaccount/), SSA may have a unique opportunity to prompt consumers to think about early or late claiming, gather consumers' thoughts about claiming, and see how their thoughts relate to their actual claiming behavior.

At the same time, it is important for researchers to continue exploring other process interventions, such as encouraging people to consider decisions in advance and precommit to a given option with the ability to choose differently later. Comparing different kinds of interventions and their effectiveness should be an active area of research both within the domain of retirement decisionmaking and beyond. For example, determining why changing the graphs in Study 2 did not shift participants' thoughts about the claiming decision could help clarify whether graphs are an effective way of communicating benefits information. Such comparisons will also help to determine how different interventions affect a heterogeneous population in which the ideal claiming age differs across individuals and many, but by no means all, people would benefit from delaying claiming.

More broadly, our studies underscore the point that different types of interventions have different strengths and weaknesses. On the one hand, representation interventions that change the display of choice information tend to require very little effort on the part of decisionmakers; in fact, these interventions are often most helpful for quick or automatic decisions.[49] For example, rearranging grocery store displays so that fruit is more accessible than candy helps people quickly reach for a healthy snack without thinking much about the decision. On the other hand, representation interventions tend to be very specific and need to be customized to fit each decision—rearranging grocery store displays to encourage healthier eating does not help people make sound retirement decisions.

In contrast, process interventions that change the way people approach decisions may teach a skill that, once learned, can be generalized. Training people to consider an alternative option first is a general skill that can apply to many situations, whether it is considering healthy food before considering junk food or considering saving for tomorrow before considering spending today. Process interventions often ask more from decisionmakers because they must change their decisionmaking process to some degree. But there may be ways to reduce the amount of effort needed. For example, we are currently researching whether preference checklists can function as a low-effort substitute for type-aloud protocols; initial results suggest that asking participants to simply read and respond to lists of claiming-related thoughts has an effect similar to that of asking participants to type aloud their own thoughts. With their different strengths, representation interventions and process interventions can be used to complement and reinforce each other, helping policymakers design useful interventions. These interventions, in turn, will help individuals make choices to improve their welfare in many different arenas, including retirement benefit claiming.

author affiliation

Knoll, Office of Retirement Policy, Social Security Administration; Appelt, Johnson, and Westfall, Center for Decision Sciences, Columbia Business School. Corresponding author's e-mail: eric.johnson@columbia.edu

author note

Melissa A. Z. Knoll is now at the Consumer Financial Protection Bureau Office of Research. Jonathan E. Westfall is now at the Division of Counselor Education & Psychology at Delta State University. Support for this research was provided by a grant from the Social Security Administration as a supplement to National Institute on Aging Grant 3R01AG027934-04S1 and a grant from the Russell Sage/Alfred P. Sloan Foundation Working Group on Consumer Finance. The views expressed in this article are those of the authors and do not represent the views of the Social Security Administration. This article is the result of the authors' independent research and does not necessarily represent the views of the Consumer Financial Protection Bureau or the United States. The authors thank participants at the Second Boulder Conference on Consumer Financial Decision Making for comments.

supplemental material

- http://behavioralpolicy.org/supplemental-material
- Methods & Analysis
- Additional Figures & Tables
- Additional References

References

1. Topoleski, J. J. (2013, July). *U.S. household savings for retirement in 2010* (Congressional Research Service Report for Congress No. R43057). Washington, DC: Congressional Research Service.
2. Bengen, W. P. (1994). Determining withdrawal rates using historical data. *Journal of Financial Planning, 7,* 171–180.
3. Social Security Administration. (2014). Modeling income in the near term, Version 6 (MINT6). Retrieved July 29, 2014, from http://www.ssa.gov/retirementpolicy/projection-methodology.html
4. U.S. Department of Health and Human Services, National Institutes of Health, National Institute on Aging. (2007). *Growing older in America: The health & retirement study* (NIH Publication No. 07-5757). Retrieved from http://www.nia.nih.gov/sites/default/files/health_and_retirement_study_0.pdf
5. Social Security Administration. (2010, April). *Income of the aged chartbook, 2008* (SSA Publication No. 13-11727). Retrieved from http://www.socialsecurity.gov/policy/docs/chartbooks/income_aged/2008/iac08.pdf
6. Reno, V. P., & Lavery, J. (2009). *Economic crisis fuels support for Social Security: Americans' views on Social Security*. Retrieved from National Academy of Social Insurance website: http://www.nasi.org/research/2009/economic-crisis-fuels-support-social-security
7. Burman, L. E., Coe, N. B., & Gale, W. G. (1999). *What happens when you show them the money? Lump sum distributions, retirement income security, and public policy* (Final Report 06750-003). Retrieved from Urban Institute website: http://www.urban.org/url.cfm?ID=409259
8. Bütler, M., & Teppa, F. (2005). *Should you take a lump-sum or annuitize? Results from Swiss pension funds* (CESifo Working Paper Series No. 1610). Retrieved from Social Science Research Network website: http://ssrn.com/abstract=834465
9. Warner, J. T., & Pleeter, S. (2001). The personal discount rate: Evidence from military downsizing programs. *American Economic Review, 91*(1), 33–53. http://dx.doi.org/10.1257/aer.91.1.33
10. Burtless, G., & Quinn, J. F. (2002). *Is working longer the answer for an aging workforce?* (Issue in Brief No. 2). Retrieved from Center for Retirement Research at Boston College website: http://crr.bc.edu/briefs/is_working_longer_the_answer_for_an_aging_workforce.html
11. Wise, D. A. (1997). Retirement against the demographic trend: More older people living longer, working less, and saving less? *Demography, 34,* 83–95. http://dx.doi.org/10.2307/2061661
12. Favreault, M. M., & Johnson, R. W. (2010, July). *Raising Social Security's retirement age* (Urban Institute Fact Sheet on Retirement Policy). Retrieved from Urban Institute website: http://www.urban.org/uploadedpdf/412167-Raising-Social-Security.pdf
13. Helman, R., Copeland, C., & VanDerhei, J. (2010, March). *The 2010 Retirement Confidence Survey: Confidence stabilizing, but preparations continue to erode* (Issue Brief No. 340). Retrieved from Employment Benefit Research Institute website: http://www.ebri.org/pdf/briefspdf/EBRI_IB_03-2010_No340_RCS.pdf
14. Thaler, R. H., & Benartzi, S. (2004). Save More Tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy, 112,* S164–S187.
15. Dushi, I., & Iams, H. M. (2008). Cohort differences in wealth and pension participation of near-retirees. *Social Security Bulletin, 68,* 45–66.
16. Muldoon, D., & Kopcke, R. W. (2008). *Are people claiming Social Security benefits later?* (Issue in Brief No. 8-7). Retrieved from Center for Retirement Research at Boston College website: http://crr.bc.edu/briefs/are_people_claiming_social_security_benefits_later.html
17. Song, J., & Manchester, J. (2007). Have people delayed claiming retirement benefits? Responses to changes in Social Security rules. *Social Security Bulletin, 67,* 1–23.
18. Coile, C., Diamond, P., Gruber, J., & Jousten, A. (2002). Delays in claiming social security benefits. *Journal of Public Economics, 84,* 357–385. http://dx.doi.org/10.1016/S0047-2727(01)00129-3
19. Munnell, A., Buessing, M., Soto, M., & Sass, S. A. (2006). *Will we have to work forever?* (Issue in Brief No. 4). Retrieved from Center for Retirement Research at Boston College website: http://crr.bc.edu/briefs/will-we-have-to-work-forever/
20. Employee Benefit Research Institute. (2008, July). *How long do workers consider retirement decision?* (FFE No. 91). Retrieved from http://www.ebri.org/pdf/fastfact07162008.pdf
21. Brown, J. R., Kapteyn, A., & Mitchell, O. S. (2011). *Framing effects and expected Social Security claiming behavior* (NBER Working Paper No. 17018). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w17018.pdf
22. Liebman, J. B., & Luttmer, E. F. P. (2009). *The perception of Social Security incentives for labor supply and retirement: The median voter knows more than you'd think* (Working Paper No. 08-01). Retrieved from National Bureau of Economic Research website: http://www.nber.org/~luttmer/ssperceptions.pdf
23. Gustman, A. L., & Steinmeier, T. L. (2002). Retirement and wealth. *Social Security Bulletin, 64,* 66–91.
24. Knoll, M. A. Z., & Olsen, A. (2014). Incentivizing delayed claiming of Social Security retirement benefits before reaching the full retirement age. *Social Security Bulletin, 74,* 21–43.
25. Knoll, M. A. Z. (2011). Behavioral and psychological aspects of the retirement decision. *Social Security Bulletin, 71,* 15–32.
26. National Commission on Fiscal Responsibility and Reform. (2010). *The moment of truth: Report of the National Commission on Fiscal Responsibility and Reform*. Retrieved from http://www.fiscalcommission.gov/news/moment-truth-report-national-commission-fiscal-responsibility-and-reform
27. Lynch, J. G., Jr., & Zauberman, G. (2007). Construing consumer decision making. *Journal of Consumer Psychology, 17,* 107–112. http://dx.doi.org/10.1016/S1057-7408(07)70016-5
28. Frederick, S., Loewenstein, G., & O'Donoghue, T. (2002). Time discounting and time preference: A critical review. *Journal of Economic Literature, 40,* 351–401. http://dx.doi.org/10.1257/002205102320161311
29. Behaghel, L., & Blau, D. M. (2010). *Framing Social Security reform: Behavioral responses to changes in the full retirement age* (IZA Discussion Paper No. 5310). Retrieved from Social Science Research Network website: http://ssrn.com/abstract=1708756
30. Social Security Administration. (2014, February). *Annual statistical supplement to the Social Security Bulletin, 2013* (SSA Publication No. 13-11700). Retrieved from http://www.ssa.gov/policy/docs/statcomps/supplement/2013/6b.html#table6.b5
31. Choi, J. J., Laibson, D., & Madrian, B. C. (2004). Plan design and 401(k) savings outcomes. *National Tax Journal, 57,* 275–298.
32. Hershfield, H. E., Goldstein, D. G., Sharpe, W. F., Fox, J., Yeykelis, L., Carstensen, L. L., & Bailenson, J. N. (2011). Increasing saving behavior through age-progressed renderings of the future self. *Journal of Marketing Research, 48,* 23–37.
33. Knoll, M. A. Z. (2010). The role of behavioral economics and behavioral decision making in Americans' retirement savings decisions. *Social Security Bulletin, 70,* 1–23.
34. Dinner, I., Johnson E. J., Goldstein, D., & Liu, K. (2011). Partitioning default effects: Why people choose not to choose. *Journal of Experimental Psychology: Applied, 17,* 332–341. http://dx.doi.org/10.1037/a0024354
35. Hardisty, D. J., Johnson, E. J., & Weber, E. U. (2010). A dirty

word or a dirty world? Attribute framing, political affiliation, and query theory. *Psychological Science, 21,* 86–92. http://dx.doi.org/10.1177/0956797609355572
36. Johnson, E. J., Häubl, G., & Keinan, A. (2007). Aspects of endowment: A query theory of value construction. *Journal of Experimental Social Psychology: Learning, Memory, and Cognition, 33,* 461–474. http://dx.doi.org/10.1037/0278-7393.33.3.461
37. Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice: A query-theory account. *Psychological Science, 18,* 516–523. http://dx.doi.org/10.1111/j.1467-9280.2007.01932.x
38. Weber, E. U., & Johnson, E. J. (2011). Query theory: Knowing what we want by arguing with ourselves. *Behavioral and Brain Sciences, 34,* 91–92. http://dx.doi.org/10.1017/S0140525X10002797
39. Appelt, K. C., Hardisty, D. J., & Weber, E. U. (2011). Asymmetric discounting of gains *and* losses: A query theory account. *Journal of Risk and Uncertainty, 43,* 107–126. http://dx.doi.org/10.1007/s11166-011-9125-1
40. Figner, B., Weber, E. U., Steffener, J., Krosch, A., Wager, T. D., & Johnson, E. J. (2015). *Framing the future first: Brain mechanisms of enhanced patience in intertemporal choice.* Manuscript in preparation.
41. Choi, J. J., Haisley, E., Kurkoski, J., & Massey, C. (2012). *Small cues change savings choices* (NBER Working Paper No. 17843). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w17843.pdf
42. Goda, G. S., Manchester, C. F., & Sojourner, A. (2012). *What will my account really be worth? An experiment on exponential growth bias and retirement saving* (NBER Working Paper No. 17927). Retrieved from National Bureau of Economic Research website: http://www.nber.org/papers/w17927
43. Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.
44. Harrison, G. W., & List, J. A. (2004). Field experiments. *Journal of Economic Literature, 42,* 1009–1055. http://dx.doi.org/10.1257/0022051043004577
45. Benhabib, J., Bisin, A., & Schotter, A. (2010). Present-bias, quasi-hyperbolic discounting, and fixed costs. *Games and Economic Behavior, 69,* 205–223. http://dx.doi.org/10.1016/j.geb.2009.11.003
46. Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics, 112,* 443–477. http://dx.doi.org/10.1162/003355397555253
47. Phelps, E. S., & Pollak, R. A. (1968). On second-best national savings and game-equilibrium growth. *Review of Economic Studies, 35,* 185–199. http://dx.doi.org/10.2307/2296547
48. Social Security Administration. (2014, June). *Monthly statistical snapshot, May 2014.* Retrieved from http://www.ssa.gov/policy/docs/quickfacts/stat_snapshot/2014-05.pdf
49. Johnson, E. J., & Goldstein, D. G. (2012). Decisions by default. In E. Shafir (Ed.), *Behavioral foundations of public policy* (pp. 417–427). Princeton, NJ: Princeton University Press.


ILLING
New charges
£8.69
out of plan
ET MORE

Review

Designing better energy metrics for consumers

Richard P. Larrick, Jack B. Soll, & Ralph L. Keeney

Summary. Consumers are often poorly informed about the energy consumed by different technologies and products. Traditionally, consumers have been provided with limited and flawed energy metrics, such as *miles per gallon,* to quantify energy use. We propose four principles for designing better energy metrics. Better measurements would describe the amount of energy *consumed* by a device or activity, not its energy efficiency; relate that information to important *objectives*, such as reducing costs or environmental impacts; use *relative* comparisons to put energy consumption in context; and provide information on *expanded* scales. We review insights from psychology underlying the recommendations and the empirical evidence supporting their effectiveness. These interventions should be attractive to a broad political spectrum because they are low cost and designed to improve consumer decisionmaking.

Consider a family that owns two vehicles. Both are driven the same distance over the course of a year. The family wants to trade in one vehicle for a more efficient one. Which option would save the most gas?

A. Trading in a very inefficient SUV that gets 10 miles per gallon (MPG) for a minivan that gets 20 MPG.
B. Trading in an inefficient sedan that gets 20 MPG for a hybrid that gets 50 MPG.

Most people assume option B is better because the difference in MPG is bigger (30 MPG vs. 10 MPG), as is the percentage of improvement (150% vs. 100%). But to decipher gas use and gas savings, one must convert MPG, a common efficiency metric, to actual consumption. Dividing 100 miles by the MPG values given above, our family can see that option A reduces gas consumption from 10 gallons to 5 every hundred miles, whereas option B reduces gas consumption from 5 gallons to 2 over that distance.

Making rates of energy consumption clear is more important than ever given the urgent need to reduce fossil fuel use globally. People around the world are dependent on fossil fuels, such as coal and oil. But emissions from burning fossil fuels are modifying Earth's climates in risky ways, from raising average temperatures to transforming habitats on land and in the oceans. Although individual consumer decisions have a large effect on emissions—passenger vehicles

Larrick, R. P., Soll, J. B., & Keeney, R. L. (2015). Designing better energy metrics for consumers. *Behavioral Science & Policy, 1*(1), pp. 63–75.

and residential electricity use account for nearly half of the greenhouse gas emissions in the United States—consumers remain poorly informed about how much energy they consume.[1–3] Behavioral research offers many insights on how to inform people about their energy consumption and how to motivate them to reduce it.[4] One arena in which this research could be immediately useful is on product labels, where energy requirements could be made clearer for consumers faced with an abundance of choices.

The current U.S. fuel economy label for automobiles (revised in 2013) includes a number of metrics associated with energy. The familiar MPG metric is most prominent, but one can also see gallons per 100 miles (GPHM), annual fuel cost, a rating of greenhouse gas emissions, and a five-year relative cost or savings figure compared with what one would spend with an average vehicle (see Figure 1). The original label introduced in the 1970s contained two MPG figures (see Figure 2). As the label was being redesigned for 2013, there was praise for including new information and criticism for providing too much information.[5–7] The new fuel economy label raises two general questions that apply to many settings in which consumers are informed about energy use, such as on appliance labels, smart meter feedback, and home energy ratings:

- What energy information should be given to consumers?
- How much is the right amount?

How information is presented always matters. More often than not, people pay attention to what they see and fail to think further about what they really want to know. In his best-selling book *Thinking, Fast and Slow,* Nobel prize–winning psychologist Daniel Kahneman reviewed decades of research on biases in decision-making and found a common underpinning: "What you see is all there is."[8] Too often, people lack the awareness, knowledge, and motivation to consider relevant information beyond what is presented to them. This can produce problems. In the case of judging energy use, incomplete or misleading metrics leave consumers trapped with a poor understanding of the true consequences of their decisions. But this important communication can be improved.

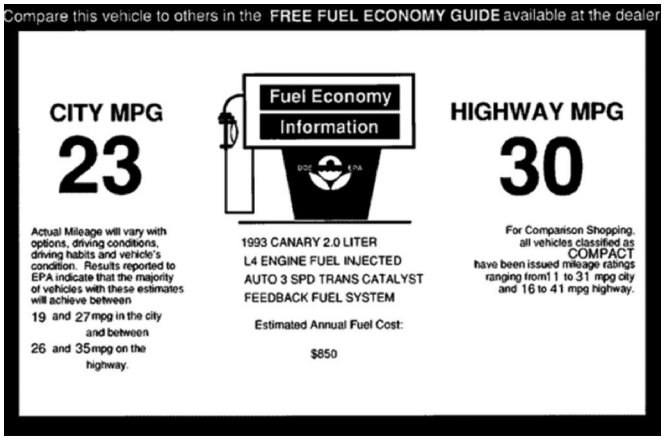
A CORE Approach to Better Decisionmaking

How people learn and how they make decisions is less of a mystery than ever before. Insights from psychology, specifically, are now used to help consumers make better decisions for themselves and for society.[9,10] In this context, we have created four research-based principles, which we abbreviate as CORE, that could be employed to better educate people about energy use and better prepare them to make informed decisions in that domain. They include:

- CONSUMPTION: Provide consumption rather than efficiency information.
- OBJECTIVES: Link energy-related information to objectives that people value.
- RELATIVE: Express information relative to meaningful comparisons.
- EXPAND: Provide information on expanded scales.

Figure 1. Revised fuel economy label (2013)

Figure 2. Original fuel economy label (from 1993)



Consumption: An Alternative to Efficiency Information

Our first principle is to express energy use in consumption terms, not efficiency terms. It is common practice in the United States to express the energy use of many products as an efficiency metric. For example, just as cars are rated on MPG, air conditioners are given a seasonal energy efficiency rating (SEER), which measures BTUs of cooling divided by watt-hours of electricity. Efficiency metrics put the energy unit, such as gallons or watts, in the denominator of a ratio. Unfortunately, efficiency metrics such as MPG and SEER produce false impressions because consumers use inappropriate math when reasoning about efficiency.

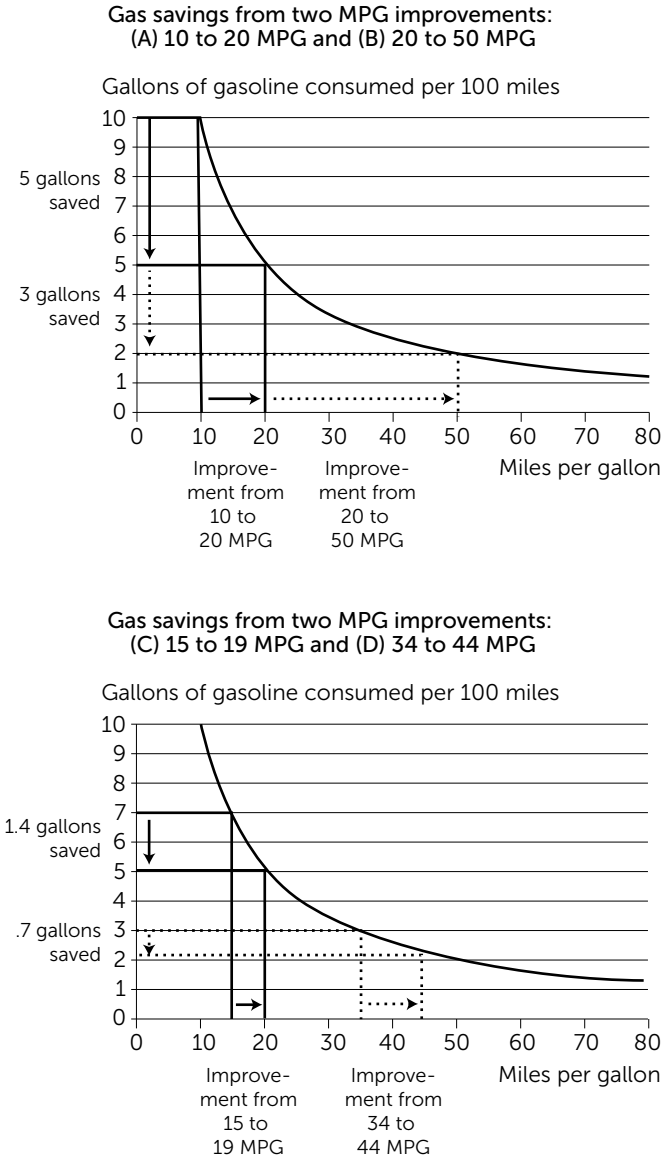
At the most basic level, efficiency metrics such as MPG do convey some crystal clear information: Higher is better. However, as our opening example showed, the metrics create a number of problems when people try to use them to make comparisons between energy-consuming devices. Consider a town that owns an equal number of two types of vehicles that differ in their fuel efficiency. All of the vehicles are driven the same distance each year. The town is deciding which set of vehicles to upgrade to a hybrid version:

C. Should it upgrade the fleet of 15-MPG vehicles to hybrids that get 19 MPG?
D. Or should it upgrade the fleet of 34-MPG vehicles to hybrids that get 44 MPG?

Larrick and Soll presented these options to an online sample of adults.[11] Seventy-five percent incorrectly picked option D over option C. In fact, option C saves nearly twice as much as gas as option D does. Figure 3 plots the highly curvilinear relationship between MPG and gas consumption. The top panel shows the gas savings from the upgrades described in the opening example. The bottom panel shows the gas savings from each of the upgrades described in C and D. Larrick and Soll called the tendency to underestimate the benefits of MPG improvements on inefficient vehicles (and to overestimate them on efficient vehicles) the "MPG illusion."[11]

The confusion caused by MPG is avoided, however, when the energy unit is put in the numerator of a ratio. When the same decision also included a GPHM number, people could see clearly that replacing the

Figure 3. Gas consumed per 100 miles of driving as a function of miles per gallon (MPG)



15-MPG (6.67-GPHM) vehicles with 19-MPG (5.26-GPHM) hybrids saved twice as much gas as replacing the 34-MPG (3.00-GPHM) vehicles with 44-MPG (2.27-GPHM) hybrids.[11]

Consumption metrics are more helpful than efficiency metrics because they not only convey what direction is better (lower) but also provide clear insights about the size of improvements. A consumption perspective (see Table 1) reveals that replacing a 10-MPG car with an 11-MPG car saves about as much gas as replacing a 34-MPG car with a 50-MPG car (1 gallon per 100 miles). A cash-for-clunkers program in

Table 1. Converting miles per gallon (MPG) to gas consumption metrics

MPG	Gallons per 100 miles	Gallons per 100,000 miles
10	10	10,000
11	9	9,000
12.5	8	8,000
14	7	7,000
16.5	6	6,000
20	5	5,000
25	4	4,000
33	3	3,000
50	2	2,000
100	1	1,000

the United States in 2009 was ridiculed for seeming to reward small changes[12]—such as trade-ins of 14-MPG vehicles that were replaced by 20-MPG vehicles—but a consumption perspective reveals that this is actually a substantial improvement of 2 gallons every 100 miles. Moving consumers from cars with MPGs in the teens into cars with MPGs in the high 20s is where most of society's energy savings will be achieved.

Although consumption measures may be unfamiliar in the consumer market, they are common in other settings. For example, U.S. government agencies transform MPG to gallons per mile to calculate fleet MPG ratings. Europe and Canada use a gas consumption measure (liters per 100 kilometers). Recently, the National Research Council argued that policymakers need to evaluate efficiency improvements in transportation using a consumption metric.[13,14] The MPG illusion motivated the addition of the GPHM metric to the revised fuel economy label (see Figure 1).

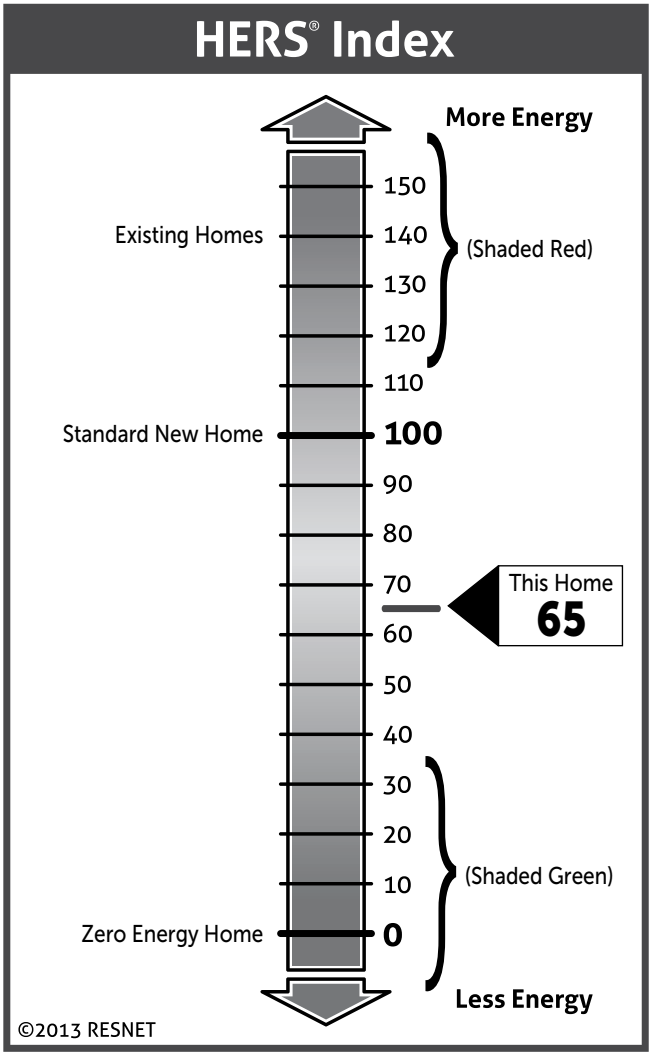
MPG is a well-known energy measure with the wrong number on top, but it is not the only metric that needs improvement. Several important energy ratings similarly place performance on top of energy use, including those for air-conditioning, home insulation, and IT server ratings.[15] These efficiency ratings also distort people's perceptions. Older homes may have air-conditioning units that are rated at 8 SEER (a measure of cooling per watt-hour of electricity) and the most efficient (and expensive) new units have SEER ratings above 20. For a given space and outdoor-temperature difference, energy consumption is once again an inverse: 1/SEER. Trading in an outdated 10-SEER air conditioner for a 13-SEER air conditioner yields large energy savings—more than the trade-in of a 14-SEER unit for a 20-SEER unit for the same space and conditions.

There is no name for the metric 1/SEER, and, unlike GPHM, the basic units in SEER (watts and BTUs) are unfamiliar to most people. Still, it is possible to be clearer. For air conditioners, the consumption metric might need to be an index, expressed as percentage of savings from an initial baseline measure (e.g., 8 SEER). As an example, consider the consumption index created by the Residential Energy Services Network called the *Home Energy Rating System* (HERS) index. A standard home is set at a unit of 100; homes that consume more energy have a higher score and are shaded in red in visual depictions of the index; homes that consume less energy have a lower score and are shaded in green (see Figure 4). A home rated at 80 uses 20% less energy than a home comparable in size and location. The HERS label, therefore, needs to be adapted to specific circumstances. Those circumstances can be explored at http://www.resnet.com. By comparison, a similar label for air conditioners actually could be more general.

Although a large home in Florida uses more air-conditioning than a small home in Minnesota does, the same consumption index can provide an accurate picture of relative energy savings possible from a more efficient air-conditioning unit. For example, Floridians know that their monthly electricity bill is high in the summer and roughly by what amount (perhaps $200 per month). A consumption index would allow them to quantify the savings available from greater efficiency (a 20% reduction in my $200 electricity bill is $40 per month). Minnesotans, on the other hand, have a smaller air-conditioning bill and would recognize that a 20% reduction yields smaller benefits. More precise cost savings could be provided at the point of purchase on the basis of additional information about effects from local electricity costs, home size, and climate, including the number of days when air-conditioning is likely needed in different regions.

In sum, the problem with MPG, SEER, and other efficiency metrics is that one cannot compare the energy savings between products without first inverting the numbers and then finding the difference. The main benefit of a consumption metric is that it does the math for people. There is no loss of information, and consumption measures help people get an accurate

Figure 4. Home Energy Rating System label



picture of the amount of energy use and savings.

Objectives: Make Cost and Environmental Impact Clear

Our second principle is to translate energy information into terms that show how energy use aligns with personal goals, such as minimizing cost or reducing the environmental impact of consumption. Theoretically, people would not require such a translation because both cost and environmental impact are often directly related to energy use. In the case of driving, for instance, as gas consumption goes up, gasoline costs and carbon dioxide (CO_2) emissions rise at exactly the same rate. Realistically, however, people may not know that these relationships are so closely aligned or stop to think about how energy usage affects the goals they care about. For example, burning 100 gallons of gas emits roughly one ton of CO_2. That outcome is invisible when people stop at "what you see is all there is."

Some consumers may care about MPG as an end in itself, but the measure is more often a proxy for other concerns, such as the cost of driving a car, its impact on the environment, or its impact on national security. Keeney argued that decisionmakers need to distinguish "means objectives" such as MPG from "fundamental objectives" such as environmental impact so that they can see how their choices match or do not match their values.[16] Providing consumers with cost and environmental translations directs their attention to these end objectives and helps them see how a means objective—energy use—affects those ends.

There is a tension, however, between offering translations and overwhelming people with information. In the redesign process for the fuel economy label, expert marketers counseled the Environmental Protection Agency (EPA) to "keep it simple."[5] However, the new EPA label for automobiles (see Figure 1) provides a number of highly related attributes, including MPG, GPHM, annual fuel costs, and a greenhouse gas rating. Is this too much information?

Ungemach and colleagues have argued that multiple translations are critical in helping consumers recognize and apply their end objectives when making choices among consumer products such as cars or air conditioners.[17] Translations have two effects. The first is what is called a *counting effect,* meaning that preferences grow stronger for choices that look favorable in more than one category.[18] For instance, multiple translations of fuel efficiency increase preference for more efficient vehicles because consumers see that the more efficient car seems to be better on three dimensions: It gets more MPG, has lower fuel costs, and is more helpful to the environment. But MPG is a not a distinct dimension from fuel costs and environmental impact, so the effect of translation is partly attributable to a double counting.

In addition, Ungemach and colleagues have found that translations have a *signpost* effect by reminding people of an objective they care about and directing them on how to reach it.[17] In one study, Ungemach and fellow researchers measured participants' attitudes toward the environment and willingness to engage in behaviors that protect the environment.[17] Participants had to choose between two cars: one that was a more efficient and more expensive car and one that was a

Table 2. Examples of choice options

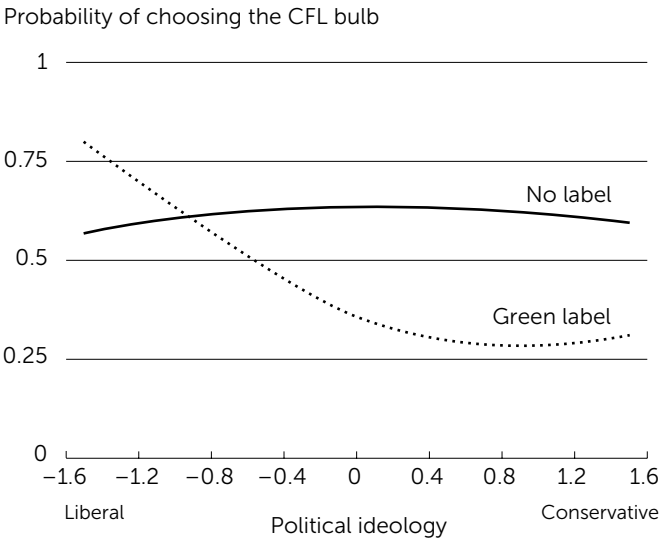
Options	Annual fuel cost	Gallons per 100 miles	Price of car
Car A	$3,964	7	$29,999
Car B	$2,775	5	$33,699

Options	Annual fuel cost	Greenhouse gas ratings (out of 10 = best)	Price of car
Car A′	$3,964	5	$29,999
Car B′	$2,775	7	$33,699

less efficient and less expensive car (see Table 2). When vehicles were described in terms of both annual fuel costs and greenhouse gas ratings, environmental attitudes strongly predicted preference for the more efficient option. However, when vehicles were described in terms of annual fuel costs and gas consumption, environmental attitudes were not correlated with preference for the more efficient option. Although both annual fuel cost and gas consumption are perfect proxies for greenhouse gas emissions, they were inadequate as signposts for environmental concerns. They neither reminded people of something they cared about nor helped them act on those concerns. The explicit translation to greenhouse gas ratings was necessary to enable people to act on their values. Additional studies demonstrated signpost effects for choices regarding air conditioners[17] by varying whether the energy metric was labeled *BTUs per watt, Seasonal Energy Efficiency Rating,* or *Environmental Rating.* Only *Environmental Rating* evoked choices in line with subjects' attitudes toward the environment.

One problem with translating energy measures into end objectives is that some consumers may be hostile to the promoted goals.[19] For example, in the United States, political conservatives and liberals alike believe that reducing personal costs and increasing national security are valid reasons to favor energy-efficient products. But conservatives find the goal of diminishing climate change to be less persuasive than do liberals.[20] As a result, emphasizing the environmental benefits of energy-efficient products may backfire with some people. Gromet and colleagues found a backlash effect in a laboratory experiment in which 200 participants were given $2 to spend on either a standard incandescent light bulb or a more efficient compact fluorescent

Figure 5. Probability of buying a more expensive compact fluorescent light (CFL) bulb when it has a green label ("protect the environment") or not as a function of political ideology



light (CFL) bulb.[20] All participants were informed about the cost savings of using a CFL. In one condition, the CFL came with a "protect the environment" label. Compared with participants in a control condition with no label, liberals showed a slightly higher rate of CFL purchase, but the purchase rate for independents and conservatives dropped significantly (see Figure 5). With no label, the economic case was equally persuasive to conservatives and liberals. The presence of the label forced conservatives to trade off a desired economic outcome with an undesired political expression.

Thus, there is a potential tension when using multiple translated attributes—they may align with a consumer's concerns but may also increase the chances of triggering a consumer's vexation. One option for navigating this tension is to target translations to specific market segments. Environmental information can be emphasized in more liberal communities and omitted in more conservative ones. Another option is to provide environmental information along a continuum rather than as an either–or choice. The environmental label described above backed consumers into a corner. People were forced to choose between a product that seemed to endorse environmentalism and one that did not. In contrast, the greenhouse gas rating on the new EPA label is continuous (for example, 6 vs. 8 on a 10-point scale) and is less likely to appear as an

endorsement of a political view.

Relative: Provide Information with Meaningful Comparisons

Our third principle is to express energy-related information in a way that allows consumers to compare their own energy use with meaningful benchmarks, such as other consumers or other products. This principle is illustrated nicely in a series of large-scale behavioral interventions conducted by the company OPower across many areas of the United States. The company applied social psychological research on descriptive norms to reduce energy consumption.[21] In field studies, OPower presented residential electricity consumers with feedback on how their energy use compared with the energy use of similar neighbors (thereby largely holding constant housing age, size, and local weather conditions). Consumers who see that they are using more energy than those in comparable homes are motivated to reduce their energy use. To offset complacency in homes performing better than average, OPower couples neighbor feedback with a positive message, such as a smiley face, to encourage sustained performance. Feedback about neighbors alone—in the absence of any changes in price or incentives—reduces energy consumption by about 2%, which is roughly the reduction one would expect if prices were increased through a 20% tax increase.[22] Other studies have shown that feedback about neighbors can produce small but enduring savings for natural gas[23]and water consumption.[24,25] Moreover, there is no evidence that consumers ignore or tire of feedback over time.[26] Although many OPower interventions combine neighbor feedback with helpful advice on how to reduce energy use, research suggests that norm information alone is effective in motivating change.[27]

The benefits of comparative information are often attributed to people's intrinsic competitiveness. Homeowners want to "keep up with the Joneses" in everything, including their energy conservation. Competition plays an important part, but we believe that the neighbor feedback effect demonstrates a more basic psychological point. Energy consumption (for example, kilowatts or ergs) and even energy costs (for example, $73.39) are difficult to evaluate on their own. Is $73.39 a lot of money or a little? Feedback about neighbors' energy consumption provides a reference point that helps people judge the magnitude of the outcomes of their actions, as when they learn that they spend $40 more per month on natural gas than their neighbors do. Providing information so that it can be seen as relatively better or worse than a salient comparison measure, such as neighborhood norms, the numbered scale for HERS (see Figure 4), or the greenhouse gas ratings on the EPA label (see Figure 1), helps consumers better understand an otherwise abstract energy measure.[28,29]

Reference points also have a second effect, which is to increase motivation. Decades of research have shown that people strongly dislike the feelings of loss, failure, and disappointment. Further, the motivation to eliminate negative outcomes is substantially stronger than the motivation to achieve similar positive outcomes.[30,31] Because reference points allow people to judge whether outcomes are good or bad, they strongly motivate those who are coming up short to close the gap: Being worse than the neighbors or ending up "in the red" (see Figure 4) leads people to work to avoid those outcomes.

Of course, about half of the people in an OPower study would be given the positive feedback that they are better than average, which can lead to complacency. An alternative is to have people focus on stretch goals instead of the average neighbor.[32] Carrico and Riemer studied the energy use in 24 buildings on a college campus.[33] The occupants of half of the buildings were randomly assigned to meet a goal of a 15% reduction in energy use and received monthly feedback in graphic form. Occupants of the remaining buildings received the same goal but no feedback on their performance. There were no financial incentives tied to meeting the goal, and none of the occupants personally bore any of the energy costs. Nevertheless, those who received feedback on whether they met the goal achieved a 7% reduction in energy use; those who received no feedback showed no reduction in energy use.

OPower uses a similar logic when it lists the energy consumption of the 10% most efficient homes in a neighborhood, in addition to the energy consumption of the average home. This challenging reference point introduces a goal and gives residents with better than average energy consumption habits a target that they currently fall short of and can aim for.

Research on self-set goals has also found beneficial effects. In a study of 2,500 Northern Illinois homes,

Harding and Hsiaw found that homeowners who set realistic goals for reducing their electricity use (goals up to 15%) reduced their consumption about 11% on average, which is substantially more than the reductions achieved by homeowners who set no goals or who set unrealistically ambitious goals and abandoned them.[34]

Of the many possible reference points that could be used, which ones best help reduce energy consumption? Focusing on typical numbers (such as neighbor averages) helps consumers know where they stand; deviating from the typical may motivate consumers to explore why they are inferior or superior to others. As we have noted, however, superiority can also lead to complacency. If continued energy reduction is desired, policymakers or business owners should identify a realistic reference point that casts current levels of consumption as falling short. Both realistic goals, say a 10% reduction, and social comparisons to the best performers, such as the 10% of neighbors who use the least energy, create motivation for those already performing better than average.

The most extreme form of relative comparison is when all energy information is converted to a few ranked categories, such as with a binary certification system (for example, Energy Star certified or not) or using a limited number of colors and letter grades (e.g., European Union energy efficiency labels).[5,29,35] If used alone, these simple rankings are likely to be effective at changing behavior,[29] but they may generate some undesirable consequences. For example, ranked categories exaggerate the perceived difference between two similar products that happen to fall on either side of a threshold (for example, B vs. C or green vs. yellow) and thereby distort consumer choice.[29,35] Other challenges arise when there are multiple product categories, such as SUVs and compact vehicles—should an efficient SUV be graded against all vehicles (and score poorly) or against other SUVs (and score highly)? We recommend that simple categories not be used alone but rather be combined with richer information on cost and energy consumption so that consumers can make a decision that best fits their personal goals and preferences.

Expand: Provide Information on Larger Scales

Our fourth principle is to express energy-related information on expanded scales, which allows the impact of a change to be seen over longer periods of time or over greater use. For example, the cost of using an appliance could be expressed as 30 cents per day, $109.50 per year, or $1,095 over 10 years. Fundamentally, these expressions are identical. However, a growing body of research shows that people pay more attention to otherwise identical information if it is expressed on expanded scales (such as cost over 10 years) rather than contracted scales (cost per day). As a result, they are more likely to choose options that look favorable on the expanded dimensions.[36–39] When people compare two window air-conditioning units that differ in their energy use, small scales such as cost per hour make the differences look trivial—savings are within pennies of each other (for example, 30 cents vs. 40 cents per hour). Large scales such as cost per year, however, reveal costs in the hundreds of dollars (e.g., $540 vs. $720 per year). The problem of trivial costs raises questions about the benefits of smart meters. If real-time energy and cost feedback are expressed in terms of hourly consumption, for example, all energy use can seem inconsequential.

A number of studies have shown that providing cost information over an extended period of time, such as the cost of energy over the expected lifetime operation of a product, increases preferences for more expensive but more efficient products.[37,38] Camilleri and Larrick tested the benefits of scale expansion directly by giving people (n = 424) hypothetical choices between six pairs of cars in which a more efficient car cost more than a less efficient car.[40] Participants saw vehicle gas consumption stated for one of three distances: 100 miles, which is the distance used to express consumption on the EPA car label; 15,000 miles, which is the distance used to express annual fuel costs on the EPA car label; or 100,000 miles, which is roughly equivalent to a car's lifetime driving distance (see Table 3).

The researchers presented some participants with a gas-consumption metric and others with a cost metric. Participants were most likely to choose the efficient car when they were given cost information (an end objective) and when it was scaled over 100,000 miles. In a second study, when the gas savings from the efficient car did not cover the difference in upfront price (over 100,000 miles of driving), interest in the efficient car naturally dropped, but it remained highest when cost was expressed on the 100,000 miles scale.

Table 3. Three examples from Camilleri and Larrick (2014) of expanding gas costs over different distances (100 miles, 15,000 miles, 100,000 miles)

Options	Cost of gas per 100 miles of driving	Price of car
Car A	$20	$18,000
Car B	$16	$21,000

Options	Cost of gas per 15,000 miles of driving	Price of car
Car A′	$3,000	$18,000
Car B′	$2,400	$21,000

Options	Cost of gas per 100,000 miles of driving	Price of car
Car A″	$20,000	$18,000
Car B″	$16,000	$21,000

Hardisty and colleagues presented people with varied cost information for three time scales—one year, five years, and 10 years—for light bulbs, TV sets, furnaces, and vacuum cleaners.[37] Control subjects received no cost information. Providing cost information increased people's choice of the more expensive, energy-efficient product. The tendency to choose the more efficient product increased as the time scale increased. However, results varied according to the product. This suggests the importance of testing design changes,[41] even in hypothetical studies, to uncover context-specific psychological effects.

A major benefit of expressing energy consumption and energy costs over larger time spans is that it counteracts people's tendency to be focused on the present in their decisionmaking. A large body of research in psychology finds that people heavily discount the future; for instance, they focus more on immediate out-of-pocket costs and do not consider delayed savings.[42] Expanded scales help people to consider the future more clearly by doing the math for them.[43] However, costs that are delayed long into the future may need to be expressed in terms of current dollars to take into account the time value of money.

What is the best time frame to use? Although the results suggest that larger numbers have more psychological impact, there are several reasons to strive for large but reasonable numbers. The magnitude of gas savings appears even larger if scaled to 300,000 miles of driving, but that is not a realistic number of miles that one vehicle will accumulate. Consumers might see it as manipulative. Also, at some point, numbers become so large that they become difficult to relate to (try considering thousands of pennies per year). All of these factors suggest a basic design principle, which is that scale expansion best informs choice if the expansion is set to a large but meaningful number, such as the expected lifetime of an appliance.

Combining CORE Principles

We have largely discussed the effectiveness of the four proposed CORE principles when applied separately. But how do they work in combination? Multiple principles often are being used at once in labeling. The revised EPA label (see Figure 1), for instance, includes a new metric that combines three principles. The label contains a five-year (75,000-mile) figure that displays a vehicle's gas costs or savings compared with an average vehicle. For an SUV that gets 14 MPG, this figure is quite large: It is roughly $10,000 in extra costs to own the vehicle. This new metric combines scale expansion (75,000 miles), translation to an end objective (cost), and a relative comparison (to an average vehicle) that makes good and bad outcomes more salient. On the basis of our research, there is reason to believe that combining principles in this way should better inform car buyers, but the benefits of the combination approach have not been empirically tested. Existing field research on the use of descriptive norms and of energy savings goals finds reductions between 2% and 10%.[22–27] Empirical tests are needed to assess whether different combinations of the four principles could increase energy savings further.

One challenge in redesigning the EPA label was the need to create a common metric that allows the comparison of traditional vehicles that run on gasoline and newer vehicles that run on electricity. The solution was to report a metric called *MPGe,* which stands for *MPG equivalent.* Equivalence is achieved by calculating the amount of electricity equal to the amount of energy produced by burning a gallon of gasoline and then calculating the miles an electric vehicle can drive on that amount of electricity. On the basis of the principles we have proposed, this metric is a poor one. First, it inherits all of the problems of MPG—it leads people to underestimate the benefits of improving inefficient

vehicles and to overestimate the benefits of improving efficient vehicles. Second, it completely obscures both the cost and the environmental implications of the energy source, which are buried in the denominator. A better approach would be to express the cost and environmental implications of the energy source over a given distance of driving. This is not a trivial undertaking because the cost and environmental implications of electricity vary widely across the United States depending on regulation and the relative reliance on coal, natural gas, hydropower, or other renewables to produce electricity (to address this challenge, the U.S. Department of Energy provides a zip code–based cost and carbon calculator for all vehicles: http://www.afdc.energy.gov/calc/). Despite the challenges, this information would be more useful to consumers than the confusing MPGe metric.

Although we have proposed the CORE principles in the context of energy consumption information, the same principles may be useful when providing information about a wide range of consumer choices. For example, the federal Affordable Care Act requires chain restaurants to provide calorie information about their menu items by the end of 2015. Although some studies have found that calorie labeling reduces calorie consumption,[44] the results across studies have been a mix of beneficial and neutral effects.[45,46] The provision of calorie information has a larger effect, however, if a relative comparison is offered, such as when there is a list of alternatives from high to low calorie;[47] when calorie counts are compared with recommended daily calorie intake;[48] or when calorie levels are expressed using traffic light colors of green, yellow, and red.[49] There is also limited evidence that translating calories to another objective, the amount of exercise required to burn an equivalent number of calories, also reduces consumption.[50,51] Although we know of no existing studies testing it, the expansion principle might also be of use in the food domain. For example, phone apps that count calories consumed and burned in a given day could provide estimates of weight loss or weight gain if those same behaviors occurred over a month. Dieters might be motivated by seeing a small number scaled up to something relevant to an objective as important as expected weight loss. Research exploring how the principles influence choices in disparate domains, such as energy consumption and obesity-reduction projects, might be useful to both areas.

CORE can also be applied to more consumer domains if the *C* is broadened from *consumption* to include *calculations* of many kinds. MPG is a misleading measure because its relationship to gas consumption is highly nonlinear. A GPHM metric is helpful because it does the math for consumers. There are other nonlinear relationships that consumers face for which calculations would be helpful. Consumers systematically underestimate the beneficial effects of compounding on retirement savings[52] and the detrimental effects of compounding on unpaid credit card debt.[53] Explicitly providing these calculations is helpful in both cases. A familiar product, sunscreen, also has a misleading curvilinear relationship. Sunscreen is measured using a *sun-protection-factor* (SPF) score that might range in value from 15 to 100, which captures the number of minutes a consumer could stay in the sun to achieve the same level of sunburn that results from one minute of unprotected exposure. A more meaningful number, however, is the percentage of radiation blocked by the sunscreen. This is calculated by subtracting 1/SPF from 1 and reveals the similarity of all sunscreens above 30 SPF. A 30-SPF sunscreen blocks 97% of UV radiation, and a 50-SPF sunscreen blocks 98% of UV radiation. Dermatologists consider any further differentiation above 50-SPF pointless,[54] and regulators in Japan, Canada, and Europe cap SPF values at 50.[55]

When one is trying to make the most of the CORE principles described above, it is important to consider how much as well as what kind of information to provide to help people choose. Too much information can be overwhelming. Consider food nutrition labels. They contain dozens of pieces of information that are hard to evaluate and hard to directly translate to end objectives such as minimizing weight gain or protecting heart health. Thus, we believe that *simplicity* is also an important principle when providing information (and can be added as the first letter in a modified acronym, *SCORE*). Simplicity is at odds with multiple translations. To reconcile this conflict, we propose the idea of *minimal coverage:* striving to cover diverse end objectives with a minimum of information. The revised EPA label succeeds here. It is not too cluttered and conveys a minimal set of distinct information (energy, costs, and greenhouse gas impacts) to allow consumers with different values to recognize and act on objectives they care about. Of course, a focus on one primary thing—energy use—requires only a few possible translations.

Feasibility and Acceptability

Thanks to the best-selling 2008 book *Nudge: Improving Decisions About Health, Wealth, and Happiness* by Thaler and Sunstein,[10] behavioral interventions to help consumers are often termed *nudges* because they encourage a change in behavior without restricting choice. However, there has been recent debate over both the ethics and the political feasibility of implementing nudges to influence consumer behavior. We believe it is useful to evaluate nudges in terms of how they operate psychologically. Some nudges steer behavior by tapping known psychological tendencies that people have but are not aware of. Others try to guide decisionmakers by improving their decision processes. Perhaps the best known steering nudge is the use of default options to influence choice. Decisionmakers who are required to start with one choice alternative, such as being enrolled in a company retirement plan[56] or being registered as an organ donor,[57] tend to stick with the first alternative—the default—when given the option to opt out. Consequently, those who must opt out end up selecting the default option at a much higher rate than those who must actively opt in to get the same alternative. Defaults tap a number of known psychological tendencies such as a bias for the status quo and inertia, which people exhibit without being aware they are doing so.[58] Guiding nudges, on the other hand, tend to offer information that consumers care about and make it easy to use—examples include informing credit card users that paying the minimum each month will trap them in debt for 15 years and double their total interest costs compared with paying an amount that would allow them to pay off the debt in three years.[53]

Two of the CORE principles we propose are guiding nudges. Both consumption metrics and expanded scales improve information processing by delivering relevant, useful math. The two remaining principles, however, both guide and steer. Translating energy to costs and environmental impacts improves the decisionmaking process by calling people's attention to objectives they care about and providing a signpost for achieving them. The practice also taps into a basic psychological tendency, counting, that makes efficient options more attractive. The revised EPA label, for instance, may encourage counting when it displays multiple related benefits of efficient vehicles. Similarly, relative comparisons improve information processing by providing a frame of reference for evaluating otherwise murky energy information. However, comparison also taps into a powerful psychological tendency: the desire to achieve good outcomes and the even stronger desire to avoid bad ones. As we have explained, there are many possible comparisons, such as the energy used by an average neighbor or an energy reduction goal, and no comparison is obviously the right one to use.

We emphasize that although the CORE principles we advance are designed to make energy information more usable, they may not always yield stronger preferences for energy reduction. For example, consumption metrics make clear that improvements on inefficient technologies can yield large reductions in consumption (and in costs and environmental impact). They also make clear that large efficiency gains on already efficient technologies, such as trading in a 50-MPG hybrid for a 100-MPGe plug-in or a 16-SEER air-conditioning unit for a 24-SEER air-conditioning unit, will be very expensive but yield only small absolute savings in energy and cost. If some car buyers who would have bought a 16-MPG vehicle now see the benefits of choosing a 20-MPG vehicle, other buyers may no longer trade in their 30-MPG sedan for a 50-MPG hybrid.[59] An interesting empirical question is whether other motivations, such as a strong interest in the environment, will keep the already efficiency-minded segment pushing toward the most efficient technologies for intrinsic reasons. Alternatively, consumers who value environmental conservation may choose to shift their attention from one technology to another (from automobiles to household energy use, for instance) once it is apparent they have achieved a low level of energy consumption in the first technology.[60]

We recognize that better energy metrics can have only limited impact. Better metrics can improve and inform decisions and remind people of what they value, but they may do little to change people's attitudes about energy or the environment. There is a growing literature on political differences in environmental attitudes and the motivations that lead people to be open to or resist energy efficiency as a solution to climate change.[19,20,61,62] An understanding of what motivates people to be concerned with energy use complements this article's focus on how best to provide information. In addition, better energy metrics will not influence

behavior as powerfully as policy levers such as raising the Corporate Average Fuel Economy standards to 54.5 MPG, for example, or raising fossil fuel prices to reflect their environmental costs. However, designing better energy metrics is politically attractive because they represent a low-cost intervention that focuses primarily on informing consumers while preserving their freedom to choose.

Even though the benefit of any given behavioral intervention may be modest,[22] pursuing and achieving benefits from multiple interventions can have a large impact as larger political and technological solutions are pursued.[4,63] Moreover, better energy metrics can make future political and technological developments more powerful. If cultural shifts produce greater concern for the environment, or political shifts lead to mechanisms that raise the cost of fossil fuels to reflect their environmental impacts, a clear understanding of energy consumption and its impacts would empower consumers to respond more effectively to such policy changes.

author affiliation

Larrick, Soll, and Keeney, Fuqua School of Business, Duke University. Corresponding author's e-mail: larrick@duke.edu

References

1. Attari, S. Z., DeKay, M. L., Davidson, C. I., & de Bruin, W. B. (2010). Public perceptions of energy consumption and savings. *PNAS: Proceedings of the National Academy of Sciences, USA, 107,* 16054–16059.
2. Gillingham, K., Newell, R., & Palmer, K. (2006). Energy efficiency policies: A retrospective examination. *Annual Review of Environment and Resources, 31,* 161–192.
3. Gillingham, K., & Palmer, K. (2014). Bridging the energy efficiency gap: Policy insights from economic theory and empirical evidence. *Review of Environmental Economics and Policy, 8,* 18–38.
4. Dietz, T., Gardner, G. T., Gilligan, J., Stern, P. C., & Vandenbergh, M. P. (2009). Household actions can provide a behavioral wedge to rapidly reduce US carbon emissions. *PNAS: Proceedings of the National Academy of Sciences, USA, 106,* 18452–18456.
5. Office of Transportation and Air Quality & National Highway Traffic Safety Administration. (2010). *Environmental Protection Agency fuel economy label: Expert panel report* (EPA-420-R-10-908). Retrieved from http://www.epa.gov/fueleconomy/label/420r10908.pdf
6. Hincha-Ownby, M. (2010). Consumers confused by proposed EPA car labels. Retrieved from http://www.mnn.com/green-tech/transportation/stories/consumers-confused-by-proposed-epa-car-labels
7. Doggett, S. (2011). EPA unveils smart new fuel economy labels. Retrieved from http://www.edmunds.com/autoobserver-archive/2011/05/epa-unveils-smart-new-fuel-economy-labels.html
8. Kahneman, D. (2011). *Thinking, fast and slow.* New York, NY: Farrar, Straus and Giroux.
9. Johnson, E. J., Shu, S.B., Dellaert, B. G. C., Fox, C.R., Goldstein, D.G., Haubl, G., . . . Weber, E. U. (2012). Beyond nudges: Tools of a choice architecture. *Marketing Letters, 23,* 487–504.
10. Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving decisions about health, wealth, and happiness*. New Haven, CT: Yale University Press.
11. Larrick, R. P., & Soll, J. B. (2008, June 20). The MPG illusion. *Science, 320,* 1593–1594.
12. Plumer, B. (2011, November 5). Was "cash for clunkers" a clunker? [Blog post]. Retrieved from http://www.washingtonpost.com/blogs/wonkblog/post/was-cash-for-clunkers-a-clunker/2011/11/04/gIQA42EhpM_blog.html
13. National Research Council. (2010). *Technologies and approaches to reducing the fuel consumption of medium- and heavy-duty vehicles.* Washington, DC: National Academies Press.
14. National Research Council. (2011). *Assessment of fuel economy technologies for light-duty vehicles.* Washington, DC: National Academies Press.
15. Larrick, R. P., & Cameron, K. W. (2011). Consumption-based metrics: From autos to IT. *Computer, 44,* 97–99.
16. Keeney, R. L. (1992). *Value-focused thinking: A path to creative decision making*. Cambridge, MA: Harvard University Press.
17. Ungemach, C., Camilleri, A. R., Johnson, E. J., Larrick, R. P., & Weber, E. U. (2014). *Translated attributes as a choice architecture tool*. Durham, NC: Duke University.
18. Weber, M., Eisenführ, F., & Von Winterfeldt, D. (1988). The effects of splitting attributes on weights in multiattribute utility measurement. *Management Science, 34,* 431–445.
19. Costa, D. L., & Kahn, M. E. (2013). Energy conservation "nudges" and environmentalist ideology: Evidence from a randomized residential electricity field experiment. *Journal of the European Economic Association, 11,* 680–702.
20. Gromet, D. M., Kunreuther, H., & Larrick, R. P. (2013). Political identity affects energy efficiency attitudes and choices. *PNAS: Proceedings of the National Academy of Sciences, USA, 110,* 9314–9319.
21. Schultz, P. W., Nolan, J., Cialdini, R., Goldstein, N., & Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science, 18,* 429–434.
22. Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics, 95,* 1082–1095.
23. Ayres, I., Raseman, S., & Shih, A. (2013). Evidence from two large field experiments that peer comparison feedback can reduce residential energy usage. *Journal of Law, Economics, and Organization, 29,* 992–1022.
24. Ferraro, P. J., & Miranda, J. J. (2013). Heterogeneous treatment effects and mechanisms in information-based environmental policies: Evidence from a large-scale field experiment. *Resource and Energy Economics, 35,* 356–379.
25. Ferraro, P. J., Miranda, J. J., & Price, M. K. (2011). The persistence of treatment effects with norm-based policy

instruments: Evidence from a randomized environmental policy experiment. *American Economic Review, 101,* 318–322.
26. Allcott, H., & Rogers, T. (in press). The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation. *American Economic Review.*
27. Dolan, P., & Metcalfe, R. (2013). *Neighbors, knowledge, and nuggets: Two natural field experiments on the role of incentives on energy conservation* (CEP Discussion Paper CEPDP1222). London, United Kingdom: London School of Economics and Political Science, Centre for Economic Performance.
28. Hsee, C. K., Loewenstein, G. F., Blount, S., & Bazerman, M. H. (1999). Preference reversals between joint and separate evaluations of options: A review and theoretical analysis. *Psychological Bulletin, 125,* 576–590.
29. Newell, R. G., & Siikamaki, J. (2014). Nudging energy efficiency behavior: The role of information labels. *Journal of the Association of Environmental and Resource Economists, 1,* 555–598.
30. Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47,* 263–292.
31. Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *Quarterly Journal of Economics, 106,* 1039–1061.
32. Heath, C., Larrick, R. P., & Wu, G. (1999). Goals as reference points. *Cognitive Psychology, 38,* 79–109.
33. Carrico, A. R., & Riemer, M. (2011). Motivating energy conservation in the workplace: An evaluation of the use of group-level feedback and peer education. *Journal of Environmental Psychology, 31,* 1–13.
34. Harding, M., & Hsiaw, A. (2014). Goal setting and energy conservation. *Journal of Economic Behavior and Organization, 107,* 209–227.
35. Houde, S. (2014). *How consumers respond to environmental certification and the value of energy information* (NBER Working Paper w20019). Cambridge, MA: National Bureau of Economic Research.
36. Burson, K. A., Larrick, R. P., & Lynch, J. G., Jr. (2009). Six of one, half dozen of the other: Expanding and contracting numerical dimensions produces preference reversals. *Psychological Science, 20,* 1074–1078.
37. Hardisty, D. J., Shim, Y., & Griffin, D. (2014). *Encouraging energy efficiency: Product labels activate temporal tradeoffs*. Vancouver, British Columbia, Canada: University of British Columbia Sauder School of Business. Contact David Hardisty.
38. Kaenzig, J., & Wüstenhagen, R. (2010). The effect of life cycle cost information on consumer investment decisions regarding eco-innovation. *Journal of Industrial Ecology, 14,* 121–136.
39. Pandelaere, M., Briers, B., & Lembregts, C. (2011). How to make a 29% increase look bigger: The unit effect in option comparisons. *Journal of Consumer Research, 38,* 308–322.
40. Camilleri, A. R., & Larrick, R. P. (2014). Metric and scale design as choice architecture tools. *Journal of Public Policy and Marketing, 33,* 108–125.
41. Sunstein, C. R. (2011). Empirically informed regulation. *University of Chicago Law Review, 78,* 1349–1429.
42. Hardisty, D. J., & Weber, E. U. (2009). Discounting future green: Money versus the environment. *Journal of Experimental Psychology: General, 138,* 329–340.
43. Weber, E. U., Johnson, E. J., Milch, K. F., Chang, H., Brodscholl, J. C., & Goldstein, D. G. (2007). Asymmetric discounting in intertemporal choice: A query-theory account. *Psychological Science, 18,* 516–523.
44. Bollinger, B., Leslie, P., & Sorenson, A. (2011). Calorie posting in chain restaurants. *American Economic Journal: Economic Policy, 3,* 91–128.
45. Kiszko, K. M., Martinez, O. D., Abrams, C., & Elbel, B. (2014). The influence of calorie labeling on food orders and consumption: A review of the literature. *Journal of Community Health, 39,* 1–22.
46. Sinclair, S. E., Cooper, M., & Mansfield, E. D. (2014). The influence of menu labeling on calories selected or consumed: A systematic review and meta-analysis. *Journal of the Academy of Nutrition and Dietetics, 114,* 1375–1388.
47. Liu, P. J., Roberto, C. A., Liu, L. J., & Brownell, K. D. (2012). A test of different menu labeling presentations. *Appetite, 59,* 770–777.
48. Roberto, C. A., Larsen, P. D., Agnew, H., Baik, J., & Brownell, K. D. (2010). Evaluating the impact of menu labeling on food choices and intake. *American Journal of Public Health, 100,* 312–318.
49. Thorndike, A. N., Sonnenberg, L., Riis, J., Barraclough, S., & Levy, D. E. (2012). A 2-phase labeling and choice architecture intervention to improve healthy food and beverage choices. *American Journal of Public Health, 102,* 527–533.
50. Bleich, S. N., Herring, B. J., Flagg, D. D., & Gary-Webb, T. L. (2012). Reduction in purchases of sugar-sweetened beverages among low-income Black adolescents after exposure to caloric information. *American Journal of Public Health, 102,* 329–335.
51. James, A., Adams-Huet, B., & Shah, M. (2014). Menu label displaying the kilocalorie content or the exercise equivalent: Effects on energy ordered and consumed in young adults. *American Journal of Health Promotion, 29,* 294–302.
52. McKenzie, C. R. M., & Liersch, M. J. (2011). Misunderstanding savings growth: Implications for retirement savings behavior. *Journal of Marketing Research, 48*(SPL), S1–S13.
53. Soll, J. B., Keeney, R. L., & Larrick, R. P. (2013). Consumer misunderstanding of credit card use, payments, and debt: Causes and solutions. *Journal of Public Policy and Marketing, 32,* 66–81.
54. Harris, G. (2011, June 14). F.D.A. unveils new rules about sunscreen claims. *The New York Times*. Retrieved from http://www.nytimes.com
55. EWG. (n.d.). What's wrong with high SPF? Retrieved from http://www.ewg.org/2014sunscreen/whats-wrong-with-high-spf/
56. Thaler, R. H., & Benartzi, S. (2004). Save More Tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy, 112,* S164–S187.
57. Johnson, E. J., & Goldstein, D. (2003, November 21). Do defaults save lives? *Science, 302,* 1338–1339.
58. Smith, N. C., Goldstein, D. G., & Johnson, E. J. (2013). Choice without awareness: Ethical and policy implications of defaults. *Journal of Public Policy and Marketing, 32,* 159–172.
59. Allcott, H. (2013). The welfare effects of misperceived product costs: Data and calibrations from the automobile market. *American Economic Journal: Economic Policy, 5,* 30–66.
60. Truelove, H. B., Carrico, A. R., Weber, E. U., Raimi, K. T., & Vandenbergh, M. P. (2014). Positive and negative spillover of pro-environmental behavior: An integrative review and theoretical framework. *Global Environmental Change, 29,* 127–138.
61. Campbell, T., & Kay, A. C. (2014). Solution aversion: On the relation between ideology and motivated disbelief. *Journal of Personality and Social Psychology, 107,* 809–824.
62. Feinberg, M., & Willer, R. (2013). The moral roots of environmental attitudes. *Psychological Science, 24,* 56–62.
63. Pacala, S., & Socolow, R. (2004, August 13). Stabilization wedges: Solving the climate problem for the next 50 years with current technologies. *Science, 305,* 968–972.

Finding

Payer mix & financial health drive hospital quality: Implications for value-based reimbursement policies

Matthew Manary, Richard Staelin, William Boulding, & Seth W. Glickman

Summary. Documented disparities in health care quality in hospitals have been associated with patients' race, gender, age, and insurance coverage. We used a novel data set with detailed hospital-level demographic, financial, quality-of-care, and outcome data across 265 California hospitals to examine the relationship between a hospital's financial health and its quality of care. We found that *payer mix*, the percentage of patients with private insurance coverage, is the key driver of a hospital's financial health. This is important because a hospital's financial health influences its quality of care and patient outcomes. Government policies that financially penalize hospitals on the basis of care quality and/or outcomes may disproportionately impair financial performance and quality investments at hospitals serving fewer privately insured patients. Such policies could exacerbate health disparities among patients at greatest risk of receiving substandard care.

In recent years, the availability of data measuring the quality of health care in hospitals has expanded dramatically. One important observation is that hospitals with higher numbers of racial minorities and poor people in their patient populations provide lower quality care. A critical question for policymakers is this: Where do these disparities originate? Do they primarily reflect differences in treatment based on patient demographic factors? We explore a second explanation, that disparities may be driven by the underlying financial health of hospitals. Minority and poorer populations are more likely to be under- or uninsured. If hospitals receive lower reimbursements for their services to these populations, they are less able to make the investments that hospitals need to ensure quality care for all patients. Testing for such a possibility requires the right kind of data (demographic, financial, and clinical) and a robust analysis that looks at multiple relevant variables over time.

We began our research into this area aware of evidence that financial health may be a very important driver of quality of care. For one, studies that look at health care quality measures within individual hospitals find much smaller correlations between patients' race

Manary, M., Staelin, R., Boulding, W., & Glickman, S. W. (2015). Payer mix & financial health drive hospital quality: Implications for value-based reimbursement policies. *Behavioral Science & Policy, 1*(1), pp. 77–84.

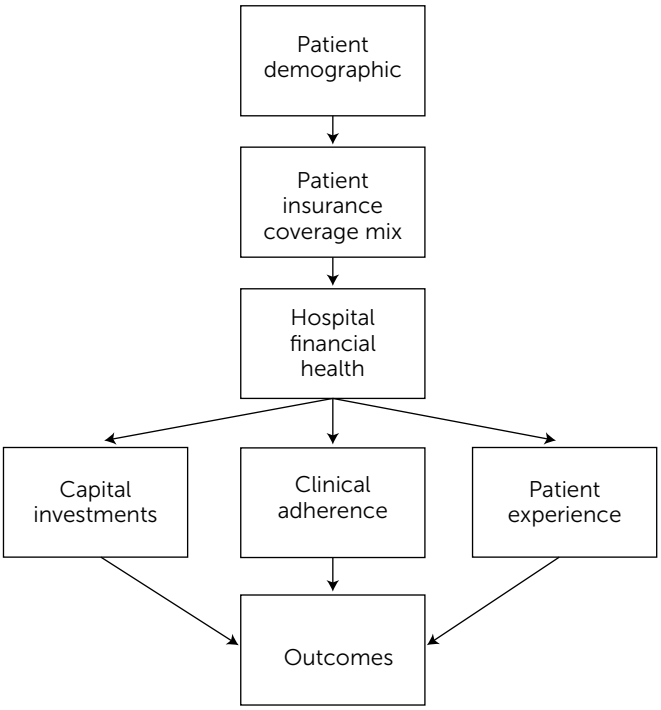
or income and lower quality than do cross-sectional studies that look for relationships by comparing performance across hospitals.[1–3] Another clue is research by Dranove and White dating back to the 1990s.[4] In a longitudinal analysis of how multiple hospitals reacted to Medicare and Medicaid payment reductions in the 1980s and early 1990s, they found that hospitals did not compensate for these reductions by raising prices for patients with private insurance. Instead, they tended to treat the quality of care as a somewhat consistently provided *public good* within their hospital. Thus, the quality of care declined for all patients, albeit more for Medicaid and Medicare patients.

Understanding what causes these disparities is vital today. Medicare, for instance, is shifting from a payment structure based solely on quantity or intensity of services at hospitals to one that creates incentives for improving the quality of health care services.[5,6] For example, the Hospital Value-Based Purchasing Program of the Centers for Medicare & Medicaid Services (CMS) ties hospital Medicare payments to performance in quality measures, outcomes, efficiency, and patient experience. Because these policies are designed also, in part, to limit costs, the incentive programs by design create a system of winners (those that receive financial rewards for high quality) and losers (those that receive financial penalties for low quality). Our findings suggest that such penalties could unintentionally drive quality even lower at already low-performing hospitals. That is, the current rewards and penalties system may lead to institutionalizing inferior health care at hospitals that serve patients at the greatest risk of receiving lower quality care.

What Drives Health Outcomes?

To better understand the factors that ultimately impact health outcomes, we developed a model that recognizes the complex interplay between patient characteristics, reimbursement, organizational behavior, and quality of care and health outcomes. We extended a classic quality assessment framework by Donabedian,[7] which identifies measurable components that contribute to the quality of care in hospitals. This approach allowed us to relate quality of care and health outcomes to organizational behaviors as expressed through capital investments, clinical adherence to standard guidelines, and reported patient experiences. Our resulting hospital quality framework (see Figure 1) was built on the premise that the demographics of a hospital's patient population are significantly correlated with its payer mix, called here the patient insurance coverage mix. Data showing that Spanish-speaking and African American patients are significantly less likely than White patients to have health care insurance support this approach.[8] Caring for substantial numbers of patients without insurance decreases a hospital's revenue. Less income may degrade a hospital's financial health, which leads to lower investment in personnel, information technology, and other key contributors to quality care. Therefore, changes in a hospital's demographic or financial structure (possibly among other factors, many of which we control for in our analyses) will affect downstream institutional processes and, consequently, the quality of care (see Figure 1).

Figure 1. Hospital quality framework



We built our model using a variety of health care quality data from four major sources. The first was the California Office of Statewide Health Planning and Development (COSHPD), from whose website (http://www.oshpd.ca.gov/Healthcare-Data.html) we pulled information for general and acute care hospitals with at

least two years of consecutive data from 2005 to 2011. This source provided detailed audited financial data, which helps overcome the limitations of using cost-accounting data from Medicare cost reports.[9] We also accessed information on payer insurance coverage, patient characteristics such as race, and hospital controls (for example, ownership status, capital investment changes, and licensed bed count).

Our second data source was Yale University's Center for Outcomes Research & Evaluation, which provided annual hospital 30-day risk-standardized readmission and mortality rates for three clinical areas (acute myocardial infarction, heart failure, and pneumonia) for the period 2005–2010. Using annual data rather than CMS's publicly available three-year aggregate data allowed us to better control for unobserved factors and test for causality.

Our third source was the Hospital Compare database compiled by the U.S. Department of Health and Human Services: http://www.medicare.gov/hospitalcompare/search.html. From this database, we obtained data on annual adherence to clinical guidelines for the same three clinical areas for the calendar years 2005–2010.

The fourth source was the annual Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) survey for the period 2007–2010, from which we obtained patient assessments of their in-hospital care experiences. Note that these experiences were not limited to the above-mentioned clinical areas. Survey scores were adjusted by CMS to account for factors believed to affect patient responses but do not control for patient ethnicity or form of payment.[10]

From these sources, we used multiple measures whenever possible for each component of the quality framework shown in Figure 1. Thus, our results reflect an aggregate view of a hospital's performance and are not indicators of any individual patient status, experience, or outcome, nor do they reveal the performance of a specific clinical area within a hospital.

Our model required annual financial and patient information for the hospitals included in our study. We constructed our data set through a process of elimination. First, we identified 485 health care facilities that reported in California's COSHPD financial database and 515 health care facilities that reported patient demographics, payer coverage, and hospital characteristics (not all facilities were acute care hospitals). We cross-referenced the additional data sources (see above and the Supplemental Material) to find 30-day risk-standardized readmission and mortality rates, adherence to clinical guidelines, and patient surveys. Our final study population was 265 acute care hospitals in California that had complete information for at least two consecutive years and also maintained a one-to-one relation with a Medicare provider number from 2005 through 2010.

This final data set allowed us to draw on the strengths of comparisons both within and between hospitals. In general, analyses across multiple institutions can be useful for identifying correlations between factors such as health outcomes and patient demographics. However, they cannot determine if one factor causes another because they cannot control for unobserved factors that affect the dependent variable of interest and that differ between institutions.[11] In contrast, analyses conducted within a single hospital are more revealing of causal relationships because they hold fixed many of these unobserved factors. That said, considerations unique to each institution might limit the ability to generalize the results. Having data from the same hospitals over multiple years allowed us to control for unobserved fixed and autocorrelated effects while increasing the number and breadth of the hospitals analyzed, thereby allowing us to identify relationships applicable across a variety of health care organizations.

An overview of our data set confirmed that the sample contained data points across a wide enough range for each variable to allow us to estimate relationships. We also compared the general characteristics of our California hospital sample with those of the national hospital data set. Statistical tests show that for the majority of variables recorded, there were no significant differences between our sample and the national sample. However, the hospitals in our sample were larger overall and had lower clinical adherence for pneumonia, higher mortality rates for pneumonia, and lower patient satisfaction. With this noted, we observe that these comparisons suggest that the relationships we identified here are likely to apply to a wider range of health care organizations as well. (Much more detail on our measures and tables of our results are available for review in our Supplemental Material.)

Patient Populations and Hospital Performance

We used several common metrics, described briefly below, to assess different aspects of patient populations and hospital performance.

Patient Demographics and Patient Insurance Coverage

Using the COSHPD database, we calculated the annual percentage of patients covered by private insurers for each hospital (the patient insurance coverage mix), the percentage of underrepresented minorities (African American, Hispanic, and Native American) served by the hospital, and the percentage of a hospital's patients who were 60 years of age or older.

Financial Health

We measured the financial health of a hospital in any given year using the DuPont System, which is widely used in financial statement analysis to assess the overall financial health of an institution.[12] The DuPont System includes three key financial ratios that reflect different aspects of financial health. *Current ratio* provides information about the institution's ability to meet its short-term financial obligations. *Gross operating margin* is a good indicator of the institution's ability to generate profits. And *return on assets* captures how efficiently the institution uses its assets. As detailed in our Supplemental Material, we standardized and combined these ratios to create a single measure of the hospital's annual financial health. This measure reflects a hospital's access to the resources needed to deliver high-quality care, such as staff, managerial talent, and physical assets. Higher scores indicate better financial performance.

Clinical Adherence

We used care performance measures from CMS's Hospital Compare database to report how well a hospital met the objective standards associated with high-quality medical care for each of three clinical areas: acute myocardial infarction, heart failure, and pneumonia. As described further in our Supplemental Material, we created a single measure of the hospital's clinical quality in a given year relative to the other 264 hospitals in our database. For this measure, higher scores reflect greater adherence to clinical standards, an indicator of better care.

Patient Experience

The HCAHPS database contains average patient assessments on 10 dimensions of patient care, derived from 18 survey questions. To generate a single annual hospital value for overall patient experience, we combined responses to two hospital-specific questions ("How do you rate the hospital overall?" and "Would you recommend the hospital to friends and family?"). These two dimensions reflect overall service quality[13,14] and have been found to capture patients' overall satisfaction with their hospital experience.[15] They are also important predictors of health outcomes such as mortality and readmission, as observed across multiple clinical areas and hospital services.[16,17] These yearly aggregated measures were then standardized (see the Supplemental Material for details). As with HCAPHS, better patient experiences are associated with higher scores for this measure.

Hospital Infrastructure

Prior work has shown that hospital investment in infrastructure such as equipment is related to outcomes and quality screens.[18–20] We captured each hospital's new annual capital investment on the basis of annual percentage of change in equipment and net depreciation as determined from audited financial records, which we then standardized across the population within each year. Larger values are associated with greater levels of investment.

Hospital Outcomes

We used two common quality measures, hospital-level 30-day risk-standardized mortality rates and readmission rates, which control a particular hospital's outcome rates for patient demographics (gender and age), cardiovascular condition, and other existing health conditions. As detailed in the Supplemental Material, we combined these two measures for each of our three clinical areas to create a single hospital-wide quality index for each hospital and each year. As with the above measures, this measure should be viewed as a good but not perfect

hospital-level measure of the quality of health care. In this case, smaller values represent better outcomes.

Control Measures

We also controlled for other hospital-observed factors that are not of primary interest in our model but are commonly used in hospital financial research,[9,21] including number of licensed beds, teaching hospital status, ownership (for example, investor, government, or nonprofit), and presence of 24-hour emergency services.

Hospital Finances and Health Care Outcomes

Our primary objective was to identify links between a hospital's patient population and its quality of care, then evaluate whether those relationships are mediated by the financial health of the hospital. We first looked at our data set for evidence that variation in patient demographics, including ethnicity, correlated with variations in health care quality. Using a regression analysis statistical approach, we tested whether the percentage of underrepresented minorities was directly associated with the three performance measures that CMS uses in its pay-for-performance programs: clinical adherence, patient experience, and hospital outcomes. (Note that CMS controls for age when reporting patient experience and outcomes.) Much like the previous studies we mentioned earlier, we found highly statistically significant results showing that hospitals that treated higher percentages of minority patients reported lower clinical adherence scores, worse patient experiences, and poorer health outcomes. However, this regression analysis is designed only to show correlation between factors, not whether one directly causes another.

Given our interest in assessing causality, we next defined a series of linear models to test the relationships we proposed in Figure 1. We used these models to address four main issues. First, the models help identify factors that might separately explain an observed correlational relationship between the variables in question. They do this by controlling for some aspects of unobserved variables (such as managerial expertise) that might cut across equations and/or are related to the independent and dependent variables and thus could affect both. Second, the models test whether an observed statistical association (such as between ethnic status and measures of financial health) can be accounted for by an intermediate variable (such as insurance status). Third, the models test whether our results might be explained by unaccounted-for contemporaneous factors (for example, economic shocks that lead to lower employment levels, which, in turn, lead to sicker patients because of postponed health care). And finally, the models are used to test for causality among the factors described in Figure 1. We analyzed causality using a methodology proposed by Clive Granger that uses past observations of the dependent variable (such as quality of health care) as a control and then looks to see if an independent variable (such as insurance reimbursements) causes changes in the dependent variable after including additional control variables (such as demographics).[22] The models testing the Figure 1 relationships and their main findings are described below.

1. *Is a hospital's patient insurance coverage mix determined by its patient demographics?* We found that hospitals that treated higher percentages of patients from underrepresented minority populations had fewer privately insured patients.

2. *Is a hospital's financial health determined by its patient insurance coverage mix?* Institutions with a higher percentage of privately insured patients also demonstrated better financial performance. Although hospitals that treat greater numbers of older patients and underrepresented minorities have poorer financial health, these effects are completely mediated once the percentage of privately insured patients is included in the model. That is, the age and racial composition of a patient population are not related to the financial health of a hospital once the insurance coverage of the patients is known. When we tested for causality, we found that the percentage of privately insured patients significantly affects hospital financial performance in the subsequent year. This latter point highlights the potentially complex and long-lasting impact payer coverage has on a hospital's financial health and, indirectly, its ability to provide quality care both today and in the future.

3–5. *Are patient experiences, clinical adherence, and investment in equipment, respectively, determined by the hospital's financial health?* Together, these three separate analyses showed that a hospital's financial health seems to have widespread impact on institutional decisionmaking and structure. Both clinical

performance and changes in equipment investment correlated with the institution's financial health, although patient experiences did not. However, when we tested for causality, we found that last year's financial health negatively affected not only this year's investment in equipment and clinical performance but also this year's patient experience scores.

6. *Are hospital outcomes determined jointly by the hospital's patient experiences, clinical adherence, and investment in equipment?* We found that better adherence to clinical guidelines and positive patient experiences were associated with better hospital-wide outcomes, even after controlling for the effects of the other factors (including investment in equipment).

Implications for Health Care Policy

Our analyses, which are very supportive of the relationships proposed in Figure 1, provide a number of important insights useful to policymakers and researchers. Our results show empirically that the payer mix of a hospital's patients affects the quality of its services and patient outcomes. This is largely due to the payer mix's effects on a hospital's financial condition rather than its patient demographic profile. Controlling for payer coverage absorbed most if not all of the relationship between patient demographics and quality measures. We say "most" because the percentage of privately insured patients did not mediate the relationship between minority percentage and clinical adherence. However, when the percentage of privately insured patients was exchanged for the percentage of payers on Medicaid, demographics were no longer significant. Moreover, because our data do not allow us to identify payment coverage by demographic group within a hospital, we cannot say that demographics play no part in determining quality of care; however, failing to account for payment sources will likely overstate demographic effects.

To provide insights into the magnitude of impact that the hospital's financial health has on downstream measures of performance and outcomes, we segmented our sample into three groups: hospitals in the top 20% of financial health in 2007 (our first year with complete measures), hospitals in the bottom 20%, and those in between. We compared the average performance in patient HCAHPS scores, clinical adherence, and investment in equipment for the top and bottom groups to show the actual average performance for these three downstream measures. Hospitals in the top 20% of financial health, for instance, invested more heavily in equipment (9.3% vs. 8.1%), scored 7 points higher on HCAHPS (80 vs. 73), and scored higher in clinical adherence for heart attack, heart failure, and pneumonia (3.5, 7.7, and 6.7 points higher, respectively). For an average-sized hospital from our sample, our model predicts that being in the top 20% of infrastructure investment, clinical adherence performance, and HCAHPS scores in aggregate in a given year resulted in 6.5 fewer deaths that year (0.4 heart attack, 1.1 heart failure, 5.0 pneumonia) and 11.2 fewer readmissions (1.4 heart attack, 4.1 heart failure, 5.7 pneumonia) compared with an average-sized hospital in the bottom 20%. Note that these differences represent the impact on just the 797 patients treated annually in these three clinical areas in this average hospital; the impact of increased financial health on a hospital's full patient population will likely be much greater.

Taken together, these findings imply that failing to adjust CMS's Hospital Value-Based Performance Program (HVBP) and Readmission Reduction Program (RRP) domain scores to account for patient demographics or payer mix could have unintended consequences. That is, it could set up a cycle of imposing financial penalties on already struggling hospitals, which would cause even worse subsequent relative performance, lower HVBP and RRP scores, and further reductions in reimbursement. In their current form, HVBP and RRP may inadvertently institutionalize substandard care for people already at risk of receiving poorer care.[23,24]

A critical facet of fairly administering health care funding programs is to risk-adjust outcome measures to control for factors that are beyond the control of a hospital. That includes the presence and/or severity of certain diseases such as diabetes, so-called *exogenous* factors, but not for hospital characteristics that are within their control, so-called *endogenous* factors.[25] CMS and other quality assessment bodies such as the National Quality Forum do not risk-adjust for factors such as race and socioeconomic status because they do not want to hold hospitals with different patient demographics to different performance standards.[26] Adjusting for race or socioeconomic status could also obscure real differences that would be important to identify wherever they exist. While valid, these concerns need to be balanced against our findings that failing to

adjust for payer mix or demographic factors could have unintended negative effects on organizational finances and resulting health care quality for underserved populations.

Recent findings show that safety-net hospitals in California already are more likely than other hospitals to be penalized financially by hospital-based quality reimbursement programs such as HVBP, RRP, and the electronic health record meaningful-use program.[27] One potential solution is to handle such hospitals, which treat high proportions of underinsured patients, as a discrete cohort for the purposes of calculating Value-Based Purchasing reimbursement adjustments. Policy-makers could channel a greater proportion of incentive payments to these safety-net hospitals and potentially make some of these payments contingent on specified organizational investments in quality management and systems.

Another option would be to directly incorporate patient insurance coverage profiles into the value-based reimbursement formula for hospitals. This risk-adjustment methodology could be separated from formal reporting of quality and outcome metrics to avoid CMS's and the National Quality Forum's explicit concerns about concealing disparities. Finally, the adverse effects that decreasing insurance payments are likely to have on the quality of care for all patients deserve greater attention. That is particularly true in states that have elected not to expand Medicaid under the Affordable Care Act, as also has been highlighted by Gilman et al.[27] In an era of unsustainable cost increases, hospitals are unlikely to be able to shift costs to the private sector at historical levels.[28] Instead, many hospitals may respond by cutting costs in ways that are likely to reduce their ability to provide quality health care,[29] which could adversely affect care for all patients, regardless of their insurance status.

author affiliation

Manary, Staelin, Boulding, Duke University Fuqua School of Business; Glickman, University of North Carolina School of Medicine. Corresponding author's e-mail: swglick@med.unc.edu

author note

This work was supported by research funds from the Fuqua School of Business. The authors thank Dr. Harlan Krumholz and Yale University's Center for Outcomes Research & Evaluation for their support and for providing the team with access to annual hospital-level outcomes data.

supplemental material

- http://behavioralpolicy.org/supplemental-material
- Methods & Analysis
- Data, Analyses & Results
- Additional References

References

1. Dozier, K. C., Miranda, M. A., Kwan, R. O., Cureton, E. L., Sadjadi, J., & Victorino, G. P. (2010). Insurance coverage is associated with mortality after gunshot trauma. *Journal of the American College of Surgeons, 210,* 280–285.
2. Neureuther, S. J., Nagpal, K., Greenbaum, A., Cosgrove, J. M., & Farkas, D. T. (2013). The effect of insurance status on outcomes after laparoscopic cholecystectomy. *Surgical Endoscopy and Other Interventional Techniques, 27,* 1761–1765.
3. Taghavi, S., Jayarajan, S. N., Duran, J. M., Gaughan, J. P., Pathak, A., Santora, T. A., . . . Goldberg, A. J. (2012). Does payer status matter in predicting penetrating trauma outcomes? *Surgery, 152,* 227–231.
4. Dranove, D., & White, W. D. (1998). Medicaid-dependent hospitals and their patients: How have they fared? *Health Services Research, 33,* 163–185.
5. Centers for Medicare & Medicaid Services. (2008). *Roadmap for implementing value driven healthcare in the traditional Medicare fee-for-service program*. Retrieved from http://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/QualityInitiativesGenInfo/downloads/vbproadmap_oea_1-16_508.pdf
6. Conrad, D. A., & Perry, L. (2009). Quality-based financial incentives in health care: Can we improve quality by paying for it? *Annual Review of Public Health, 30,* 357–371.
7. Donabedian, A. (1978, May 26). Quality of medical care. *Science, 200,* 856–864.
8. Mead, H., Cartwright-Smith, L., Jones, K., Ramos, C., Woods, K., & Siegel, B. (2008). *Racial and ethnic disparities in U.S. health care: A chartbook* (Commonwealth Fund Pub. No. 1111). Retrieved from Commonwealth Fund website: http://www.commonwealthfund.org/usr_doc/mead_racialethnicdisparities_chartbook_1111.pdf
9. Bazzoli, G. J., Chen, H. F., Zhao, M., & Lindrooth, R. C. (2008). Hospital financial condition and the quality of patient care. *Health Economics, 17,* 977–995.
10. Centers for Medicare & Medicaid Services. (2008). *Mode and Patient-Mix Adjustment of the CAHPS Hospital Survey (HCAHPS)*. Retrieved from http://www.hcahpsonline.org/files/Final%20Draft%20Description%20of%20HCAHPS%20Mode%20and%20PMA%20with%20bottom%20box%20modedoc%20April%2030,%202008.pdf
11. Boulding, W., & Staelin, R. (1995). Identifying generalizable effects of strategic actions on firm performance: The case of demand-side returns to R&D spending. *Marketing Science, 14*(3, Suppl.), G222–G236.
12. Foster, G. (1978). *Financial statement analysis*. Englewood Cliffs, NJ: Prentice Hall.
13. Boulding, W., Kalra, A., Staelin, R., & Zeithaml, V. A. (1993). A dynamic process model of service quality: From expectations to behavioral intentions. *Journal of Marketing Research, 30,* 7–27.
14. Boulding, W., Kalra, A., & Staelin, R. (1999). The quality double whammy. *Marketing Science, 18,* 463–484.
15. White, B. (1999). Measuring patient satisfaction: How to do it and why to bother. *Family Practice Management, 6*(1), 40–44.
16. Boulding, W., Glickman, S. W., Manary, M. P., Schulman, K. A., & Staelin, R. (2011). Relationship between patient satisfaction with inpatient care and hospital readmission within 30 days. *American Journal of Managed Care, 17,* 41–48.
17. Glickman, S. W., Boulding, W., Manary, M., Staelin, R., Roe, M. T., Wolosin, R. J., . . . Schulman, K. E. (2010). Patient satisfaction and its relationship with clinical quality and inpatient mortality in acute myocardial infarction. *Circulation: Cardiovascular Quality and Outcomes, 3,* 188–195.
18. Cleverley, W. O., & Harvey, R. K. (1992). Is there a link between hospital profit and quality? *Healthcare Financial Management: Journal of the Healthcare Financial Management Association, 46*(9), pp. 40, 42, 44–45.
19. Kuhn, E. M., Hartz, A. J., Gottlieb, M. S., & Rimm, A. A. (1991). The relationship of hospital characteristics and the results of peer review in six large states. *Medical Care, 29,* 1028–1038.
20. Levitt, S. W. (1994). Quality of care and investment in property, plant, and equipment in hospitals. *Health Services Research, 28,* 713–727.
21. Bazzoli, G. J., Clement, J. P., Lindrooth, R. C., Chen, H. F., Aydede, S. K., Braun, S. K., & Loeb, J. M. (2007). Hospital financial condition and operational decisions related to the quality of hospital care. *Medical Care Research and Review, 64,* 148–168.
22. Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica, 37,* 424–438.
23. Fiscella, K., Franks, P., Gold, M. R., & Clancy, C. M. (2000). Inequality in quality: Addressing socioeconomic, racial, and ethnic disparities in health care. *JAMA, 283,* 2579–2584.
24. Karve, A. M., Ou, F.S., Lytle, B. L., & Peterson, E. D. (2008). Potential unintended financial consequences of pay-for-performance on the quality of care for minority patients. *American Heart Journal, 155,* 571–576.
25. Medicare Program; Hospital Inpatient Value-Based Purchasing Program, 76 Fed. Reg. 26,490 (May 6, 2011) (to be codified at 42 C.F.R. pts. 422 and 480).
26. National Quality Forum. (2012). *Measure evaluation criteria.* Retrieved from http://www.qualityforum.org/docs/measure_evaluation_criteria.aspx
27. Gilman, M., Adams, E. K., Hockenberry, J. M., Wilson, I. B., Milstein, A. S., & Becker, E. B. (2014). California safety-net hospitals likely to be penalized by ACA value, readmission, and meaningful-use programs. *Health Affairs, 33,* 1314–1322.
28. Avalere Health & American Hospital Association. (2014). *Trends affecting hospitals and health systems* (TrendWatch Chartbook 2014). Retrieved from http://www.aha.org/research/reports/tw/chartbook/ch1.shtml
29. Robinson, J. (2011). Hospitals respond to Medicare payment shortfalls by both shifting costs and cutting them, based on market concentration. *Health Affairs, 30,* 1265–1271.

editorial policy

Behavioral Science & Policy (BSP) is an international, peer-reviewed publication of the Behavioral Science & Policy Association and Brookings Institution Press. BSP features short, accessible articles describing actionable policy applications of behavioral scientific research that serves the public interest. Articles submitted to BSP undergo a dual-review process: For each article, leading disciplinary scholars review for scientific rigor and experts in relevant policy areas review for practicality and feasibility of implementation. Manuscripts that pass this dual-review are edited to ensure their accessibility to policy makers, scientists, and lay readers. BSP is not limited to a particular point of view or political ideology.

Manuscripts can be submitted in a number of different formats, each of which must clearly explain specific implications for public- and/or private-sector policy and practice.

External review of the manuscript entails evaluation by at least two outside referees—at least one in the policy arena and at least one in the disciplinary field.

Professional editors trained in BSP's style work with authors to enhance the accessibility and appeal of the material for a general audience.

Behavioral Science & Policy charges a $50 fee per submission to defray a portion of the manuscript processing costs. For the first volume of the journal, this fee has been waived.

Each of the sections below provides general information for authors about the manuscript submission process. We recommend that you take the time to read each section and review carefully the BSP Editorial Policy before submitting your manuscript to *Behavioral Science & Policy*.

Manuscript Formats

Manuscripts can be submitted in a number of different formats, each of which must clearly demonstrate the empirical basis for the article as well as explain specific implications for (public and/or private-sector) policy and practice:

- Proposals (≤ 2,500 words) specify scientifically grounded policy proposals and provide supporting evidence including concise reports of relevant studies. This category is most appropriate for describing new policy implications of previously published work or a novel policy recommendation that is supported by previously published studies.
- Findings (≤ 4,000 words) report on results of new studies and/or substantially new analysis of previously reported data sets (including formal meta-analysis) and the policy implications of the research findings. This category is most appropriate for presenting new evidence that supports a particular policy recommendation. The additional length of this format is designed to accommodate a summary of methods, results, and/or analysis of studies (though some finer details may be relegated to supplementary online materials).
- Reviews (≤ 5,000 words) survey and synthesize the key findings and policy implications of research in a specific disciplinary area or on a specific policy topic. This could take the form of describing a general-purpose behavioral tool for policy makers or a set of behaviorally grounded insights for addressing a particular policy challenge.
- Other Published Materials. BSP will sometimes solicit or accept *Essays* (≤ 5,000 words) that present a unique perspective on behavioral policy; *Letters* (≤ 500 words) that provide a forum for responses from readers and contributors, including policy makers and public figures; and *Invitations* (≤ 1,000 words with links to online Supplemental Material), which are requests from policy makers for contributions from the behavioral science community on a particular policy issue. For example, if a particular agency is facing a specific challenge and seeks input from the behavioral science community, we would welcome posting of such solicitations.

Review and Selection of Manuscripts

On submission, the manuscript author is asked to indicate the most relevant disciplinary area and policy area addressed by his/her manuscript. (In the case of some papers, a "general" policy category designation may be appropriate.) The relevant Senior Disciplinary Editor and the Senior Policy Editor provide an initial screening of the manuscripts. After initial screening, an appropriate Associate Policy Editor and Associate Disciplinary Editor serve as the stewards of each manuscript as it moves through the editorial process. The manuscript author will receive an email within approximately two weeks of submission, indicating whether the article has been sent to outside referees for further consideration. External review of the manuscript entails evaluation by at least two outside referees. In most cases, Authors will receive a response from BSP within approximately 60 days of submission. With rare exception, we will submit manuscripts to no more than two rounds of full external review. We generally do not accept re-submissions of material without an explicit invitation from an editor. Professional editors trained in the BSP style will collaborate with the author of any manuscript recommended for publication to enhance the accessibility and appeal of the material to a general audience (i.e., a broad range of behavioral scientists, public- and private-sector policy makers, and educated lay public). We anticipate no more than two rounds of feedback from the professional editors.

Standards for Novelty

BSP seeks to bring new policy recommendations and/or new evidence to the attention of public and private sector policy makers that are supported by rigorous behavioral and/or social science research. Our emphasis is on novelty of the policy application and the strength of the supporting evidence for that recommendation. We encourage submission of work based on new studies, especially field studies (for Findings and Proposals) and novel syntheses of previously published work that have a strong empirical foundation (for Reviews).

BSP will also publish novel treatments of previously published studies that focus on their significant policy implications. For instance, such a paper might involve re-working of the general emphasis, motivation, discussion of implications, and/or a re-analysis of existing data to highlight policy-relevant implications or prior work that have not been detailed elsewhere.

In our checklist for authors we ask for a brief statement that explicitly details how the present work differs from previously published work (or work under review elsewhere). When in doubt, we ask that authors include with their submission copies of related papers. Note that any text, data, or figures excerpted or paraphrased from other previously published material must clearly indicate the original source with quotation and citations as appropriate.

Authorship

Authorship implies substantial participation in research and/or composition of a manuscript. All authors must agree to the order of author listing and must have read and approved submission of the final manuscript. All authors are responsible for the accuracy and integrity of the work, and the senior author is required to have examined raw data from any studies on which the paper relies that the authors have collected.

Data Publication

BSP requires authors of accepted empirical papers to submit all relevant raw data (and, where relevant, algorithms or code for analyzing those data) and stimulus materials for publication on the journal web site so that other investigators or policymakers can verify and draw on the analysis contained in the work. In some cases, these data may be redacted slightly to protect subject anonymity and/or comply with legal restrictions. In cases where a proprietary data set is owned by a third party, a waiver to this requirement may be granted. Likewise, a waiver may be granted if a dataset is particularly complex, so that it would be impractical to post it in a sufficiently annotated form (e.g. as is sometimes the case for brain imaging data). Other waivers will be considered where appropriate. Inquiries can be directed to the BSP office.

Statement of Data Collection Procedures

BSP strongly encourages submission of empirical work that is based on multiple studies and/or a meta-analysis of several datasets. In order to protect against false positive results, we ask that authors of empirical work fully disclose relevant details concerning their data collection practices (if not in the main text then in the supplemental online materials). In particular, we ask that authors report how they determined their sample size, all data exclusions (if any), all manipulations, and all measures in the studies presented. (A template for these disclosures is included in our checklist for authors, though in some cases may be most appropriate for presentation online as Supplemental Material; for more information, see Simmons, Nelson, & Simonsohn, 2011, *Psychological Science, 22,* 1359-1366).

Copyright and License

Open Access

BSP posts each accepted article on our website in an open access format at least until that article has been bundled into an issue. At that point, access is granted to journal subscribers and members of the Behavioral Science & Policy Association. Questions regarding institutional constraints on open access should be directed to the editorial office.

Supplemental Material

While the basic elements of study design and analysis should be described in the main text, authors are invited to submit Supplemental Material for online publication that helps elaborate on details of research methodology and analysis of their data, as well as links to related material available online elsewhere. Supplemental material should be included to the extent that it helps readers evaluate the credibility of the contribution, elaborate on the findings presented in the paper, or provide useful guidance to policy makers wishing to act on the policy recommendations advanced in the paper. This material should be presented in as concise a manner as possible.

Embargo

Authors are free to present their work at invited colloquia and scientific meetings, but should not seek media attention for their work in advance of publication, unless the reporters in question agree to comply with BSP's press embargo. Once accepted, the paper will be considered a privileged document and only be released to the press and public when published online. BSP will strive to release work as quickly as possible, and we do not anticipate that this will create undue delays.

Conflict of Interest

Authors must disclose any financial, professional, and personal relationships that might be construed as possible sources of bias.

Use of Human Subjects

All research using human subjects must have Institutional Review Board (IRB) approval, where appropriate.

sponsors

The Behavioral Science & Policy Association is grateful to the sponsors and partners who generously provide continuing support for our non-profit organization.

Alfred P. Sloan FOUNDATION

BROOKINGS

Duke SOCIAL SCIENCE RESEARCH INSTITUTE

MacArthur Foundation

IRiSS

Berkeley UNIVERSITY OF CALIFORNIA

UCLA Anderson SCHOOL of MANAGEMENT

USC University of Southern California

THE WILLIAM AND FLORA HEWLETT FOUNDATION

To become a Behavioral Science & Policy Association sponsor, please contact BSPA at bsp@behavioralpolicy.org or 919-681-5932.

behavioral
science & policy

A publication of the Behavioral Science & Policy Association

join bspa

Behavioral Science & Policy Association's global community of scholars, practitioners, policy makers, and students is dedicated to fostering collaboration in the application of insights from rigorous behavioral science research in ways that serve the public interest. Visit http://behavioralpolicy.org/membership/ to become a BSPA member.

call for submissions

An international, peer-reviewed journal, *Behavioral Science & Policy* features short, accessible articles describing actionable policy applications of behavioral science research. As part of our dual-review process, leading disciplinary scholars assess articles for scientific rigor; at the same time, experts in relevant policy areas evaluate manuscripts for feasibility of implementation. Authors whose articles pass this dual-review work with editors trained in BSP's style to ensure their accessibility to scientists, policy makers, and lay readers. To submit your manuscript to *Behavioral Science & Policy,* visit http://behavioralpolicy.org/journal.

Behavioral Science & Policy Association
P.O. Box 51336
Durham, NC 27717-1336

ISBN 978-0-8157-2508-4
52000>
9 780815 725084