

Original Scholarship

Dark Nudges and Sludge in Big Alcohol: Behavioral Economics, Cognitive Biases, and Alcohol Industry Corporate Social Responsibility

MARK PETTICREW,^{*} NASON MAANI,^{*,†}
LUISA PETTIGREW,^{*} HARRY RUTTER,[‡]
and MAY CI VAN SCHALKWYK^{*}

^{*}*London School of Hygiene and Tropical Medicine;* [†]*Boston University School of Public Health;* [‡]*University of Bath*

Policy Points:

- Nudges steer people toward certain options but also allow them to go their own way. “Dark nudges” aim to change consumer behavior against their best interests. “Sludge” uses cognitive biases to make behavior change more difficult.
- We have identified dark nudges and sludge in alcohol industry corporate social responsibility (CSR) materials. These undermine the information on alcohol harms that they disseminate, and may normalize or encourage alcohol consumption.
- Policymakers and practitioners should be aware of how dark nudges and sludge are used by the alcohol industry to promote misinformation about alcohol harms to the public.

Context: “Nudges” and other behavioral economic approaches exploit common cognitive biases (systematic errors in thought processes) in order to influence behavior and decision-making. Nudges that encourage the consumption of harmful products (for example, by exploiting gamblers’ cognitive biases) have been termed “dark nudges.” The term “sludge” has also been used to describe strategies that utilize cognitive biases to make behavior change harder. This

The Milbank Quarterly, Vol. 00, No. 0, 2020 (pp. 1-39)

© 2020 The Authors. *The Milbank Quarterly* published by Wiley Periodicals LLC on behalf of The Millbank Memorial Fund

This is an open access article under the terms of the Creative Commons Attribution-Non-Commercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

study aimed to identify whether dark nudges and sludge are used by alcohol industry (AI)–funded corporate social responsibility (CSR) organizations, and, if so, to determine how they align with existing nudge conceptual frameworks. This information would aid their identification and mitigation by policymakers, researchers, and civil society.

Methods: We systematically searched websites and materials of AI CSR organizations (e.g., IARD, Drinkaware, Drinkwise, Éduc'alcool); examples were coded by independent raters and categorized for further analysis.

Findings: Dark nudges appear to be used in AI communications about “responsible drinking.” The approaches include social norming (telling consumers that “most people” are drinking) and priming drinkers by offering verbal and pictorial cues to drink, while simultaneously appearing to warn about alcohol harms. Sludge, such as the use of particular fonts, colors, and design layouts, appears to use cognitive biases to make health-related information about the harms of alcohol difficult to access, and enhances exposure to misinformation. Nudge-type mechanisms also underlie AI mixed messages, in particular alternative causation arguments, which propose nonalcohol causes of alcohol harms.

Conclusions: Alcohol industry CSR bodies use dark nudges and sludge, which utilize consumers’ cognitive biases to promote mixed messages about alcohol harms and to undermine scientific evidence. Policymakers, practitioners, and the public need to be aware of how such techniques are used to nudge consumers toward industry misinformation. The revised typology presented in this article may help with the identification and further analysis of dark nudges and sludge.

Keywords: behavioral economics, nudge, sludge, commercial determinants of health, public health, alcohol.

ALCOHOL CONSUMPTION IS A MAJOR CONTRIBUTOR TO GLOBAL death and disability, accounting for nearly 10% of global deaths among populations aged 15–49 years¹ and contributing to both communicable and noncommunicable diseases.^{1,2} For most alcohol-related diseases and injuries, there is a dose-response relationship between the volume of alcohol consumed and the risk of a given harm.³ Among those aged 50 years and older, cancers account for a large proportion of total alcohol-attributable deaths; 27.1% of total alcohol-attributable female deaths and 18.9% of male deaths.¹ The link with cancer in particular has been an extensive target of AI misinformation.^{4–6}

The evidence on alcohol harms has led to calls for strong alcohol control policies to protect public health. Among the cost-effective approaches is regulation of advertising and marketing.³ Alcohol marketing

is an important target because of its influence on the amount, frequency, timing, and contexts of alcohol consumption, and because it takes place through an increasingly wide variety of venues and formats, including the internet and social media, as well as traditional media.^{2,7} It has also long been known that marketing does not need to rely on the conscious awareness of the consumer to be effective, and often exploits cognitive biases (systematic errors in our thought processes) and information deficits, drawing on nudge-type approaches to changing behavior, and behavioral economics theory more generally.⁸⁻¹⁰

Thaler and Sunstein describe nudges as liberty-preserving approaches that steer people toward certain options but also allow them to go their own way:

A nudge, as we will use the term, is any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives.⁸

Thaler, Sunstein, and other behavioral economists have also discussed how companies exploit our cognitive biases and information deficits.¹¹ Such approaches are well documented in the marketing literature, which offers guidance, for example, on "how to build habit-forming products."¹² Cialdini's seminal books on how to influence purchasing in both physical and online environments draw strongly on behavioral economics.^{9,10}

The rest of this article describes how nudge-type approaches are used by alcohol industry (AI)-funded corporate social responsibility (CSR) organizations in the form of dark nudges and sludge. It shows how organizations use *dark nudges* to exploit well-known cognitive biases to promote misinformation and to subvert accurate information. It incorporates Sunstein's more recent discussion of *sludge*, a type of friction that makes choices and behaviors more difficult. Note that we use the term *misinformation* throughout rather than *disinformation*. *Disinformation* is frequently defined as the deliberate creation and sharing of false and/or manipulated information with the intention to mislead, for example for the purposes of political, personal, or financial gain.¹³ *Misinformation* refers to the inadvertent sharing of false information.^{13,14} We use *misinformation* to reflect that further evidence on the provenance and purpose of such misinformation, and the explicit intentions behind it, is needed to discern the difference between misinformation and disinformation

in AI and AI CSR materials. A summary of nudge approaches and an explanation of dark nudges is presented first, before an analysis of their use in AI CSR materials.

Nudge and Choice Architecture

Thaler and Sunstein point to ten broad types of nudges (Box 1).⁸

Box 1. Ten Important Types of Nudges

1. Default rules (e.g., automatic enrollment in education, savings, or other programs)
2. Simplification (e.g., removal of unnecessary complexity to promote take-up of social programs)
3. Use of social norms that emphasize what most people do (e.g., “nine out of ten hotel guests reuse their towels”)
4. Increases in ease or convenience (e.g., placing low-cost options or healthy foods in the easiest-to-reach locations)
5. Disclosure (e.g., of economic or environmental costs associated with energy use)
6. Warnings, graphic or otherwise (e.g., large fonts, bold letters, and bright colors to trigger people’s attention)
7. Precommitment strategies in which people commit to a certain course of action (e.g., to stop drinking or smoking)
8. Reminders (e.g., emails or text messages)
9. Eliciting implementation intentions (e.g., “do you plan to vote?”)
10. Informing people of the nature and consequences of their past actions (e.g., when utility companies are forced to let you know about trends in your consumption, as a nudge for you to amend your behavior)

Nudges typically exploit a wide range of cognitive biases in human information processing. These biases arise because human judgments are often fast and automatic. In this context, Kahneman and others have referred to two mental systems that govern human information processing: System 1, which operates quickly and automatically, with little or

no effort and no sense of voluntary control,¹⁵ and System 2, which allocates attention to effortful mental activities that demand it, such as performing complex computations and checking the validity of complex logical arguments. System 2 is associated with the subjective experience of agency, choice, and concentration.¹⁵ In some circumstances, particularly when people are cognitively overloaded, System 1 is in operation and rapid decisions are based on unconscious rules of thumb (heuristics). While often necessary for day-to-day functioning, these heuristics are often susceptible to cognitive biases and errors. Proponents of nudging argue that a nudge can be used to work *with* individuals' cognitive biases to achieve the desired behavior.⁸

Nudges typically take the form of changes to the *choice architecture*, which involves designing the context in which people make choices, to influence decision-making and make certain choices easier.^{8,16} Examples include changes to the physical environment, such as putting fruit at eye level in a cafeteria to prompt purchase and presenting information in different ways to influence decisions.¹⁶⁻¹⁸

Marteau and colleagues have noted that nudging builds on a long tradition of psychological and sociological theory explaining how environments shape and constrain human behavior. It draws on the more novel approaches of behavioral economics and social psychology and in doing so challenges the rational behavior model of classical economics.¹⁹

Hollands and colleagues have developed a typology of nudge-type interventions, called the Typology of Interventions in Proximal Physical Micro-Environments (TIPPME), that involve changing the environment in which they are available (*placement*), within settings such as shops, restaurants, bars, and workplaces, and the characteristics of products themselves (*properties*).¹⁶ Examples from this typology include providing additional healthier options to select from, placing less healthy options further away from potential consumers, and/or altering the portion size of food, alcohol, and tobacco products.¹⁶ Their typology identifies six different choice architecture intervention types, namely availability, position, functionality, presentation, size, and information (Table 1).

Dark Nudges and Sludge

Much of the academic discussion of nudges, particularly in the field of public policy, focuses on how they may be used to benefit individuals,

Table 1. Summary of TIPPME Typology of Choice Architecture Interventions^a

Class	Intervention Type (TIPPME Description)
Placement	<p><i>Availability</i> (Add or remove some or all products to increase, decrease, or alter their range, variety, or number)</p> <p><i>Position</i> (Alter the position, proximity, or accessibility of products)</p>
Properties	<p><i>Functionality</i> (Alter functionality or design of products to change how they work, or guide or constrain how people use or physically interact with them)</p> <p><i>Presentation</i> (Alter visual, tactile, auditory, or olfactory properties of products)</p> <p><i>Size</i> (Alter size or shape of products)</p> <p><i>Information</i> (Add, remove, or change words, symbols, numbers, or pictures that convey information about the product or its use)</p>

^aAdapted from Hollands and colleagues.¹⁶

such as public health policy to promote healthier lifestyles. However, nudge and behavioral economics more broadly also provide a valuable framework for understanding and analyzing how activities of harmful industries can promote actions that harm health. Newall provides an example from gambling research, where he has identified a range of dark nudges.²⁰ He describes how the industry exploits gamblers' cognitive biases, for example by framing losses as wins, or by removing friction from the gambling experience through touchscreen buttons that minimize the physical effort of long gambling sessions. Similarly, Schull's ethnography *Addiction by Design* details how characteristics of casinos and gambling machines are explicitly designed to keep people gambling, from the physical architecture (e.g., designing passageways to guide people toward machines) to sensory characteristics of the gambling environments (temperature, light, sound, and color) and gambling machines (slant of the screen; use of symbols, characters, and bonus wins).²¹

Sludge also utilizes cognitive biases, but rather than encourage behavior change it aims to restrict it through the use of cognitive biases that

favor the status quo and default options. Sludge has been defined by Sunstein as “excessive or unjustified frictions, such as paperwork burdens, that cost time or money; that may make life difficult to navigate; that may be frustrating, stigmatizing, or humiliating; and that might end up depriving people of access to important goods, opportunities, and services.”²² Friction can make choices and behavior change more difficult. The UK Government’s Behavioural Intervention Team (jointly owned by the UK Cabinet Office and the innovation charity Nesta and its employees) notes:

Requiring even small amounts of effort (‘friction costs’) can make it much less likely that a behavior will happen. For example, making it just slightly more difficult to obtain large amounts of over-the-counter drugs has been shown to greatly reduce overdoses.²³

Sludge has been observed in relation to the gambling industry’s voluntary self-exclusion schemes, which allow people to exclude themselves from gambling outlets and/or online betting. In practice, the schemes involve complex steps, including in some cases visiting betting shops to fill in the forms.²⁴

In the rest of this article, we use the term *dark nudges* to refer to the range of nudges available to industry, including nudges toward activities and information favorable to industry. We interpret sludge to be a type of supporting and synergistic architecture to dark nudges, because of the way sludge can nudge consumers away from (or hinder access to) behavior or information beneficial to the consumer but unfavorable to industry. We recognize that the terms may be used inconsistently in the academic and mainstream literature.

There has been little systematic analysis of the use of dark nudges or sludge. This is an important area for further exploration because dark nudges and sludge, in particular when used under the guise of CSR, can undermine public health efforts, waste public health resources, maintain environments favorable to industry interests, and result in harm to the public if used to disseminate misinformation and stall progress on addressing noncommunicable diseases. They may also increase inequalities if they have a greater impact on the most vulnerable in society.

Alcohol consumption is a particularly suitable area in which to explore harmful use of cognitive biases for two reasons. First, alcohol is a product that disproportionately harms the most vulnerable, and the AI is financially reliant on harmful use of its products.^{25,26} Second, there is

growing evidence that AI CSR organizations act to misrepresent the evidence on alcohol harms.²⁷⁻³⁰ It has recently been shown that such bodies misrepresent the scientific evidence on alcohol harms in pregnancy.³¹ To explore this further we conducted an analysis of the health-related information disseminated by AI-funded CSR organizations, to assess whether and how dark nudges and sludge-type approaches are used.

In the context of a well-developed body of literature documenting the role of CSR bodies in supporting industry as opposed to public interests, we hypothesized that AI and other industry CSR bodies adopt dark nudges and use sludge that serve the interests of their industry funders. Based on previous literature,²⁷⁻³⁰ we hypothesized that these would take multiple forms, draw upon multiple mechanisms, and serve to create an illusion of providing health information while simultaneously distorting or diluting the scientific evidence, nudging toward misinformation, toward more favorable attitudes to industry, and toward the consumption of their products.

Our specific objectives were to (1) determine whether dark nudges and sludge are evident in AI CSR activities and in their public-facing health information, and, if so, (2) locate these within existing behavioral economics frameworks to help understand the mechanisms through which they may exploit cognitive biases, and (3) propose an initial framework for classifying dark nudges and sludge in CSR materials.

The main source of data for these analyses consisted of the websites and social media accounts of CSR organizations, through which they frequently disseminate health-related information. This is an appropriate data source for analyzing AI nudges because websites (as noted by the UK government's Behavioural Insights Team) frequently exploit users' cognitive biases by using friction and by manipulating users' information deficits.¹¹

Methods

Using Hollands and colleagues' TIPPME (see Table 1) and Thaler and Sunstein's ten types of nudges (see Box 1) as a guide, we developed a comprehensive inventory of dark nudges and sludge through the exploration of recent systematic reviews and primary studies analyzing the activities of AI CSR organizations.^{4,28,29,31-35} We identified dark

nudges and sludge independently through a process of interrogating the literature using the framework, and reaching consensus among researchers through open discussion. We adopted an iterative approach, in which, guided by the evidence from behavioral economics, we identified mechanisms used by AI CSR to influence behavior through cognitive biases. We then used the framework and inventory to assess the dedicated websites and social media platforms of AI CSR bodies to (1) document contemporary use of dark nudges and sludge and (2) validate the use of the framework and inventory for this purpose.

Data Sources and Approach to the Analysis

AI-funded CSR activities include the dissemination of health information through industry-funded organizations. Examples of organizations include IARD (the International Alliance for Responsible Drinking, formerly ICAP), an alcohol producer's "responsible drinking" body. Such organizations are active in many countries, including Drinkaware (in the UK and Ireland), Drinkwise in Australia (funded by a wide range of Australian and international alcohol producers), Éduc'alcool in Quebec, and many others.⁵ These organizations have been shown to disseminate misinformation via their public-facing websites about the health risks of alcohol consumption.^{4,5,31,34,35} These organizations have the stated intention of informing consumers about health and encouraging "responsible drinking."³⁰

We initially searched three sources: systematic reviews and primary studies, Twitter content (January 2016 to February 2020), and website material of CSR organizations. However, there were no relevant examples in the systematic reviews because they focused on broad political and CSR strategies, rather than specific mechanisms. Therefore, our main sources were Twitter content (January 2016 to February 2020) and website material (September 2016 to February 2020). The 23 CSR organizations included were identified from two previous analyses.^{5,31}

Next, three authors (MP, NM, MvS) reanalyzed data from previously published studies and then discussed the results with the wider team to reach consensus on whether or not the examples were relevant to the framework and which if any of the nudge categories were relevant. Illustrative examples were rejected if not agreed on by all authors. All examples and interpretation were discussed with the wider team,

including authors with public health and clinical expertise who had not been involved in previous analyses of AI activities (HR, LP).

We include illustrative examples of how the industry in question uses a range of dark nudges and sludge that exploit cognitive biases to influence consumers' understanding and use of health information. Finally, we use these examples alongside the TIPPME typology to propose a revised taxonomy of AI CSR dark nudges, which we suggest may also be applied more broadly to other industries and contexts.

Findings

Searches of the AI CSR websites and examination of previous studies identified a range of dark nudges and sludge utilizing and building on the TIPPME framework. These include (1) alterations to the **placement**, including *availability* and *positioning* of information, to make it more or less accessible, and the selective *omission* of relevant information; (2) altering the **properties**, including the *functionality* of CSR websites, how information is *presented* on them, the *size* of products, and what *information* is conveyed; and (3) **misinformation strategies**, such as *priming* and *framing* effects, including the use of images and wording to prime drinking, *positioning and sequencing*, *stimulus incompatibility*, and *dilution effects*, which create confusion about the message, alongside *social norming* (or "social proofing"). Examples of these are described below. A full list, with a modified version of the TIPPME framework, appears in Table 2 and is summarized in Figure 1.

Placement

Availability and Positioning of Information. The ordering and positioning of information can strongly influence judgments: Kahneman describes how the weight of the first impression is such that subsequent information is mostly wasted.¹⁵ Positioning can also effectively mean that information is overlooked. This can be seen in AI-funded information about pregnancy harms. In one example on the website of Drinkaware UK (an industry-funded charity), the sections on pregnancy harms appear on the webpage titled "Health effects of alcohol." The page has (at time of writing) 45 such sections, of which the sections on pregnancy, breastfeeding, fetal alcohol syndrome, and fertility are

Table 2. Dark Nudges and Sludge in Alcohol Industry–Funded CSR^a

Class	Intervention Type (Based on TIPPME Description Where Applicable)	AI CSR Example
Placement	<p>Availability Add, remove, or physically obscure relevant information to increase, decrease, or alter its range, variety, or amount</p> <p>Position Alter the position, proximity, or accessibility of health-related information</p> <p>Omission Omission of relevant information, leading the reader to assume that they have been given complete information (WYSIATI)</p>	<p>Positioning of information on risk of breast cancer and pregnancy in Drinkaware materials and website; for example, where information about significant health harms is placed well below trivia. (The placement of the information may also mean that it is undermined by dilution and framing; see below.)</p> <p>Placing warning labels on the back of products, at the bottom of labels, or on disposable packaging;³⁶ placing website information “below the fold”; using difficult-to-read color/text/font combinations; placing pregnancy warnings below long sections of nonhealth trivia.³¹</p> <p>Avoiding mention of cancer on AI “health” websites,⁵ noninclusion of women in the Drinkaware Ireland infographic, meaning that breast cancer is not included.³⁷</p>

Continued

Table 2. *Continued*

Intervention Type (Based on TIPPME Description Where Applicable)	AI CSR Example
<p>Properties</p> <p>Alter functionality or design of information materials and websites/webpages to change how they work or to guide or constrain how people use or physically interact with them: “Requiring even small amounts of effort (‘friction costs’) can make it much less likely that a behavior will happen.”^{23 (p4)}</p>	<p>Making websites or webpages less functional to make them more difficult to interact with; making it more difficult to access health information by requiring registration processes/supply of personal information; requiring users to scroll through many pages to get to pregnancy and cancer information, for example, on the DrinkIQ website users have to scroll down four pages to read the health information; placing information on alcohol’s effects on the body “below the fold,” requiring users to page down and click on a link, taking them to a page where they have to page down again to get to the health harms, including breast cancer; placing statements about the benefits of drinking above information about harms: “The social benefits are immediate – especially when keeping the amount you consume within the recommended limit. It’s a great social lubricant, and when paired with food it can enhance celebratory occasions.”³⁸ (Also see framing, below.)</p> <p>Some AI-funded organizations require users to order information and to complete online forms. Other sites use multiple split moving screens; moving information makes navigating the health/consumption sections difficult. (See, for example, Bacardi-Martini: https://www.slowdrinking.com/us/en/#first.)</p> <p>Information overload—providing so much factual information that it may be difficult to process or difficult to find the most relevant information (e.g., inclusion of large amount of minor and trivial information on industry websites, swamping or displacing information on major health harms).</p>

Continued

Table 2. *Continued*

Class	Intervention Type (Based on TIPPME Description Where Applicable)	AI CSR Example
	<p>Presentation</p> <p>Alter visual properties of information (e.g., on webpages or documents)</p>	<p>Placing health warning labels in very small font and in difficult-to-read color combinations;³⁶ using difficult-to-read color/text/font combinations in printed or web material (e.g., black on purple text in a Drinkaware breast cancer leaflet).</p>
	<p>Size</p> <p>Alter size or shape of products</p>	<p>Single drink cans that exceed daily alcohol consumption guidelines; use of larger wine glasses as default serving.</p>
	<p>Information</p> <p>Add, remove, or change words, symbols, numbers, or pictures that convey information about the product, its use, or risks</p>	<p>Noninclusion of chief medical officers' health information on alcohol products.</p>
	<p>Priming</p> <p>(Not in TIPPME) Use of words, symbols, images, or numbers to prime drinking</p>	<p>Use of written cues to drink: "Didn't drink at all during January and looking forward to drinks tonight?"; "It's tradition to ring in the New Year with a glass of bubbly, but what's in champagne?"; "Blue Monday research shows that many drink alcohol to forget their problems, cheer themselves up when in a bad mood and because it helps when they feel depressed or nervous" (Drinkaware³⁹⁻⁴¹).</p>
	<p>Misinformation strategies:</p> <p>undermining, subverting, or contradicting health messages through ordering and sequencing of the messages, by placing alongside counter-messages, or by using other wording, arguments, and/or images to undermine the health message or make it more difficult to interpret; may also involve denialism and use of pseudoscientific language/arguments.</p>	

Continued

Table 2. *Continued*

Intervention Type (Based on TIPPME Description Where Applicable)	AI CSR Example
Class	<p>Priming drinkers by offering verbal and pictorial cues to drink, sometimes while simultaneously appearing to warn about harms, as in Drinkaware Twitter messaging.³⁴</p> <p>Alcohol-free drinks may also involve priming: the branding, design, bottles, and location in the supermarket send stimuli that remind the drinker of drinking alcohol.</p>
<p>Framing (Not in TIPPME) "Large changes in preferences ...are sometimes caused by inconsequential variations in ... wording."¹⁵</p>	<p>AI-funded websites frequently frame information on alcohol in terms of the benefits, even when discussing harms (e.g., Éduc alcool frames health information in terms of benefits of alcohol consumption on its webpage on "the 8 benefits of moderate drinking," mentioning no specific health risks.⁴²</p> <p>AI materials often frame health information in terms of uncertainty and complexity, and offer alternative explanations for alcohol harms: "There are many other factors that increase the risk of developing breast cancer, some of which we can't control like: Age; you're more likely to develop it as you get older; A family history of breast cancer; Being tall; A previous benign breast lump. However, in addition to alcohol, other lifestyle factors such as being overweight and smoking are thought to increase your risk of developing breast cancer" (Drinkaware).²</p>
<p>Positioning and sequencing (Not in TIPPME)</p>	<p>Information on harms on AI websites is often positioned below information on benefits; for example, Diageo's "DrinkIQ" website places information on breast cancer below statements emphasizing social and other benefits, such as "It's a great social lubricant, and when paired with food it can enhance celebratory occasions."³⁸</p>

Continued

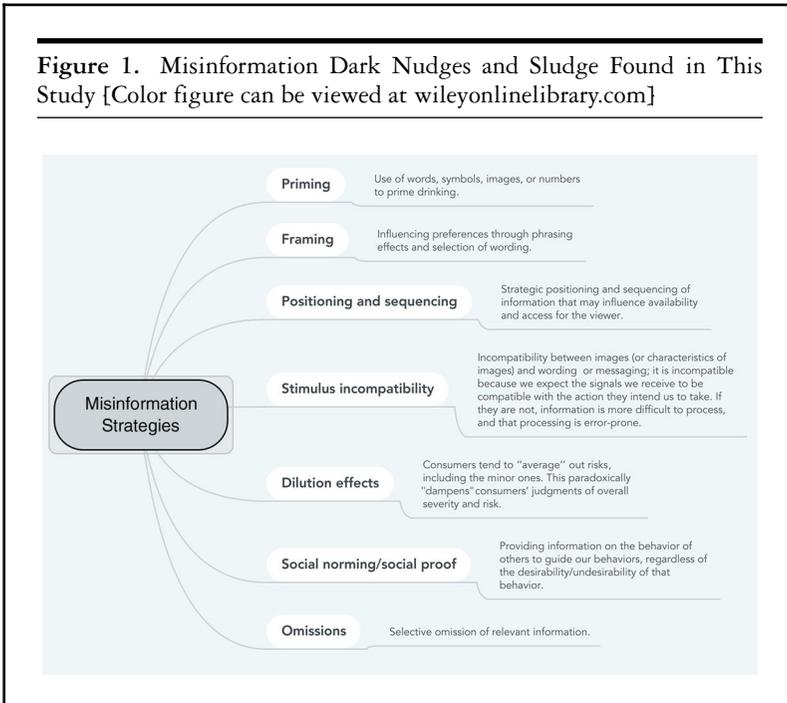
Table 2. Continued

Intervention Type (Based on TIPPME Description Where Applicable)	AI CSR Example
<p>Stimulus incompatibility Incompatibility between images (or characteristics of images) and wording or messaging; we expect the signals we receive to be compatible with the action they intend us to take. If they are not, information is more difficult to process, and that processing is error-prone.</p>	<p>CSR messages that appear to warn readers about consumption, placed alongside trivia about alcohol, may involve stimulus incompatibility (in addition to distraction, dilution, framing, and other biases). The Drinkaware “Drink Free” days campaign messaging shows a woman having fun in the swimming pool, alongside upbeat music. The spelling and spacing of the phrase “Enjoy more Drink Free” (as opposed to “Drink-Free”) may, however, convey a pro-drinking message. The clear separation of the words “Enjoy more” from “Drink Free,” in addition to the different coloring of the word “Days,” may prompt the reader to break these phrases into two separate messages, and thus interpret the message as “Enjoy more: Drink free.” Thus, “Drink Free” has a dual meaning, including “Drink freely.”⁴³</p>
<p>Dilution effects Consumers tend to “average” out risks, including the minor ones. This paradoxically “dampens” consumers’ judgments of overall severity and risk.</p>	<p>AI websites often including a wide range of minor and trivial effects of alcohol alongside (and sometimes displacing) severe harms. This may result in dilution effects (as well as framing and positioning effects). For example, the Drinkaware website includes information on “Why does alcohol make you pee more?” alongside (and placed well before) information on fetal alcohol syndrome.</p>
<p>Social norming/social proof (Not in TIPPME) Providing information on the behavior of others to guide our behaviors, regardless of the desirability/undesirability of that behavior</p>	<p>“Today’s the day when people most commonly give up on their Dry January efforts” and “On holiday, everybody loves to drink, you’re in a good vibe, the alcohol’s cheaper, that’s always going to happen” (Drinkaware^{44,45}).</p>

Abbreviations: AI, alcohol industry; CSR, corporate social responsibility; TIPPME, Typology of Interventions in Proximal Physical Micro-Environments, WYSIATI, what you see is all there is.

^aNote that some well-known biases and heuristics are not included; for example, the base-rate fallacy and loss aversion.

Figure 1. Misinformation Dark Nudges and Sludge Found in This Study [Color figure can be viewed at wileyonlinelibrary.com]



the last four. Thus, the user needs to scroll through approximately nine pages to access it—or many more pages on handheld devices.⁴⁶ This positioning means that the information on pregnancy harms—one of the most important and well-established harms of alcohol consumption—is placed well below, for example, sections on “How does alcohol affect my beer belly?” and “Why does alcohol make you pee more?”⁵¹ On other AI websites, accessibility is reduced by not allowing health information to be viewed directly on screen; instead it has to be requested separately or ordered online.

Selective Omission of Information. Framing effects can also be achieved by the omission of information. As Entman notes, “Most frames are defined by what they omit as well as what they include, and the omissions of potential problem definitions, explanations, evaluations and recommendations may be as critical as the inclusions in guiding the audience.”⁴⁷ The tendency of readers to ignore absent evidence has been

named WYSIATI (“what you see is all there is”¹⁵) by Kahneman, who identifies it as a key source of bias.¹⁵

WYSIATI is a common source of bias on AI webpages, particularly in relation to cancer: Mention of specific cancers is often omitted, particularly breast cancer and colorectal cancer (which the industry has been found to deny is causally associated with alcohol consumption²⁶), while other chronic diseases such as liver disease are included.^{4,5} WYSIATI in industry-funded materials can also involve cherry-picking particular health outcomes (for example, by stating that alcohol consumption reduces the incidence of bladder cancer), while ignoring other negative outcomes (such as the increase in risk of other much more common cancers, in particular breast cancer and colorectal cancer).⁵

In one example, the Drinkaware Ireland infographic “The Body” omits women entirely, showing only a man.³⁷ In a later addition, while the other health harms (e.g., “liver disease,” “gastritis”) are shown close to the site of the harm (e.g., the words “liver disease” are placed by the torso), the word “cancers” appears by the right ankle. The result of this “men-only” infographic is that the most common alcohol-attributable cancer in women, breast cancer, is not clearly noted as a significant alcohol-related harm.³⁷ The accompanying text, “The effects shown here are the same for men and women,” is also misleading. This omission of breast cancer from a female infographic was also previously found on the industry-funded Drinkwise website in Australia.⁵ The omission or misrepresentation of breast and colorectal cancer is a common feature of AI materials more generally.⁵

WYSIATI can also be seen in documents from AI-funded organizations that discuss risk factors for underage drinking, and harmful alcohol consumption in general. These usually point to the causes of alcohol harms in young people as being variously “peer pressure,” “parents,” and “culture,” but generally do not mention alcohol advertising or AI activities. Examples of this omission can be seen in Drinkaware UK and Drinkaware Ireland materials on young people’s drinking.⁴⁸ Drinkaware Ireland’s list of reasons points to a range of individual-level factors and the influence of parents and peers, but does not mention the role of the AI.⁴⁹ Similarly, the Australian Drinkwise website, in its “Facts for parents” section, points to peers, parents, grandparents, and the community as the main influence on young people’s drinking, without mentioning the role of the AI. The evidence is clear that alcohol

advertising influences young people to start drinking and influences consumption in those who already drink.⁵⁰

Properties

Alterations to the Functionality of CSR Websites and to the Presentation of Information. Colors, fonts, and graphic design can be used to nudge people's choices in certain directions. Kahneman describes how fonts can influence judgments about the truth of statements and affect the ease of processing information.^{15(p70)} This has been known by graphic designers for decades and is a standard part of web-designers' skillset.⁵¹ One well-known example is the placement of information "above the fold" on websites—placing key information high on the first page of a website to relieve users of the effort of scrolling. A Google study has found that viewability rapidly drops off for information placed below the fold; information above the fold has >70% visibility, whereas information below it has around 44% visibility.⁵²⁻⁵⁴ Similarly, eye-tracking studies show that people scan for information on websites in an F-shaped pattern, in which they focus mainly on the headline and first paragraph.⁵⁵

In some Drinkaware materials, statements emphasizing uncertainty about whether alcohol causes cancer are placed first, while factual statements about the risk of specific cancers are placed further down the page. The latter section is printed in a less-legible black font on a dark purple background, compared with the much more legible section on uncertainties (black font on light purple).³⁶ The design appears to nudge the reader toward the uncertainties and away from information on the risk of specific cancers.

Graphic design also appears to be used to present mixed messages in alcohol labeling. In the UK, alcohol labeling is self-regulated by the industry itself. Its labeling guidance includes a voluntary warning about the risks of drinking while pregnant, either in the form of text or a circular warning sign that consists of a very small (mean diameter <6mm) logo showing the silhouette of a pregnant woman holding a glass of what is presumed to be alcohol with a line struck across it (Figure 2).⁵⁶ There is, however, a major inconsistency in the logo's color. Warning signs and stop signs in the real world are usually red. But AI pregnancy warnings are red in only a small proportion (10%) of cases⁵⁷ and can be found in many other colors, including green, gray, black, and white. This



represents a form of dark nudge because we expect the signals we receive to be compatible with the action they intend us to take. Thaler and Sunstein themselves use the example of stop signs. We naturally expect a stop sign to be red. We do not expect it to be green or a “go” sign to be red. Incompatible configurations like this (“stimulus incompatibility”) are difficult to process and lead to high error rates.⁸

Misinformation Strategies

Priming Effects. Kahneman describes priming effects, in which “evidence accumulates gradually and interpretation is shaped by the emotion attached to the first impression.”¹⁵ Priming works by offering simple and apparently irrelevant cues that “prime” people toward adopting certain attitudes, beliefs, and/or emotional states or behaviors. Examples of priming include using words or images to make other words, images, or concepts easier to recall. For example if you have recently seen the word EAT, you are more likely to complete the word fragment SO_P as SOUP than to complete it as SOAP.¹⁵ Environmental triggers can also prime behaviors—for example, larger plates prompt diners to consume more.^{57,58}

Priming can also use images: the MINDSPACE report by the UK Cabinet Office and the Institute for Government notes that “if a happy

face is subliminally presented to someone drinking it causes them to drink more than a frowning face.”^{59(p25)} Drinkaware frequently accompanies its Twitter messaging about alcohol harms with marketing-style images of people drinking and laughing.³⁴ There are many other examples of imagery-based framing on AI-funded websites, particularly the use of “happy drinkers.” Almost every image on the Drinkwise website at time of writing is of happy drinkers, drinking from bottles and glasses or toasting each other.⁶⁰

Framing Effects. Framing effects occur, according to Druckman, when changes in how an issue or an event is presented produce changes of opinion.⁶¹ Different ways of presenting the same information can thus evoke different emotions and result in different levels of credibility. Framing is a powerful tool: Kahneman notes that “large changes in preferences ... are sometimes caused by inconsequential variations in ... wording,”^{15(p88)} and Hollsworth and colleagues argue that “adopting different ‘frames’ can have powerful effects on how people perceive a problem and what they consider to be relevant facts (or ‘facts’ at all).”⁶²

Approaches to framing include frame incorporation, where one side incorporates a challenging element into their own frame by creating a watered-down version of it.⁶³ An example of this might be where AI materials acknowledge that alcohol is indeed a risk factor for breast cancer (the most prevalent cancer in women) but then weaken the statement by pointing to the relationship with less common forms of breast cancer:

Recent studies indicate that alcohol consumption may be more strongly linked to a certain less common form of breast cancer (lobular cancer), than it is to the most common type of breast cancer (ductal cancer).⁵ (SABMiller)

The epidemiological evidence regarding alcohol consumption with lobular and ductal cancers is in fact unclear, and in any case the information is of no direct relevance to individual drinking decisions; drinkers are obviously unable to choose any particular form of breast cancer risk.⁵

Positioning, Sequencing, and Stimulus Incompatibility. Ordering and framing effects can frequently be seen in AI materials, such as when (sometimes accurate) statements about specific alcohol harms are placed below misleading general statements. For example, in the Drinkaware leaflet titled “Alcohol and cancer,” the very first sentence emphasizes uncertainty: “There is no scientific consensus on why some people develop cancer and some don’t.”⁶⁴ The narrative continues by emphasizing

the importance of risk factors other than alcohol, including genes (twice), as well as smoking and exercise. As noted earlier, factual statements about the risk of specific cancers (e.g., breast cancer) appear further down the page, in a less-legible section. This adds considerable “friction” to the priming effect, which prioritizes the uncertainties by placing them first. Similar “priming with uncertainty” may be seen in Drinkaware information materials on alcohol and pregnancy.³¹

Emphasizing uncertainty is also a very common feature of AI information on cancer,⁵ often regarding the specific biological mechanisms, which may undermine the presentation of accurate information. For example, information about alcohol and cancer risk disseminated by the Wine Information Council (an organization funded by the European Wine Sector) states:

All the studies show that the knowledge about the causes of breast cancer is still very incomplete and as scientists from the National Institute on Alcohol Abuse and Alcoholism in the USA, recently pointed out, some other (possible confounding) factors have not been considered in the research relating the consumption of alcoholic beverages to breast cancer.⁵

Uncertainty is also highlighted in this example:

The mechanism by which alcohol consumption may cause breast cancer is not fully known. ... The relationship ... is undergoing vigorous research. ... If and how these two factors may interact and affect risk is not completely known.⁵ (SABMiller)

There are many other AI examples of the use of uncertainty as a frame.⁴ It should be noted that framing the evidence in terms of uncertainty like this is a well-documented characteristic of AI misinformation and misinformation related to other harmful industries, including climate change misinformation.⁶⁵⁻⁶⁷

Dilution Effects. As noted earlier, when faced with a vast array of information, people fall back on rules of thumb. Some AI-funded websites include large amounts of information, much of it of little or no relevance to specific health harms, which may be designed to cause information overload and information dilution. The dilution effect may be enhanced further by the positioning of information alongside apparently irrelevant trivia. For example, in Pernod Ricard’s “Wise Drinking” app, a statement that fetal alcohol syndrome is a “pattern of physical and mental defects that can develop in a fetus in association with high levels of

alcohol consumption during pregnancy” is found alongside trivia such as “the term Champagne can only be used for wines produced in the Champagne region.” Another example of dilution may be seen on the Drinkaware website, which positions the section entitled “Is alcohol increasing your risk of cancer?” beside the section on “Is alcohol affecting your looks?”⁶⁸

Dilution is a potent form of information bias. Sivanathan and Kakkar⁶⁹ have shown that when pharmaceutical commercials list the severe, but rare side effects of those drugs along with those that are most frequent (which include both serious and minor side effects), it dilutes consumers’ judgments of the overall severity of the side effects, compared with when only the serious side effects are listed. In effect, consumers tend to average out the risks including the minor ones. This paradoxically “dampens” consumers’ judgments of overall severity and risk and increases the marketability of these drugs. Dilution may also underpin the use on AI-funded websites of irrelevant risk factors, such as “being tall,” as on the Drinkaware UK website, and genetic factors, on Drinkaware UK and many other AI websites. These factors are irrelevant to individual consumers because they are unmodifiable, with no direct relevance to an individual’s alcohol consumption.

Social Norming and Social Proof. Some dark nudges and sludge rely on the power of social influences. Social norming, for example, involves promoting behavior by telling the reader or consumer what most other people do, or by correcting misapprehensions about what other people do:⁷⁰

In particular, advertisers are entirely aware of the power of social influences. Frequently they emphasize that “most people prefer” their own product, or that “growing numbers of people” are switching from another brand.⁸

Robert Cialdini, in his seminal book on marketing, *Influence*, describes this as social proof.¹⁰ Social proof can be seen in AI-funded campaigns. The Drinkaware advice tweeted in the summer of 2018 “When the sun’s out, it can be harder to stick to your #Drinkfreedays” arguably shows the same type of social norming.⁷¹ Earlier in the year, during the “Dry January” campaign, which is run by the independent charity Alcohol Change, Drinkaware again potentially nudged drinkers in this way, advising them on January 20: “Today’s the day when people most commonly give up on their Dry January efforts.”⁷²

New Year's Eve prompted another social norms nudge with the following tweet in both 2017 and 2018: "It's tradition to ring in the New Year with a glass of bubbly, but what's in champagne?"⁷³ On September 10, 2018, Drinkaware advised followers, "Two thirds of regular drinkers say cutting down is harder than improving diet and exercise - Just some of the findings from our new campaign."⁷⁴

A 2018 tweet also appears to employ social proof to imply that drinking is an effective form of mental health support: "The New Drinkaware & @YouGov research reveals that almost 3/5 (58%) of all people (aged 18-75) who drink alcohol are doing so because it helps them to cope with the pressures of day to day life."⁷⁵ The same social norming message appears to underline the organization's Twitter message on Christmas Eve in 2019 ("Linda" is an English soap opera character): "in real life, millions of people like Linda drink to cope and it can get out of control." This contains both the message that drinking to cope is normal and extremely common, and also that it is only when it gets "out of control" that it is a problem. "Out of control" drinking, as referred to here, is a common AI framing, which stresses behavioral aspects of alcohol harms, rather than chronic harms.⁷⁶

In other contexts, it is known that spreading information about non-compliance with health-supporting behaviors, as in the above example, can undermine health messages. For instance, reporting that some parents refuse pediatric vaccines can lead parents to construe this as being a social norm and exacerbate the problem of vaccine refusal.⁷⁷

These are some of the main examples of dark nudges and sludge that we identified in AI CSR materials. We note that an entire category of AI CSR nudges focuses on misinformation. Definitions of misinformation vary, but it is clear that it is much wider than the simple presentation of false information, and it includes the mixing of facts with inaccurate information, the undermining of facts through their framing (e.g., through misleading headlines), and the use of partial or selective truths.⁷⁸ We have therefore added this new category of misinformation to the TIPPME framework in Table 2, with illustrative examples, and we present it in Figure 1. It should be noted that some types of dark nudges fall within more than one intervention type and more than one class of nudge, suggesting that some nudges operate through a range of mechanisms.

Overall, our findings add to the literature on the contradictions that exist in the discourses of the AI,⁷⁹ whereby the industry claims that it

is supporting public health and “Nudging for Good,”^{80,81} while simultaneously, through its CSR bodies, using dark nudges that are aligned with commercial rather than public health goals. Its key approach involves presenting health information, but subverting it with a range of dark nudges and sludge. These dark nudges may also act synergistically with other tactics in forwarding the interests of the AI, as they form part of a much larger web of marketing and lobbying strategies. In this way, AI CSR programs may be “designed to fail” in reducing harmful alcohol consumption and may in fact increase it.

Discussion

It is already known that AI CSR organizations are significant sources of misinformation about health harms of alcohol consumption—misrepresenting the harms and overstating any benefits. The industry does this in order to appear to be acting responsibly and as “part of the solution.”³⁰ The use of dark nudges may help ensure that any health information provided by such organizations has little or no effect on consumers’ decisions; effective, unbiased provision of health information would pose too much of a risk to the alcohol market. Multiple dark nudges are therefore used: The positioning of information adds friction and reduces information accessibility, while readers are subject to priming and framing in which minor health effects are prioritized and chronic harms are omitted and reframed. Information overload and dilution with irrelevant trivia are also employed. Sludge also plays a role, for example, through requirements to take complicated sets of actions in order to obtain information, and through the use of layouts and fonts rarely optimized for the browser window.

Alcohol consumption is an area where consumers may be particularly easy to exploit with dark nudges, because in Kahneman’s phrasing, decisions about drinking are a “choice problem.”^{15(p272)} Consumers, when deciding whether and how much to drink, are required to make decisions about the balance of risks and benefits, yet have imperfect information and limited personal experience. Most of the harms, particularly the chronic harms—including cancer, cardiovascular disease, and liver disease—have long lag times and multifactorial etiologies, while any short-term benefits of drinking (e.g., social benefits) are much more salient. Drinkers are also subject to other biases well documented by

economists, including discounting the future and over-valuing short-term gains compared to long-term harms and losses.⁸² This is fruitful territory for a choice architect who wishes to prioritize industry interests over the consumer's best interests.

One key role for nudging in these examples is to foster uncertainty. Uncertainty is commonly exploited in a similar way by the tobacco, chemical, and other harmful industries in order to spread doubt about the harms of their products.⁶⁶ For example, the tobacco industry has used uncertainty to defend itself in litigation, arguing that it is impossible to prove that smoking is really the cause of lung cancer: "There are many other potential carcinogens to which we are all exposed. The modern environment is inherently dangerous (toxic waste, industrial pollutants) and no particular cause can be pinpointed."³¹ Alternate causation arguments were found to be very influential with jurors when the tobacco industry was defended in court; to quote tobacco industry lawyers Shook, Hardy, and Bacon, "A little alternative causation evidence goes a long way."⁸³ There are many similar statements on AI-funded websites, where the independent effects of alcohol consumption are obscured by the listing of a large range of other potential and often unmodifiable risk factors (such as the earlier example from Drinkaware of "being tall").^{4,5} The cherry-picking of positive outcomes and the avoidance of mentioning alcohol-related harms (such as breast cancer) also has resonances in tobacco industry CSR materials, which avoid mentioning death as an outcome of smoking.⁸⁴ Other dark nudges resemble well-known marketing tricks. For example, social norming has frequently been used in alcohol marketing to encourage consumption; one advertising campaign for lager noted that its "new appeal lay in drinkers' perceptions of it as a 'popular, fashionable lager.'"⁸⁵

There are several implications of this analysis. First, there is growing interest in misinformation campaigns and the role of the internet and social media in spreading them, and in how to recognize and prevent them.^{13,78} Our analysis suggests that further research on the use of nudge-based techniques to spread misinformation will be of value. Second, developing effective methods of prevention will be key. In the case of AI CSR and parallel efforts by other harmful commodity industries, this should involve raising awareness of dark nudges among policymakers, public health practitioners, clinicians, and the public. Awareness-raising activities should also target organizations that are considering partnering with AI organizations (e.g., local government).

However, education and awareness raising is unlikely to be enough, and public health policymakers and practitioners also need to consider whether there is a role for sanctions for making misleading and false health claims on alcohol and other harmful commodity industry websites. Social media companies are already acting against internet-based anti-vaccine misinformation, and it may be useful to consider how this approach could be extended to other sources of health misinformation.⁸⁶

Future Research

There is also a wider research agenda on dark nudges and sludge in public health. This could include further analysis of the use of nudges in the materials and activities of the alcohol, tobacco, food, beverage, and gambling industries. It might also extend to the analysis of how cognitive biases are used to create confusion and spread vaccine misinformation and disinformation, as well as their use as a vector for climate change denialism.^{87,88} This will also require detailed analysis of the “dark money” behind dark nudges and sludge.

Given these findings, and in light of Kahneman’s warning that “the power of format creates opportunities for manipulation,”^{15(p330)} future analyses should involve collaboration with web designers and graphic designers, as well as those who have worked in marketing, to help understand how graphic design elements are used in AI materials. Research to analyze the effects of such industry techniques on their intended audiences is also needed to understand how users respond to framing, as recommended by Entman.⁴⁶ This could include analysis of the synergistic impact of multiple nudges. To aid such future analyses it may also be useful to consider dark nudges and sludge as forming a type of complementary dark nudge architecture, while recognizing that there will frequently be overlap in their use:

1. Dark nudges nudge *toward* harmful products such as alcohol (e.g., Drinkaware having bottles and wine glasses in its logo and the imagery of whiskey on the “alcohol data and facts” page of its website).
2. Sludge, and its supporting structures, generate confusion and friction, making it harder to understand and appropriately respond to independent, accurate public health messages that aim to reduce harmful consumption.

Limitations and Strengths

There are limitations to our analysis. First, the examples we identified represent a small subset of a much larger set of nudges and decision-making biases. Other well-known biases, such as confirmation biases, where people interpret or remember information (including health information, such as in the case of vaping⁸⁹) that confirms prior beliefs, may also be used by AI CSRs. For reasons of space we have not included messenger/source biases: How we respond to information is heavily influenced by the messenger or source of that information, and how we feel about them.⁵⁸ For example, the public places less value on information derived from industry-funded clinical experts when they know about their funding.⁹⁰ In AI materials, one very frequently sees experts and other AI-funded organizations quoted as a source of clinical information, without acknowledgment that many are also paid industry advisors. This may relate to the concept of “trust simulation”: building the appearance of trustworthiness.¹¹ Further analysis of AI websites and social media is needed to determine the full extent of these and other biases.⁸⁷ Second, the evidence for some of the individual nudge approaches has also been challenged; for example, a failure to replicate some of the key studies of priming,⁹¹ though this does not invalidate the fact that the industry uses approaches based on priming, even if the evidence is flawed.

Finally, we cannot be certain that the AI examples that we cite are intentional. However, the pervasive use of such techniques, across multiple organizations and across different media, and given what is already known about AI strategies, suggests that the accidental use by AI bodies of dark nudges and sludge is unlikely. Drinkaware UK has stated that it does use behavioral economics and “nudging,” but to reduce harmful drinking.^{92,93} Other AI CSR bodies have also discussed nudging to change consumer behavior; examples include the Portman Group (UK-based AI “responsible drinking” organization) and Drinkaware Ireland.⁹⁴ Independent evaluations are now needed to assess whether and how such nudges have their intended impacts.

It is also important to note that all the examples used in the study were present and as described at the time of the analysis. However, the text and images on these websites change frequently, so specific examples may not always be locatable by readers.

In these concluding paragraphs it is worth directly addressing the possible response of AI CSR organizations to this qualitative analysis.

In particular, they may claim that our examples are not representative. However, the issue of representativeness is not directly relevant because we are not trying to estimate the prevalence of the problem (i.e., nudge-based misinformation). Instead our aim is to identify examples of industry misinformation and to explain the mechanisms behind them. We do not claim (and logically we do not need to show) that these examples are representative. Nor do we need to show or claim that they are used across *all* industry CSR bodies, nor across *all* the materials produced by an individual organization. In fact, it is highly unlikely that nudge-based misinformation could be identified in all industry CSR materials because the power of misinformation to mislead comes from the fact that it is mixed with factual information. This mixing of misinformation with fact is also used as a deliberate strategy to obscure its presence.⁶⁴ Another example that may illustrate our point is the release of millions of irrelevant and trivial documents by the tobacco industry following litigation, thereby making the identification of the egregious and damning documents extremely challenging. These latter documents were identified and analyzed in depth in numerous research projects, which have transformed how we think about and regulate the industry, but there were no claims that these documents are representative of the millions in the total sample. Such claims were not needed or relevant.

Nonetheless, industry-funded organizations may claim that our examples are cherry-picked. However, the identification of examples of AI misinformation is not cherry-picking, any more than is identifying specific examples of tobacco industry misinformation. Instead, these examples represent individual pieces of evidence of misleading, and therefore harmful, information. This is not an issue of representativeness, but an issue of safety/harm. In such cases, identifying even one example of a harm is enough to warrant concern and action to address it. For example, harmful products (e.g., certain formulations of drugs or malfunctioning devices) are frequently recalled without the need to determine whether they are representative of that entire product category.

A strength of our study is that the findings are internally and externally consistent. These are not simply coincidental examples of unintended misinformation, but instead represent a clear consistency of practice within and between AI-funded organizations. There are four main pillars underpinning this conclusion:

First, the findings show internal consistency: they represent evidence of repeated misinformation and similar approaches to misinformation from within the same organization over time and across different media (websites, social media). Consider for example, Drinkware, IARD, Drinkwise and others' repeated use of misleading "alternative causation" arguments across different topics (including cancer and pregnancy harms). The use of these types of arguments is well documented, as it is a very common misinformation technique across a wide range of harmful industries.⁶⁴⁻⁶⁷

Second, the findings show external consistency: There is consistent use of the same approach to misinformation across different AI-funded organizations. The use of alternative causation arguments is again an example of this: different AI-funded organizations that claim their independence use similar "alternative causation" arguments in relation to alcohol and breast cancer.⁴ Omission bias is another example: a wide range of organizations selectively omit breast cancer and colorectal cancer from their discussions of health. We have documented this particular omission bias in the materials of organizations such as Drinkaware, Drinkaware Ireland, Éduc'alcool, and others.⁴

Third, there is theoretical consistency: The approaches that we document closely match examples from the theoretical and empirical literature on behavioral economics. The examples of social proof we identified, for example, fall into this category.

Fourth, our interpretation is strengthened by the fact that our analyses discriminate between the materials and information used by industry-funded and independently funded organizations (such as independent health agencies). For example, independent health agencies are significantly less likely to use marketing-type images,³⁴ which, as we described earlier, are examples of priming.

Consideration now urgently needs to be given to how the dark nudges and sludge of harmful commodity industries' CSR strategies may be curtailed, including by using existing regulatory or legislative measures to protect the public. There may be existing models or amendments to existing legislation that could be considered; for example, in some countries the dissemination of misinformation about cancer treatments, and about COVID-19, is prohibited. Other transferable lessons may arise from the COVID-19 experience. Facebook, Google, Twitter, and YouTube now claim to be taking active steps to stop misinformation and harmful content from spreading, though the effectiveness of these

actions is unclear. However, in the case of AI misinformation or disinformation we also need to consider the role of clinicians and others involved in advising these organizations, and whether this is consistent with their professional codes of ethics.

Conclusion

In conclusion, reducing, removing, and mitigating the impact of dark nudges and sludge should be an important priority for public health policy. In answer to a question about alcohol labeling and pregnancy, the past UK chief medical officer replied: “We need some work about what the nudges are, and how we get change in the population.”^{95(p9)} An essential complementary course of action should be to identify and remove the dark nudges and sludge that already exist.³¹ Thaler and Sunstein note that “in many areas, ordinary consumers are novices, interacting in a world inhabited by experienced professionals trying to sell them things ...people’s choices are pervasively influenced by the design elements selected by choice architects.”⁸ Our analysis suggests that in interacting with AI CSR organizations’ websites and social media accounts, consumers are exposed to extensive misinformation from the AI’s choice architects, in ways that benefit the AI and increase the risk of harm to consumers.

References

1. GBD 2016 Alcohol Collaborators. Alcohol use and burden for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study. *Lancet*. 2018;392(10152):1015–1035.
2. Esser M, Jernigan D. Marketing in a global context: a public health perspective. *Ann Rev Public Health*. 2018;39:385–401.
3. Burton R, Henn C, Lavoie D, et al. A rapid evidence review of the effectiveness and cost-effectiveness of alcohol control policies: an English perspective. *Lancet*. 2017;389(10078):1558–1580.
4. Petticrew M, Maani N, Knai C, Weiderpass E. The strategies of alcohol industry SAPROs: inaccurate information, misleading language and the use of confounders to downplay and misrepresent the risk of cancer. *Drug Alcohol Rev*. 2018;37(3):313–315.

5. Petticrew M, Maani N, Knai C, Weiderpass E. How alcohol industry organisations mislead the public about alcohol and cancer. *Drug Alcohol Rev.* 2018;37(3):293-303.
6. Connor J. Alcohol consumption as a cause of cancer. *Addiction.* 2017;112(2):222-228.
7. Jernigan D, Noel J, Landon J, Thornton N, Lobstein T. Alcohol marketing and youth alcohol consumption: a systematic review of longitudinal studies published since 2008. *Addiction.* 2016;112(S1):7-20.
8. Thaler R, Sunstein C. *Nudge: Improving Decisions About Health, Wealth and Happiness.* Penguin Group; 2009.
9. Cialdini R. *Pre-Suasion: A Revolutionary Way to Influence and Persuade.* Simon & Schuster; 2016.
10. Cialdini R. *Influence: Science and Practice.* 5th ed. Allyn and Bacon; 2008.
11. Costa E, Halpern D; The Behavioural Insights Team. The Behavioural science of online harm and manipulation and what to do about it. <https://www.bi.team/publications/the-behavioural-science-of-online-harm-and-manipulation-and-what-to-do-about-it/>. April 15, 2019. Accessed July 10, 2020.
12. Eyal N. *Hooked: How to Build Habit-Forming Products:* Portfolio; 2014.
13. House of Commons Digital, Culture, Media and Sport Committee. Disinformation and 'fake news.' Eighth Report of Session 2017–19. Final Report; February 18, 2019. <https://publications.parliament.uk/pa/cm201719/cmselect/cmcmds/1791/1791.pdf>. Accessed October 19, 2019.
14. European Commission. A multi-dimensional approach to disinformation; report of the High Level Expert Group on Fake News and Online Disinformation. <https://ec.europa.eu/digital-single-market/en/news/final-report-high-level-expert-group-fake-news-and-online-disinformation>. Published March 12, 2018. Accessed July 10, 2020.
15. Kahneman D. *Thinking, Fast and Slow.* Farrar, Straus and Giroux; 2012.
16. Hollands G, Bignardi G, Johnston M, et al. The TIPPME intervention typology for changing environments to change behaviour. *Nat Hum Behav.* 2017;1(0140). <https://doi.org/10.1038/s41562-017-0140-17>.
17. Marcano-Olivier M, Pearson R, Ruparell A, Horne P, Viktor S, Erjavec M. A low-cost behavioural nudge and choice architecture intervention targeting school lunches increases children's consumption of fruit: a cluster randomised trial. *Int J Behav Nutr Phys Act.* 2019;16(1):20.

18. Hotard M, Lawrence D, Laitin D, Hainmueller J. A low-cost information nudge increases citizenship application rates among low-income immigrants. *Nat Hum Behav.* 2019;3(7):678-683.
19. Marteau T, Ogilvie D, Roland M, Suhrcke M, Kelly M. Judging nudging: can nudging improve population health? *Br Med J.* 2011;342:d228.
20. Newall P. Gambling on the dark side of nudges. <https://behavioralscientist.org/gambling-dark-side-nudges/>. Published June 27, 2017. Accessed July 28, 2019.
21. Schüll N. *Addiction by Design: Machine Gambling in Las Vegas*. Princeton University Press; 2012.
22. Sunstein C. Sludge audits. Harvard Public Law Working Paper No. 19–21. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3379367. April 27, 2019. Accessed December 2, 2019.
23. Hallsworth M, Snijders V, Burd H, et al. Applying behavioral insights: simple ways to improve health outcomes: report of the WISH Behavioral Insights Forum 2016. [https://www.imperial.ac.uk/media/imperial-college/institute-of-global-health-innovation/Behavioral_Insights_Report-\(1\).pdf](https://www.imperial.ac.uk/media/imperial-college/institute-of-global-health-innovation/Behavioral_Insights_Report-(1).pdf). Accessed July 17, 2019.
24. Parke J, Rigbye J; Responsible Gambling Trust. Self-exclusion as a gambling harm minimisation measure in Great Britain. <https://about.gambleaware.org/media/1176/rgt-self-exclusion-report-parke-rigbye-july-2014-final-edition.pdf>. Published July 2014. Accessed July 10, 2020.
25. Bhattacharya A, Angus C, Pryce R, Holmes J, Brennan A, Meier P. How dependent is the alcohol industry on heavy drinking in England? *Addiction.* 2018;113(12):2225-2232.
26. Casswell S, Callinan S, Chaiyasong S, et al. How the alcohol industry relies on harmful use of alcohol and works to protect its profits. *Drug Alcohol Rev.* 2016;35(6): 661-664.
27. McCambridge J, Kypri K, Miller P, Hawkins B, Hastings G. Be aware of Drinkaware. *Addiction.* 2014;109(4):519-524. <https://doi.org/10.1111/add.12356>.
28. McCambridge J, Mialon M, Hawkins B. Alcohol industry involvement in policy making: a systematic review. *Addiction.* 2018;113(9):1571-1584.
29. Mialon M, McCambridge J. Alcohol industry corporate social responsibility initiatives and harmful drinking: a systematic review. *Eur J Public Health.* 2018;28(4):664-673.
30. Babor TF, Robaina K. Public health, academic medicine, and the alcohol industry's corporate social responsibility activities. *Am J Public Health.* 2013;103(2):206-214.

31. Lim A, Van Schalkwyk M, Maani N, Petticrew MP. Pregnancy, fertility, breastfeeding, and alcohol consumption: an analysis of framing and completeness of information disseminated by alcohol industry-funded organizations. *J Stud Alcohol Drugs*. 2019;80(5):524-533.
32. McCambridge J, Coleman R, McEachern J. Public health surveillance studies of alcohol industry market and political strategies: a systematic review. *J Stud Alcohol Drugs*. 2019;80:149-157.
33. Savell E, Fooks G, Gilmore AB. How does the alcohol industry attempt to influence marketing regulations? a systematic review. *Addiction*. 2016;111(1):18-32.
34. Maani N, Schalkwyk M, Thomas S, Petticrew M. Alcohol industry CSR organisations: what can their Twitter activity tell us about their independence and their priorities? a comparative analysis. *Int J Environ Res Public Health*. 2019;16(5):892.
35. Maani N, Petticrew M. What does the alcohol industry mean by “responsible drinking”? a comparative analysis. *J Public Health (Oxf)*. 2018;40(1):90-97.
36. Alcohol and cancer fact sheet. London, UK; Drinkaware. <http://www.resourcesorg.co.uk/assets/pdfs/Alcohol-and-cancer.pdf>. Accessed July 31, 2020.
37. Alcohol and the body: Drinkaware Ireland poster. https://web.archive.org/web/20171118032459/https://www.drinkaware.ie/docs/Drinkaware_Alcohol+Body_Poster%20FINAL.pdf. Accessed July 25, 2020.
38. DrinkIQ website. How alcohol affects us. <https://web.archive.org/web/20200727112253/https://www.drinkiq.com/en-gb/how-alcohol-affects-us/the-body/alcohols-effect-on-your-body/>. Accessed July 31, 2020.
39. @drinkaware. Didn't drink at all during January and looking forward to drinks tonight? Remember you've reset your tolerance. Check out what this means: <http://bit.ly/2n7hOxG>. <https://twitter.com/Drinkaware/status/1091260075350650880>. Posted February 1, 2019. Accessed August 25, 2020.
40. @drinkaware. It's tradition to ring in the New Year with a glass of bubbly, but what's in champagne? <http://bit.ly/2B8n1Lj>. <https://twitter.com/Drinkaware/status/946328623807639553>. Posted December 28, 2017. Accessed August 25, 2020.
41. Drinkaware website. Worrying number of people drinking to cope with day-to-day pressures. <https://web.archive.org/web/20200810114017/https://www.drinkaware.co.uk/professionals/press/worrying-number-of-people-drinking-to-cope-with-day-to-day-pressures>. January 11, 2018. Accessed August 25, 2020.

42. The 8 benefits of moderate drinking. Édu'alcool website. <http://educalcool.qc.ca/en/facts-tips-and-tools/facts/8-benefits-of-moderate-drinking/#.XYi7xnt7k2w>. Accessed July 10, 2020.
43. @drinkaware. More Drink Free Days each week can help you lose weight and give you more time to get active. Why not get down your local pool and have a splash about. #DrinkFree-Days #NoAlcoholidays. <https://twitter.com/Drinkaware/status/1175726548080648197>. Posted September 22, 2019. Accessed August 25, 2020.
44. @drinkaware. Today's the day when people most commonly give up on their Dry January efforts. If you've fallen off the wagon, then don't despair – you can get back on it, no problem. Find out more about lapses here: <http://bit.ly/2SOiXGg>. <https://twitter.com/Drinkaware/status/1086911314277138433>. Posted January 20, 2019. Accessed August 25, 2020.
45. @drinkaware. Before things start spinning, switch to water. What's your top tip for staying safe on holiday? Great advice courtesy of Travel Aware #StickWithYourMates. <https://twitter.com/Drinkaware/status/1154058732302733312>. Posted July 24, 2019. Accessed August 25, 2020.
46. Drinkaware website. Health effects of alcohol. <https://web.archive.org/web/20200102012907/https://www.drinkaware.co.uk/alcohol-facts/health-effects-of-alcohol/>. Accessed August 25, 2020.
47. Entman R. Framing: toward clarification of a fractured paradigm. *J Commun.* 1993;43(4):51-58.
48. Drinkaware for Education. Drinkaware website. <https://resources.drinkaware.co.uk/support-our-campaigns/drinkaware-for-education/>. Accessed January 1, 2020.
49. Why do young people drink? Drinkaware Ireland website. <https://www.drinkaware.ie/parents/notice/#why-do-young-people-drink>. Accessed January 1, 2020.
50. Public Health England. The public health burden of alcohol and the effectiveness and cost-effectiveness of alcohol control policies: an evidence review. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/733108/alcohol_public_health_burden_evidence_review_update_2018.pdf. Published December 2016. Accessed July 10, 2020.
51. Sklar J. *Principles of Web Design*. 5th ed. Cengage Learning; 2011.
52. Brebion A. Above the fold vs. below the fold: does it still matter in 2019? AB Tasty. <https://www.abtasty.com/blog/above-the-fold/>. February 4, 2018. Accessed July 10, 2020.

53. Shrestha S, Lenz L. Eye gaze patterns while searching vs. browsing a website. *Usability News*. 2007;9(1). <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.561.296&rep=rep1&type=pdf>. Accessed July 10, 2020.
54. Google. The importance of being seen: viewability insights for digital marketers and publishers. November 2014. https://think.storage.googleapis.com/docs/the-importance-of-being-seen_study.pdf. Accessed July 27, 2019.
55. Zehra N. Eye tracking studies and what they say about website conversion optimization. VWO Blog. <https://vwo.com/blog/6-eye-tracking-studies-and-what-do-they-say-about-website-conversion-optimization/?rel=ugc>. Updated June 5, 2020. Accessed July 22, 2020.
56. Peticrew M, Douglas N, Knai C, Durand M, Eastmure E, Mays E. Health information on alcoholic beverage labels in the UK: has the alcohol industry's voluntary agreement to improve labelling been met? *Addiction*. 2016;111(1):51-55.
57. Dolan P. *Happiness by Design: Change What You Do, Not How You Think*: Avery; 2015.
58. Hollands G, Shemilt I, Marteau T, et al. Portion, package or tableware size for changing selection and consumption of food, alcohol and tobacco. *Cochrane Database Syst Rev*. 2015;9:CD011045. <https://doi.org/10.1002/14651858.CD011045.pub2>.
59. Cabinet Office/Institute for Government. MINDSPACE: influencing behaviour through public policy. <https://www.instituteforgovernment.org.uk/sites/default/files/publications/MINDSPACE.pdf>. Published 2010. Accessed August 23, 2019.
60. Facts about drinking. Drinkwise website. <https://drinkwise.org.au/#q=facts%20about%20drinking&r=true>. Accessed January 1, 2020.
61. Druckman J. On the limits of framing effects: who can frame? *J Polit*. 2001;63(4):1041-1066.
62. Hollsworth M, Egan M, Rutter J, McCrae J; Behavioural Insights Team. Behavioural government: using behavioural science to improve how governments make decisions. <https://www.instituteforgovernment.org.uk/publications/behavioural-government>. Published 2018. Accessed July 10, 2020.
63. Dewulf A, Bowen R. Issue framing in conversations for change: discursive interaction strategies for "doing differences". *J Appl Behav Sci*. 2012;48(2):168-193.
64. Alcohol and cancer. Drinkaware fact sheet. <http://www.resources.org.co.uk/assets/pdfs/Alcohol-and-cancer.pdf>. Accessed August 25, 2020.

65. Michaels D. *The Triumph of Doubt: Dark Money and the Science of Deception*. Oxford University Press; 2020.
66. Oreskes N, Conway E. *Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming*. Bloomsbury; 2012.
67. Proctor R. *Cancer Wars: How Politics Shapes What We Know and Don't Know About Cancer*. Basic Books; 1995.
68. Alcohol and your health. Drinkwise Australia website. <https://web.archive.org/save/https://drinkwise.org.au/#q=alcohol%20and%20your%20health&r=true>. Accessed December 2, 2019.
69. Sivanathan N, Kakkar H. The unintended consequences of argument dilution in direct-to-consumer drug advertisements. *Nat Hum Behav*. 2017;1:797-802.
70. Bonell C, McKee M, Fletcher A, Wilkinson P, Haines A. One nudge forward, two steps back. *Br Med J*. 2011;342:d401.
71. @drinkaware. When the sun's out, it can be harder to stick to your #DrinkFreeDays - but here are some quick tips on how you can stay motivated and keep on track. <https://web.archive.org/web/20190812110456/https://twitter.com/Drinkaware/status/1160853587423092737>. Posted August 12, 2019. Accessed July 24, 2020.
72. @drinkaware. Today's the day when people most commonly give up on their Dry January efforts. If you've fallen off the wagon, then don't despair - you can get back on it, no problem. <https://twitter.com/Drinkaware/status/1086911314277138433>. Posted January 20, 2019. Accessed July 24, 2020.
73. @drinkaware. It's tradition to ring in the New Year with a glass of bubbly, but what's in champagne? <https://web.archive.org/save/https://twitter.com/Drinkaware/status/1078953890643697664>. Posted December 29, 2018. Accessed July 24, 2020.
74. @drinkaware. Two thirds of regular drinkers say cutting down is harder than improving diet and exercise - Just some of the findings from our new campaign with @PHE_uk #drinkfreedays campaign. <https://twitter.com/Drinkaware/status/1039065599438217217>. Posted September 10, 2018. Accessed July 24, 2020.
75. @drinkaware. New Drinkaware & @YouGov research reveals that almost 3/5 (58%) of all people (aged 18-75) who drink alcohol are doing so because it helps them to cope with the pressures of day to day life. Read the full story here <http://bit.ly/2DdvsXl> #BlueMonday. <https://web.archive.org/save/https://twitter.com/Drinkaware/status/952819943074160640>. Posted January 20, 2018. Accessed July 24, 2020.

76. Petticrew M, Fitzgerald N, Durand M, Knai C, Davoren M, Perry I. Diageo's 'Stop Out of Control Drinking' campaign in Ireland: an analysis. *PLoS ONE* 2016;11(9):e0160379.
77. Oraby T, Thampi V, Bauch C. The influence of social norms on the dynamics of vaccinating behaviour for paediatric infectious diseases. *Proc Biol Sci.* 2014;281(1780):20133172.
78. Center for European Policy Analysis (CEPA). Disinformation techniques. <https://www.cepa.org/disinfo-techniques>. Accessed October 19, 2019.
79. Thornton M, Hawkins B. Between a rock and a hard place: economic expansion and social responsibility in UK media discourses on the global alcohol industry. *Soc Sci Med.* 2017;175:169-176.
80. @ABInBev_EU. #NudgingForGood is also the philosophy behind our #SmartDrinking goals supporting our consumers in making the right choices whatever the occasion without compromising on taste or experience! #NABLAB #BrewingBetterTogether @AIMbrands ab-inbev.com/what-we-do/sma... https://twitter.com/ABInBev_EU/status/1184909191615524864?s=20. Posted October 17, 2019. Accessed July 10, 2020.
81. Nudging for Good. European Brands Association website. <http://www.nudgingforgood.com/category/case-study/>. Accessed December 2, 2019.
82. Kahneman D, Slovic P, Tversky A, eds. *Judgement Under Uncertainty: Heuristics and Biases*. Cambridge University Press; 1982.
83. UCSF Truth Initiative. Analysis of post-verdict juror interview in *Kueper v. R.J. Reynolds Tobacco Company, et al.* <https://www.industrydocuments.ucsf.edu/tobacco/docs/#id=mrwn0050>. Accessed July 27, 2019.
84. Smith E, Malone R. Philip Morris's health information web site appears responsible but undermines public health. *Public Health Nurs.* 2008;25(6):554-564.
85. Maani N, Bertscher A, Critchlow N, et al. Recruiting the "heavy-using loyalists of tomorrow": an analysis of the aims, effects and mechanisms of alcohol advertising, based on advertising industry evaluations. *Int J Env Res Public Health.* 2019;16(21): E4092. <https://doi.org/10.3390/ijerph16214092>.
86. Larson H. The biggest pandemic risk? viral misinformation. *Nature.* 2018;562:309.
87. Meppelink C, Smit E, Fransen M, Diviani N. "I was right about vaccination": confirmation bias and health literacy in online health information seeking. *J Health Commun.* 2019;24(2):129-140.
88. Zaval L, Cornwell J. Cognitive biases, non-rational judgments, and public perceptions of climate change. In: Nisbet MC, ed.

- The Oxford Encyclopedia of Climate Change Communication*. Oxford University Press; 2016.
89. Dobbs P, Clawson A, Gowin M, Cheney M. Where college students look for vaping information and what information they believe. *J Am Coll Health*. 2020;68(4):347-356.
 90. Kanter G, Carpenter D, Lehmann L, Mello M. US nationwide disclosure of industry payments and public trust in physicians. *JAMA Netw Open*. 2019;2(4):e191947.
 91. Chivers T. What's next for psychology's embattled field of social priming. *Nature*. 2019;576(7786):200-202. <https://www.nature.com/articles/d41586-019-03755-2>. Accessed July 10, 2020.
 92. Drinkaware Impact Report 2016. London, UK; Drinkaware. https://www.drinkaware.co.uk/media/mbkp0230/da_alcohol_impact_report_2016_final.pdf. Accessed August 20, 2019.
 93. House of Lords Select Committee on Science and Technology, Inquiry on Behaviour Change. Drinkaware evidence. 2011. <https://www.parliament.uk/documents/lords-committees/science-technology/behaviourchange/BCOralandWrittenEvCompiled180711.pdf>. Accessed August 20, 2019.
 94. House of Lords Select Committee on Science and Technology, Subcommittee I: Inquiry on Behaviour Change, January 19, 2011. Written Evidence (BC98):p324.
 95. House of Commons Science and Technology Committee. Oral evidence: Work of the Chief Medical Officer, HC 2121. June 6, 2019.

Funding/Support: Mark Petticrew is a co-investigator in the SPECTRUM consortium, which is funded by the UK Prevention Research Partnership, a consortium of UK funders (UK Research and Innovation Research Councils: Medical Research Council, Engineering and Physical Sciences Research Council, Economic and Social Research Council, and Natural Environment Research Council; charities: British Heart Foundation, Cancer Research UK, Wellcome, and The Health Foundation; government: Scottish Government Chief Scientist Office, Health and Care Research Wales, National Institute of Health Research [NIHR], and Public Health Agency [NI]). Luisa Pettigrew is funded by a National Institute of Health Research (NIHR) Doctoral Research Fellowship. May Van Schalkwyk is funded by a NIHR Doctoral Fellowship. Nason Maani is supported by a Harkness Fellowship funded by the Commonwealth Fund. The views presented here are those of the authors and should not be attributed to the Commonwealth Fund; its directors, officers, or staff; or any other named funder. Similarly in the case of NIHR-funded authors, the views expressed are those of the authors and not

necessarily those of the National Health Service, the NIHR, or the Department of Health and Social Care.

Conflict of Interest Disclosures: All authors have completed the ICMJE Form for Disclosure of Potential Conflicts of Interest. No conflicts were reported.

Address Correspondence to: Mark Petticrew, London School of Hygiene and Tropical Medicine, 15–17 Tavistock Place, London, UK, WC1H 9SH (email: mark.petticrew@lshtm.ac.uk).