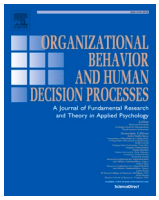
Organizational Behavior and Human Decision Processes 176 (2023) 104244

Contents lists available at ScienceDirect

Organizational Behavior and Human Decision Processes

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



Tainted nudge

Despoina Alempaki [a,*], Andrea Isoni [a,b,*], Daniel Read [a,*]

[a] Behavioural Science Group, Warwick Business School, Coventry, UK
[b] University of Cagliari, Italy

ARTICLE INFO

Keywords:
Nudges
Win–win initiatives
Profit
Prosocial actions
Deception

ABSTRACT

Nudges are increasingly used by governments and organizations to promote behaviors like healthy eating or effective financial planning. Due to their cost-effectiveness, such nudges may earn a profit for the nudger. We investigate whether this profit taints nudges, as suggested by recent research showing that altruistic acts can be regarded less favourably if they result in private benefits to the actor. Across seven preregistered experiments, we demonstrate that prosocial nudges are indeed rated less positively if a profit is earned. But this tainting is limited: prosocial but profitable nudges are evaluated much more favourably than merely profitable ones, unless profit-motivated nudgers deceptively claim their motive is prosocial. Our findings apply to both for-profit and non-profit organizations and provide behaviorally informed guidelines for the introduction of nudge interventions. We suggest organizations can avoid the potential risk of backlash by openly disclosing the win–win nature of their prosocial nudges.

1. Introduction

Nudges are used by governments to promote positive behaviors, such as healthy eating, organ donation, school attendance, vaccine uptake and saving for retirement (e.g., Benartzi et al., 2017; Chapman et al., 2021; DellaVigna and Linos, 2022; John et al., 2014; Kalil et al., 2021; OECD, 2019; Pennycook et al., 2020; Sunstein and Reisch, 2017; Vandenbroele et al., 2021). They are also finding increasing application in the private sector, with companies using what is often called "nudge theory" to improve customer and employee well-being, and to generate social benefits (e.g., Alemanno, 2016; Gosnell et al., 2020; Haugh, 2017). For example, many methods used by the beverage industry to promote low-calorie variants of their soft drinks are explicitly nudge-inspired (Strom, 2014), as illustrated by Indra Nooyi, former CEO of PepsiCo, who declared that "We've taken lessons from Richard Thaler and Cass Sunstein's book. We try to put portion-control packages out front on the shelves. We make sure our diet products are merchandised as aspirationally as our full-sugar products are" (Adi, 2015).

One widely advertised advantage of nudges is their potential to change behavior at a low cost. This opens the possibility that nudging will be profitable. For the beverage industry, artificial sweeteners cost less than sugar (e.g., Sharma et al., 2016; Tandel, 2011), so successfully nudging consumers toward diet drinks may increase profits, even if that is not the primary goal of the nudge. Similar prosocial and potentially profitable nudges include reducing missed hospital appointments via appeals to social norms (e.g., Martin et al., 2012), minimizing food waste and food consumption by downsizing plates in restaurant buffets (e.g., Kallbekken and Sælen, 2013), and decreasing transaction costs by making paperless billing the default option (e.g., Boruchowicz, 2021). Unsurprisingly, organizations are on the lookout for such "win–win" opportunities that create value both for them and for the wider society (e.g., Gosnell et al., 2020; Kallbekken and Sælen, 2013; Koh et al., 2019).

While the potential profitability of nudges makes them attractive to those implementing them, an important but largely overlooked question is how this profitability affects public support for the nudges themselves or the organization deploying them. Does it matter if nudges designed to promote social welfare also conveniently produce profits? How do individuals perceive win–win nudges[a], and how do these perceptions affect their willingness to support those nudges as well as the organizations that adopt them?

The answers to these questions are not obvious. If people care about the social good that comes from a nudge, they should be supportive of win–win nudges. It is only when nudges are profitable that they are likely to be sustainable in the long run, since organizations are more

* Corresponding authors at: Behavioural Science Group, Warwick Business School, Coventry, UK.
E-mail addresses: despoina.alempaki@wbs.ac.uk (D. Alempaki), andrea.isoni@wbs.ac.uk (A. Isoni), daniel.read@wbs.ac.uk (D. Read).

[a] A win–win nudge could in principle refer to a prosocially motivated nudge with personal benefits such as profits or a personally motivated nudge with social benefits. We focus on prosocially motivated nudges that result in profits.

https://doi.org/10.1016/j.obhdp.2023.104244
Received 30 May 2022; Received in revised form 21 February 2023; Accepted 13 March 2023
Available online 29 April 2023

likely to continue profitable initiatives than unprofitable ones (e.g., Eccles and Serafeim, 2013). This applies not merely to purely commercial organizations but even to governments and non-profits, all of whom must keep an eye on the bottom line. Pragmatic consumers might therefore welcome initiatives that do good for society while earning a profit for the nudger.

However, a now substantial line of work in the context of altruism and charitable giving suggests that earning a profit from an altruistic action reduces the esteem in which the altruistic action is held. To get full credit for altruistic acts, the purported altruist may even need to undergo some personal sacrifice (e.g., Johnson, 2018; Olivola, 2011; Lin-Healy and Small, 2013; Small and Cryder, 2016). In fact, altruism is so inherently linked to sacrifice that an actor is given less credit when there are personal benefits from their action, even when they are entirely out of the actor's control (Lin-Healy and Small, 2013). Even speaking up about one's altruism can lead to negative evaluations, due to suspicions that the altruistic act was motivated by the desire to enhance one's reputation (Berman et al., 2015). Newman and Cain (2014) use the term *tainted altruism* to describe how altruistic actions can be judged negatively if they result in profit or other personal benefits for the purported altruist.

These findings open the possibility that organizations adopting win–win nudges might be vulnerable to an analogous *tainted nudge* effect, such that profitable nudges receive negative evaluations from consumers or the general public. If profits do taint nudges, managers and policymakers may be unpleasantly surprised when win–win interventions are received less positively than expected, or even backfire.

We cannot automatically conclude from previous research that nudges will be tainted in the same way as altruistic acts, since cost-effectiveness is one of the defining characteristics of nudges (e.g., Benartzi et al., 2017; Hotard et al., 2019). Organizations will typically implement nudges as opposed to other ways of achieving the same prosocial goal because they want to achieve welfare improvements in a cost-effective manner (several examples can be found in Chapman et al., 2021). Perhaps this economic consideration is already embedded in people's mindset when they evaluate nudges. So, they may not see anything wrong in profiting from prosocial nudges, in contrast to apparently "pure" altruism, where a cost-effective maximization of welfare is usually not considered (e.g., Olivola, 2011; Olivola and Shafir, 2013; Oppenheimer and Olivola, 2011). This possibility is consistent with recent findings by Cadario and Chandon (2019), who reported that survey respondents were more likely to approve of a healthy nudge they thought would benefit both health and business, as opposed to one benefitting primarily only one of the two.

Most previous research on tainting involves identifiable individuals taking altruistic or charitable actions. Some evidence suggests that organizations implementing profitable prosocial nudges could be less morally condemned than individuals undertaking charitable actions for personal benefit. Haran (2013), for example, reported that people view the breach of contract by an individual as a moral transgression, but as a legitimate business decision when implemented by an organization. Similarly, Jago et al. (2019) showed that people attribute fewer moral convictions to organizations than to individuals, partially because they assume organizations are more self-interested and likely to be motivated by strategic rather than moral concerns.

Another difference might relate to the nature of giving in a charitable context. Altruism in the form of giving to charity may be judged more harshly if it also leads to a profit, as earning profits when doing good means that the organization could have done more for the charity. For instance, if a charity pays a fundraiser £100, then that could be seen as money withheld from the charitable purpose; and if an organization donates £100 to a charity while making £100 profit, they could have also donated this extra £100 to the charity. But if an organization nudges their employees to eat better food, it is not obvious that they could nudge them into healthier eating habits if they did not earn a profit while doing so. The link from profit to an inferior service is less direct in nudges, and as such, people might be more forgiving of a company that makes profit while nudging for a good cause.

The goal of this paper is to establish whether, and if so to what extent and under what circumstances, win–win nudges by organizations are tainted by their profitability. While previous work has focused primarily on altruism, our paper is the first to examine whether tainting also occurs for organizational nudges.

2. Background

Previous research into the tainting effect of profit has focused on altruistic acts. For organizations, these might take the form of actions that directly benefit a group that is not expressly their customers, perhaps by building a hospital, or through making a charitable contribution. For instance, in one of Newman and Cain's (2014) examples, "Daniel P.'s" organization handles fundraising events to raise donations for charity. The potential for tainting comes from the profits that Daniel P. and his staff make out of the fundraising drives.

We study whether tainting also occurs for nudges. The term "nudge" is one of the most controversial in behavioral science and has been defined in several ways. We adopt the streamlined definition by Thaler and Sunstein (2008, p. 6): "A nudge is any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives." This definition is clean because it focuses on nudging as a technology to influence behavior, and not on the specific effects of nudging. Organizational nudging differs from altruism in that the nudger influences the behavior of a group, typically its clients or those in its care, by varying the choice architecture. Under ideal circumstances, the nudge benefits the target group, but that is not a requirement. For instance, a supermarket putting fruit next to the checkout is nudging its customers, but so is one putting chocolate in the same location. A nudge is sometimes called "sludge" if the manipulation of the choice architecture is done to benefit the nudger at the expense of the nudgee (e.g., Thaler, 2018; Sunstein, 2022). When we refer to tainted nudge, we focus on the possibility that nudges intended to serve a prosocial purpose are judged poorly because they also earn a profit.

There is substantial evidence showing nudges can have significant sizeable and lasting effects on behavior (see, e.g., DellaVigna and Linos, 2022; Hummel and Maedche, 2019 for overviews), but there is still much to learn about what determines when a nudge is likely to receive public approval (e.g., Jung and Mellers, 2016; Lades and Delaney, 2022; Sunstein, 2019). Previous studies suggest that, while there tends to be overall support for nudging (e.g., Hagman et al., 2015; Sunstein and Reisch, 2019; Tannenbaum et al., 2017), there are substantial pockets of resistance, which vary from nudge to nudge. Yan and Yates (2019), for instance, found that public acceptance ranges from 20% to 90% for different types of nudges. The broad trend is for people in most countries, and with most political attitudes, to be positively disposed toward nudging, especially when it is used to promote outcomes they perceive as legitimate or worthwhile (e.g., Clavien, 2018; Hagman et al., 2015; Lades and Delaney, 2022; Sunstein and Reisch, 2019). However, when nudges are directed toward purely selfish ends, such as being elected or selling a product, they are less well received (e.g., Clavien, 2018). Many researchers and organizations have proposed codes concerning how nudging should be deployed and for what purposes, to ensure their ethicality and their broad acceptability (e.g., Jachimowicz et al., 2017; Lades and Delaney, 2022).

Research into altruistic actions and charitable giving has found evidence that, when social and personal benefits are combined, (purportedly) altruistic actors can be exposed to a backlash, with the personal benefits tainting the social good achieved (e.g., Berman et al., 2015; Carlson and Zaki, 2018; Lin-Healy and Small, 2012, 2013; Newman and Cain, 2014; see also Berman and Silver, 2022; Raihani and Power, 2021 for reviews). It is claimed that the backlash is often so large that mixed social and personal initiatives are judged less favourably than purely

personal ones (e.g., Alcala et al., 2022; Carlson and Zaki, 2018; Newman and Cain, 2014). To illustrate, the man in one of Newman and Cain's (2014) studies who volunteers at either a homeless shelter or a coffee shop with the hidden goal of obtaining a romantic connection with a co-worker was perceived as less moral when volunteering at the homeless shelter. Similarly, Carlson and Zaki (2018) found that someone giving blood to receive a gift certificate was rated less favourably than someone doing a plainly neutral action (e.g., going to see a film). Although Newman and Cain used the term "tainted altruism" when charitable actions that are personally as well as socially beneficial are judged more negatively than overtly self-interested ones that are only personally beneficial, we will (for reasons discussed shortly) refer to this as *strong* tainted altruism.

A proposed explanation for strong tainted altruism is asymmetric counterfactual thinking (e.g., Newman and Cain, 2014), a concept with a long history in cognitive and social psychology (e.g., Galinsky and Moskowitz, 2000; Kahneman, 2014; Kahneman and Miller, 1986; Miller and McFarland, 1986; Roese, 1997; Roese and Olson, 2014). Intendedly charitable actions that also personally benefit the actor are contrasted with ones that only produce that personal benefit. Here, the personal benefit is not a purely psychological one like a "warm glow" from giving, but a tangible benefit such as obtaining a romantic connection or earning extra money in the form of commissions or profits (see Carlson and Zaki, 2018, for explicit tests of this assumption). When people observe an (apparently) prosocial action that also brings personal benefits, this automatically brings to mind alternative actions that are equally prosocial but without the personal benefits. As Newman and Cain put it, "People consider the same behavior as it might occur in the absence of self-interest and ultimately conclude the person (or organization) did not behave as altruistically as he or she could have" (p. 649).

Therefore, the personal benefit added to the prosocial action cues individuals to remember or mentally construct alternative prosocial actions without the personal benefit. When people consider an action with only a single consequence, such as one that is *merely* prosocial or *merely* self-interested, the cue is absent and so the alternatives are not easily brought to mind. Consequently, while the prosocial action that yields a personal benefit is likely to be regarded unfavourably because it is personally beneficial, the *merely* personal or prosocial action is not judged unfavourably, since individuals will not automatically consider whether the actor could have been, respectively, more prosocial or more selfish.

We apply this logic to prosocial nudges and investigate whether a prosocial nudge that is profitable for the nudger is judged less favourably than a merely prosocial one, leading to what we call a *tainted nudge* effect. Moreover, as suggested by some of the key findings in the tainted altruism literature, we investigate whether a prosocial but also profitable nudge might be unfavourably evaluated compared to a merely profitable one, leading to a *strong* tainted nudge effect.

3. Research overview and formal statement of hypotheses

We conducted our experiments in the domains of health and financial planning. Experiments 1a through 3 focus on attitudes towards organizations nudging their clients to eat more healthily. Promoting healthy behaviors is one of the most prominent targets of nudge-based interventions (e.g., Arno and Thomas, 2016; Cadario and Chandon, 2020; Cohen et al., 2016; Mertens et al. 2022; Vlaev et al., 2016), and many organizations already nudge their customers and employees toward healthy eating (see, for example, Chance et al., 2016 for the approach followed by Google). Moreover, health-related nudges are amongst the most universally acceptable (Hagman et al., 2015). Experiments 4a and 4b extend our work to the context of financial planning, another domain where nudges are widely used to improve long-term savings and debt management, and decrease errors in personal investment decisions (e.g., Choi et al., 2003; Madrian and Shea, 2001; Thaler and Benartzi, 2004).

We examine nudges that produce either a "social benefit" (i.e., prosocial nudges that benefit those being nudged), a "profit" (i.e., nudges benefitting the nudging organization), or both. For short, we use the terms "Social" and "Profit" to refer to these benefits. Because the benefits of a nudge are in principle distinct from its underlying motive (e.g., a nudge motivated by a social benefit may also bring profit), when discussing nudges with multiple benefits, we will use compound terms such as "Social + Profit" where the first term indicates the motive, and the second term an additional consequence.

In Experiments 1a through 3, participants evaluated an organization nudging their clients to change their diet. The nudge resulted in a social benefit (in the form of health for the clients), a profit for the organization, or both. In Experiments 4a and 4b, we investigated analogous scenarios involving a financial planning organization nudging their clients to make different financial decisions, which again resulted in social benefits (more successful investments for the clients), profits or both. Our hypotheses are summarized in Fig. 1.

Our first hypothesis is that a purely profitable nudge with no social benefits will be evaluated less positively than one that offers a purely social benefit. We denote this as H1: Profit < Social, to mean that key dependent variables – described in detail in the next section – will be lower in the Profit than in the Social condition. Hypothesis 1 is a logical prerequisite for tainting. If profitable nudges were regarded more favourably than social ones, it would be hard to interpret the effect of a secondary profit as tainting an intended prosocial benefit.

We tested for tainting by comparing the Social + Profit condition to the Social condition. According to Hypothesis 2, if profit taints a social nudge, that nudge will be perceived less favourably than an equivalent social nudge that earns no profit. That is H2: Social + Profit < Social (*tainted nudge*).

If the tainting is such that the Social + Profit nudge is evaluated less favourably than the Profit nudge which only benefits the organization, this would constitute a *strong tainted nudge*, as expressed in our Hypothesis 3, H3: Social + Profit < Profit.

All our experiments employed closely related scenarios to minimize the differences between stimuli and allow for precise hypothesis testing. To access sufficiently large and diverse samples, the experiments were conducted on Prolific (https://www.prolific.co). Participation was restricted to adult UK residents with a Prolific approval rate over 90% (a measure of participant reliability). Prior to data collection, we preregistered all hypotheses, sample size determination, procedures and analysis plans at Open Science Framework (https://osf.io/4g83d/). We report all conditions and all measures in all experiments and follow the preregistered analysis plans precisely. Some experimental conditions and preregistered analyses are included in the Supplementary Material. Because our hypotheses are directional, we preregistered one-sided tests. However, for the effect sizes and associated 95% confidence intervals, we report two-sided estimates, as these are bounded in both directions and easier to interpret. Any exploratory analysis is clearly indicated as such and always based on two-sided tests. All experiments have unique (non-overlapping) samples. In Experiments 1a–3, we aimed for relatively large samples and set the a priori target of recruiting at least 100 participants for each experimental condition. A power analysis suggested a minimum sample size of N = 88 participants per experimental condition to achieve at least 95% power to detect a medium-sized effect ($d = 0.50$; $\alpha = 0.05$). Power calculations were conducted using G*Power version 3.1.9.4 (Faul et al., 2009). Following a suggestion from a reviewer, in Experiments 4a and 4b we approximately doubled the sample size. Overall, including the experiments and conditions reported in the Supplementary Material, we obtained data from 4810 individuals.

4. Experiments 1a and 1b – Tests of hypotheses 1 to 3

We first established whether people evaluate solely profitable nudges less favourably than solely prosocial ones (H1). We also examined

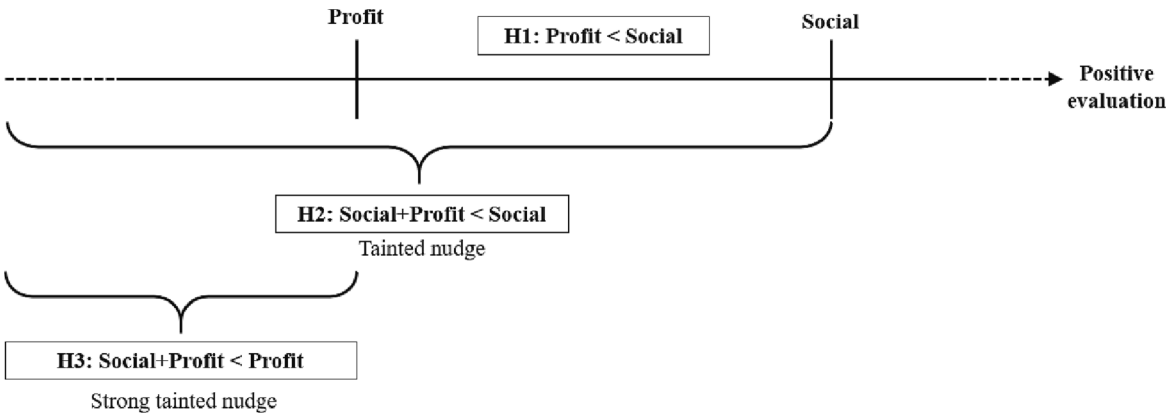


Fig. 1. Hypotheses 1–3.
Notes. The right facing arrow indicates the positive evaluation assigned to an initiative. The braces denote the range of evaluations consistent with each hypothesis.

Table 1
Scenarios used in Experiment 1a.

Wessex Evening Standard

New "default" meals at Wessex Uni prove a success

Residential students at the University of Wessex pay in advance for the hot evening meals they receive during term time. Each evening there is a "default" meal, which students get unless they fill out a form requesting something different one week in advance.

Social:	*Social + Profit:*	*Profit:*
Last September, the University launched an initiative to **increase healthy eating on campus**. They replaced the default meals with **healthier ones, lower in fat, sugar and calories, but just as tasty and nutritious.**	Last September, the University launched an initiative to **increase healthy eating on campus**. They replaced the default meals with **healthier ones, lower in fat, sugar and calories, but just as tasty and nutritious.**	Last September, the University launched an initiative to **increase the salaries they can offer to their Senior Administrators**. They replaced the default meals with ones that **cost the University significantly less to provide, but are just as tasty and nutritious.**
The annual amount paid by students for their meals remained **unchanged.** Professor Fry, the Dean of the University, said they were inspired by "nudge theory", according to which people usually stick with default options.		
"We have a **moral obligation** to ensure the students in our care eat healthily, so we replaced the defaults because we expected most students would not change their meals," said the Dean, "and in fact most didn't."	"We have a **moral obligation** to ensure the students in our care eat healthily, so we replaced the defaults because we expected most students would not change their meals," said the Dean, "and in fact most didn't."	"We replaced the defaults because we expected most students would not change their meals," said the Dean, "and in fact most didn't."
	The new default meals **cost the University significantly less to provide.** Therefore, the University was also able to **increase the salaries they can offer to their Senior Administrators.**	

whether combining an intended prosocial benefit with profit taints that nudge (H2) and whether the resulting tainting is strong (H3). We conducted two separate experiments (1a and 1b) in which we used the same basic scenarios and varied the selfishness of the initiative and the type of the nudge.[b]

4.1. Experiment 1a

4.1.1. Method

Participants were 312 UK residents (113 men, 191 women, 8 gender nondisclosed, M_{age} = 38.5, SD = 12.7, 2 age nondisclosed) recruited from Prolific in exchange for £0.50. The experiment took on average slightly less than four minutes.

Participants read a hypothetical news story describing how the fictional "University of Wessex" had nudged their students by changing their default meals. We chose the default nudge because it has been identified as the most effective at changing behavior (see e.g., the review by Hummel and Maedche, 2019). The complete scenarios are in Table 1.

Participants were randomly assigned to one of three conditions. The Social condition described a successful health-motivated nudge, the Profit condition described a successful profit-motivated nudge and the Social + Profit condition added profit to the health-motivated nudge.

In the Profit condition, we attempted to exclude potential inferences about additional benefits the nudge could create for the students. To rule out the possibility that participants inferred the profits would be used by the university to provide better services, such as library books or better classrooms, we described the change of default meals as intended to "increase **the salaries** [the University] **can offer to their Senior Administrators**.".

A similar concern relates to potential harms the nudge could bring to students. Indeed, both the public and the private sector are known to use sludges that cost nudgees time or money or deprive them of important services (see e.g., Thaler, 2018; Sunstein, 2022). This possibility is particularly important for investigating tainting effects in nudges, since in the altruism context, the relevant comparison is between altruism combined with self-interest versus self-interest alone. To rule out such confounding inferences, we made it explicit that the new meals were of similar taste and quality, so that the students did not end up worse off in terms of food consumption, and we highlighted that the amount of

[b] In Experiments 1a and 1b, we included one further "Profit+Social" condition that involved adding a social benefit to a profitable nudge. Because this condition is not directly related to our main hypotheses about the tainting effects of profit, we report it in the Supplementary Material (see S5).

money they paid for their meals remained unchanged, so that they did not end up financially worse off.

Following Makov and Newman (2016), when the primary motivation of the nudge was prosocial (Social and Social + Profit conditions), we added a reference to a "moral obligation" to further highlight the moral dimension of the initiative's motivation.[c]

In all conditions, participants rated the initiative on the three dimensions of prosociality, morality, and support[d], with each dimension measured with two 7-point Likert scale items:

Prosociality:	In your opinion, did the University of Wessex change the default meals for **SELFISH MOTIVES** [to **BENEFIT THE STUDENTS**]?
Morality:	In your opinion, how **ETHICAL** [**MORAL**] was the University of Wessex's initiative to change the default meals?
Support:	**DO YOU APPROVE** of [**DO YOU SUPPORT**] the University of Wessex's initiative to change the default meals?

The two items were defined by the variant wordings denoted by the square brackets. All bolding and capitalization was as seen by the participants. The questions were presented in a random order, determined independently for each participant. The Likert-scale was labelled at the endpoints with "Not at all" (1) and "Very much so" (7). The first prosociality item (Selfish motives) was reverse coded.

The two prosociality items were highly correlated (*Spearman's* $\rho = 0.70$, $p < 0.001$), as were the morality (*Spearman's* $\rho = 0.78$, $p < 0.001$) and support items (*Spearman's* $\rho = 0.86$, $p < 0.001$).[e] For each participant, we took an average score for each dimension by averaging the two relevant items.

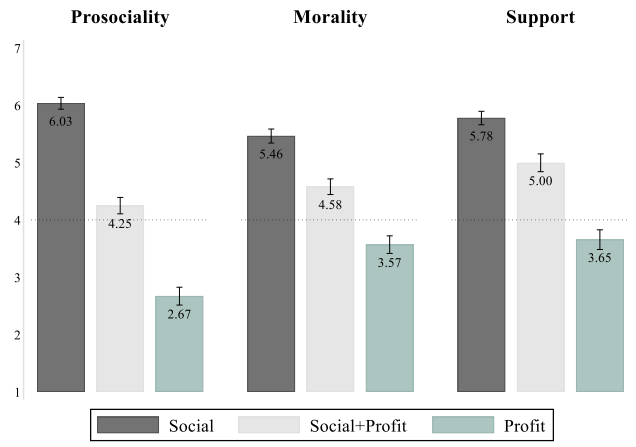


Fig. 2. Mean prosociality, morality and support per condition, Experiment 1a. *Notes*. Error bars show standard errors of the mean. Dashed line indicates scale midpoint.

4.1.2. Results

The results are depicted in Fig. 2, which reports the average of each measure in each condition. In line with Hypothesis 1, the Profit nudge was evaluated less favourably than the Social one on all three dimensions: prosociality, $t(207) = -18.07$, $p < 0.001$, $d = -2.50$, 95% CI [−2.86, −2.14], $BF_{-0} > 100$; morality: $t(207) = -9.67$, $p < 0.001$, $d = -1.34$, 95% CI [−1.64, −1.04], $BF_{-0} > 100$; and support, $t(207) = -10.22$, $p < 0.001$, $d = -1.41$, 95% CI [−1.72, −1.11], $BF_{-0} > 100$.

Whether profit tainted the nudge (H2) was assessed by comparing the Social + Profit and Social conditions. We found significant evidence for tainted nudge across all three measures: prosociality, $t(206) = -10.16$, $p < 0.001$, $d = -1.41$, 95% CI [−1.71, −1.10], $BF_{-0} > 100$; morality, $t(206) = -4.83$, $p < 0.001$, $d = -0.67$, 95% CI [−0.95, −0.39], $BF_{-0} > 100$; and support, $t(206) = -4.02$, $p < 0.001$, $d = -0.56$, 95% CI [−0.83, −0.28], $BF_{-0} > 100$. Earning a profit while doing good was regarded as less moral, less prosocial and less worthy of support than simply doing good.

While the Social + Profit nudge was tainted by profit, this was not strong tainting, as can be clearly seen in Fig. 2. In fact, the Social + Profit condition received consistently *higher* ratings than the Profit condition on all three dimensions. Although we preregistered one-sided tests that do not allow for comparisons going in the opposite direction, here we provide exploratory two-sided tests. In every case, the Social + Profit condition was rated significantly more positively than the Profit condition: prosociality, $t(205) = 7.47$, $p < 0.001$, $d = 1.04$, 95% CI [0.75, 1.33], $BF_{10} > 100$; morality, $t(205) = 4.92$, $p < 0.001$, $d = 0.68$, 95% CI [0.40, 0.96], $BF_{10} > 100$; support, $t(205) = 5.79$, $p < 0.001$, $d = 0.80$, 95% CI [0.52, 1.09], $BF_{10} > 100$.

4.2. Experiment 1b

4.2.1. Method

Participants were 310 UK residents (103 men, 201 women, 6 gender nondisclosed, $M_{age} = 36.8$, $SD = 12.8$, 3 age nondisclosed) recruited from Prolific in exchange for £0.50 (average duration below four minutes). They were randomly allocated to one of the three scenarios depicted in Table 2.

In Experiment 1b, we replaced the default nudge with the classic manipulation of choice architecture which opens Thaler and Sunstein (2008), and which has often been studied in food choice (e.g., Andersson and Nelander, 2021; Gravert and Kurz, 2021; Kurz, 2018). This was the rearrangement of the menu by either listing the healthier meals first (Social and Social + Profit conditions), or the ones that were cheaper to provide (Profit condition). Correspondingly, the change was described as intended to increase healthy eating (Social and Social + Profit) or reduce operating costs (Profit).

As in Experiment 1a, participants rated the university's initiative on prosociality (*Spearman's* $\rho = 0.65$, $p < 0.001$), morality (*Spearman's* $\rho = 0.72$, $p < 0.001$), and support (*Spearman's* $\rho = 0.85$, $p < 0.001$).

4.2.2. Results

As shown in Fig. 3, we find the same main patterns as in Experiment 1a. There was strong evidence for Hypothesis 1, as the initiative was evaluated less positively in the Profit condition than in the Social condition on all three dimensions: prosociality, $t(203) = -21.68$, $p < 0.001$, $d = -3.03$, 95% CI [−3.43, −2.62], $BF_{-0} > 100$; morality, $t(203) = -8.04$, $p < 0.001$, $d = -1.12$, 95% CI [−1.42, −0.83], $BF_{-0} > 100$; support, $t(203) = -7.80$, $p < 0.001$, $d = -1.09$, 95% CI [−1.38, −0.79], $BF_{-0} > 100$.

There was also significant evidence for tainted nudge. In line with Hypothesis 2, the Social + Profit nudge was rated less positively than the Social nudge on all three measures: prosociality, $t(206) = -8.47$, $p < 0.001$, $d = -1.17$, 95% CI [−1.47, −0.88], $BF_{-0} > 100$; morality, $t(206) = -3.43$, $p < 0.001$, $d = -0.48$, 95% CI [−0.75, −0.20], $BF_{-0} = 68.231$;

[c] We were unsure whether referring to a moral obligation would have an effect, and to examine this in Experiments 1a and 1b we included two extra conditions with no reference to the moral obligation. This modification did not have an effect. The results of these additional conditions are reported in the Supplementary Material S2.

[d] An additional item was included in Experiments 1a to 3 to examine whether the nudge intervention was perceived as manipulative (see, e.g., Sunstein, 2015, for a discussion on how whether a nudge involves the consent of the nudgee affects perceptions of manipulativeness). We did not expect any interaction between these ratings and the tainted effects, and did not formulate specific hypotheses about them. We report the questions and the ratings for perceived manipulativeness for all experiments in the Supplementary Material S6.

[e] Correlations are computed using responses from all conditions included in each experiment.

Table 2
Scenarios used in Experiment 1b.

Wessex Evening Standard		
New menu layouts at Wessex Uni prove a success		
Residential students at the University of Wessex pay in advance for the hot evening meals they receive during term time. Students are entitled to choose their meals from the large daily menus displayed in each of the University restaurants.		
Social: Last September, the University launched an initiative to **increase healthy eating on campus**. They rearranged the menus so that the **healthier dishes, lower in fat, sugar and calories were listed first.**	*Social + Profit:* Last September, the University launched an initiative to **increase healthy eating on campus**. They rearranged the menus so that the **healthier dishes, lower in fat, sugar and calories were listed first.**	*Profit:* Last September, the University launched an initiative to **reduce operating costs on campus**. They rearranged the menus so that the dishes that **cost the University less to provide were listed first.**
The annual amount paid by students for their meals remained **unchanged.** Professor Fry, the Dean of the University, said they were inspired by "nudge theory", according to which people usually order the dishes that are listed first.		
"We have a **moral obligation** to ensure the students in our care eat healthily. Rearranging the dishes in the menus had a big effect on what students ordered" said the Dean.	"We have a **moral obligation** to ensure the students in our care eat healthily. Rearranging the dishes in the menus had a big effect on what students ordered" said the Dean.	"Rearranging the dishes in the menus had a big effect on what students ordered" said the Dean.
	The dishes listed first on the new menus **cost the University less to provide**. Therefore, the University was also able to **reduce operating costs on campus.**	

support, $t(206) = -3.05$, $p = 0.001$, $d = -0.42$, 95% CI [−0.70, −0.15], $BF_{-0} = 22.233$.

But just as in Experiment 1a, there was no evidence of strong tainting (H3). For every measure, the initiative was judged *more* positively in the Social + Profit condition than in the Profit condition, contrary to the strong tainted nudge prediction: prosociality, $t(205) = 12.13$, $p < 0.001$, $d = 1.69$, 95% CI [1.37, 2.00], $BF_{10} > 100$; morality, $t(205) = 4.60$, $p < 0.001$, $d = 0.64$, 95% CI [0.36, 0.92], $BF_{10} > 100$; support, $t(205) = 4.93$, $p < 0.001$, $d = 0.69$, 95% CI [0.40, 0.97], $BF_{10} > 100$, all two-sided tests. That is, participants evaluated health nudges that earned a secondary benefit of profit more positively than nudges that brought only profit. They appeared to recognise that these were win–win initiatives and judged them accordingly.

4.3. Discussion

While we found consistent support for Hypotheses 1 and 2, we found no evidence that combining a social nudge with profit would ever be so unacceptable to be judged less positively than a purely profitable nudge. "Doing well while doing good" was always judged better than "doing well" alone. So far, these findings are good news for organizations implementing prosocial nudges: using nudging to help clients, and making money in the process, is likely to be seen as better than merely using nudging to make money.

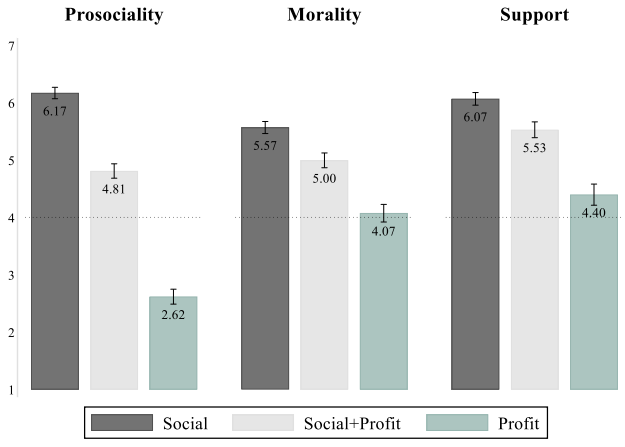


Fig. 3. Mean prosociality, morality and support per condition, Experiment 1b. *Notes*. Error bars indicate standard errors of the mean. Dashed line indicates scale midpoint.

5. Experiments 2a and 2b – The role of deception

In the context of the broader literature, one possible reading of our results is that previous research demonstrating what we would call strong tainted altruism does not generalise to nudges. Before reaching this conclusion, however, we consider if other differences between our setting and that of earlier studies might account for the different results.

In some earlier studies showing strong tainted altruism, the scenarios adopted were ones in which the "altruist" was not forthright about their true motives. For example, in Carlson and Zaki's (2018) Study 2 (Vignette 8), a man was judged negatively for returning a lost smartphone to the owner in anticipation of a cash reward. Perhaps, participants reacted negatively to the ulterior motive. Similarly, in Newman and Cain's (2014) Experiment 3, a company owner who donated a large amount to a children's hospital only to gain publicity and boost his company's reputation was judged more negatively than one who pursued the same goal by investing an equal sum in an extensive advertising campaign. Again, participants may have disliked how the company owner pretended to be charitable. Supporting evidence for this comes from Berman et al. (2015), who suggest that good deeds might backfire if people suspect one only does them to be viewed positively by others. Berman et al. show that bragging about prosociality has a negative influence on perceptions of generosity because it raises scepticism about the agent's true motives. Relatedly, Silver and Silverman (2022) suggest that people are averse to inauthenticity, and if they suspect an actor has ulterior motives for doing good, they discount their good actions.

We conjectured that we may not have observed strong tainted nudge because our scenarios made no suggestion that the nudging organization had pretended they meant to benefit the public while their true goal was to make profit. We therefore hypothesised that, if a nudger who pursues profit deceptively claims to be motivated by a prosocial goal, strong tainting would be more likely.

This hypothesis is in line with research showing that a crucial dimension of the effectiveness of win–win initiatives is public perception of how honestly motives are communicated (e.g., Alempaki et al., 2020). Individuals are often sceptical about why an organization might choose a particular prosocial activity. A leading example of this scepticism is customers' response to hotels asking them to recycle towels for the good of the environment, when this also helps the hotel save money. Chen et al. (2019) found that people are highly sceptical if this is the hotel's only "green" activity and assume it is greenwashing (i.e., the promotion

of environmentally friendly activities with a hidden purpose of increasing profits; see, e.g., Becker-Olsen and Potucek, 2013; Orange and Cohen, 2010). As Chen et al. say "When hotels employ these two practices highlighting the plight of the environment and hiding the underlying financial benefit, many consumers would naturally feel that hoteliers are being dishonest and deceptive. As a result, they might not trust the hotel" (p. 328). This argument would suggest that strong tainted nudge would be particularly likely when the nudger deceives the public about their true profit-motivated initiatives.

Thus, combining profit with prosocial nudging might backfire when profit is the true motive but a prosocial motive is *deceptively* claimed. If, for instance, people believe that a hotel's towel reuse policy is profit-motivated despite its appeal to environmental concerns, they are likely to judge the hotel negatively. Similarly, in our experiments, if an organization earns profit from giving their clients smaller meals but justifies that decision by claiming they did it for the clients' own good, then we might observe a strong tainted nudge effect. We investigated this possibility in Experiments 2a and 2b.

These experiments were based on variants of the university scenarios used in Experiments 1a and 1b. In both experiments, the university was always motivated by profit. The Profit condition was just as in the earlier experiments.

In the *deceptive* Social^{d} + Profit condition, the profit that motivated the initiative was accompanied by a prosocial benefit, which the organization deceptively claimed was the true motive. Our predictions are summarized in Fig. 4 (a reduced version of Fig. 1). Hypothesis 3^{d}, strong tainted nudge, would be supported if, for any of the dependent variables of interest, Social^{d} + Profit < Profit.

5.1. Experiment 2a

5.1.1. Method

Participants were 205 UK residents (58 men, 145 women, 2 gender nondisclosed, M_{age} = 33.9, SD = 11.4, 2 age nondisclosed) recruited from Prolific in exchange for £0.50 (average duration almost five minutes).

They were allocated to one of two conditions that manipulated the declared motivation of a change in the default meal. In both conditions, the actual motivation of the university was to increase the salaries of senior administrators. In the Profit condition, the university acknowledged their true motive and justified their actions by the need to use resources efficiently. In the Social^{d} + Profit condition, participants were told that the meals "happened" to be healthier, and that the university had (deceptively) claimed to be motivated by the concern that their students ate healthily. The scenarios are shown in Table 3 (Column 1).

It is important to note that the deception was outright, and participants were explicitly informed about it. We did this to create the most favourable conditions for strong tainting to emerge. For the same reason, we informed participants that the university was "widely praised" for their efforts, to highlight that the deception was successful, and the organization received undeserved credit for promoting healthy eating.

After reading the article, participants evaluated the nudge using the same items as in Experiments 1a and 1b to assess prosociality (*Spearman's* $\rho = 0.31$, $p < 0.001$)[f], morality (*Spearman's* $\rho = 0.79$, $p < 0.001$), and support (*Spearman's* $\rho = 0.88$, $p < 0.001$). They also answered two further questions (interspersed in random order amongst the other questions) asking how deceptive and honest they thought the initiative was. These answers were averaged to create a measure of *honesty* (*Spearman's* $\rho = 0.65$, $p < 0.001$).

[f] Because the prosociality measure shows a relatively low intercorrelation in the experiments involving deception, we report a disaggregated analysis in the Supplementary Material S1.

5.1.2. Results

The results of Experiment 2a are depicted in Fig. 5.

We begin with honesty. Participants evaluated the initiative as much less honest in the Social^{d} + Profit condition than in the Profit condition, $t(203) = -5.97$, $p < 0.001$, $d = -0.83$, 95% CI [−1.12, −0.55], $BF_{-0} > 100$. This is to be expected, since the organization was explicitly deceptive in the Social^{d} + Profit condition.

A noticeable effect of the deception manipulation is that all ratings were lower than their counterparts in Experiments 1a and 1b. Beyond this general trend, there was strong tainting only for prosociality, which was lower in the Social^{d} + Profit condition than in the Profit condition, $t(203) = -1.75$, $p = 0.041$, $d = -0.24$, 95% CI [−0.52, 0.03], $BF_{-0} = 1.213$. There was no significant difference between conditions in judged morality, $t(203) = -1.39$, $p = 0.083$, $d = -0.19$, 95% CI [−0.47, 0.08], $BF_{-0} = 0.684$. Note that this differs from Experiments 1a and 1b, in which both morality and prosociality were considerably higher in the (no deception) Social + Profit condition compared to the Profit condition.

Despite the deception, however, people were *more* supportive of the Social^{d} + Profit nudge than the Profit one, $t(203) = 2.06$, $p = 0.041$, $d = 0.29$, 95% CI [0.01, 0.56], $BF_{10} = 1.088$, two-sided test. This suggests participants responded positively to the prosocial benefit, even though it was not the true motive for the initiative while the organization had deceptively claimed it was.

5.2. Experiment 2b

5.2.1. Method

We recruited 205 UK residents (68 men, 131 women, 6 gender nondisclosed, M_{age} = 34.4, SD = 11.6, 1 age nondisclosed) from Prolific in exchange for £0.50 (average duration almost four minutes).

Participants were allocated to one of the two conditions depicted in Table 3 (Column 2). We manipulated whether the menu rearrangement was openly stated as an attempt to use resources efficiently (Profit) or deceptively disguised behind the accidental health benefits as an attempt to ensure students ate more healthily (Social^{d} + Profit).

Participants rated the initiative using the same eight items used in Experiment 2a (adjusted for the different nudge) to assess honesty (*Spearman's* $\rho = 0.63$, $p < 0.001$), prosociality (*Spearman's* $\rho = 0.49$, $p < 0.001$), morality (*Spearman's* $\rho = 0.63$, $p < 0.001$), and support (*Spearman's* $\rho = 0.79$, $p < 0.001$).

5.2.2. Results

The results, shown in Fig. 6, exhibit similar patterns as in Experiment 2a. As expected, participants evaluated the initiative as less honest in the Social^{d} + Profit than in the Profit condition, $t(203) = -4.87$, $p < 0.001$, $d = -0.68$, 95% CI [−0.96, −0.40], $BF_{-0} > 100$. Of the remaining measures, only morality showed a strong tainted nudge effect, $t(203) = -1.78$, $p = 0.038$, $d = -0.25$, 95% CI [−0.52, 0.03], $BF_{-0} = 1.277$. There was no effect for prosociality, $t(203) = -0.72$, $p = 0.235$, $d = -0.10$, 95% CI [−0.37, 0.17], $BF_{-0} = 0.295$, nor for support, $t(203) = 0.44$, $p = 0.668$, $d = 0.06$, 95% CI [−0.21, 0.33], $BF_{-0} = 0.112$.

5.3. Discussion

Experiments 2a and 2b suggest that, even in the face of outright dishonesty by the nudger, there is no consistent strong tainted nudge effect. Out of the three measures of tainting in two experiments, one in each experiment showed evidence of strong tainted nudge (prosociality in Experiment 2a, morality in Experiment 2b), and one showed the opposite effect in one experiment (support in Experiment 2a). Nonetheless, we cannot say that deception is costless, since in contrast with Experiments 1a and 1b, this time the nudge with a social benefit was not evaluated more positively than the nudge without one. The effects of that social benefit were largely reversed by the presence of deception.

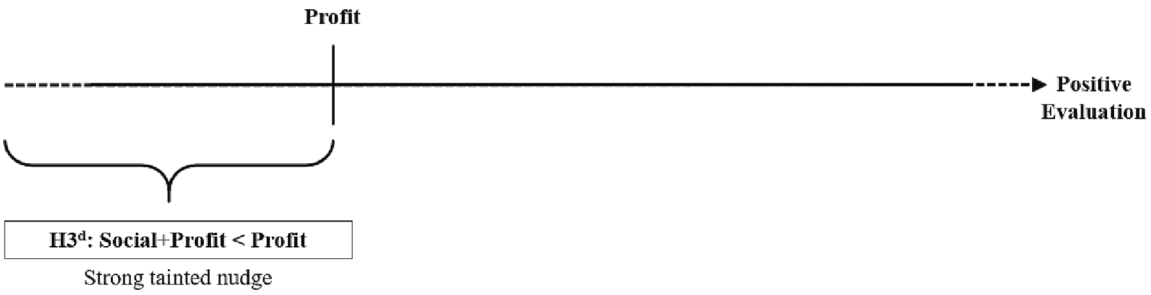


Fig. 4. Hypothesis 3d.
Notes. The right facing arrow indicates the positive evaluation assigned to an initiative. The braces denote the range of evaluations consistent with the hypothesis.

Table 3
Scenarios used in Experiments 2a and 2b.

Wessex Evening Standard			
Experiment 2a		*Experiment 2b*	
New "default" meals at Wessex Uni		**New menu layouts at Wessex Uni**	
Residential students at the University of Wessex pay in advance for the hot evening meals they receive during term time. Each evening there is a "default" meal, which students get unless they fill out a form requesting something different one week in advance. Last September, the University launched an initiative to **increase the salaries they can offer to their Senior Administrators**. They replaced the default meals with ones that **cost the University significantly less to provide, but are just as tasty and nutritious.**		Residential students at the University of Wessex pay in advance for the hot evening meals they receive during term time. Students are entitled to choose their meals from the large daily menus displayed in each of the University restaurants. Last September, the University launched an initiative to **reduce operating costs on campus.** They rearranged the menus so that the dishes that **cost the University less to provide were listed first**.	
The annual amount paid by students for their meals remained **unchanged.**			
Social^{d} + *Profit:* The new default meals happen to be **healthier, lower in fat, sugar and calories.** Because of this, Professor Fry, the Dean of the University, said: "We have a **moral obligation** to ensure that the students in our care eat healthily.	*Profit:* Professor Fry, the Dean of the University, said: "We have a **moral obligation** to ensure that our resources are used efficiently.	*Social*^{d} + *Profit:* In the new menus, the dishes listed first happen to be **healthier, lower in fat, sugar and calories**. Because of this, Professor Fry, the Dean of the University, said: "We have a **moral obligation** to ensure that the students in our care eat healthily.	*Profit:* Professor Fry, the Dean of the University, said: "We have a **moral obligation** to ensure that our resources are used efficiently.
We were inspired by 'nudge theory', according to which people usually stick with default options."		We were inspired by 'nudge theory', according to which people usually order the dishes that are listed first."	
As a result, the University has been widely praised for			
caring about healthy eating, but, in fact, their **only objective** was to **increase the salaries they can offer to their Senior Administrators.**	their attempts to **increase the salaries they can offer to their Senior Administrators.**	caring about healthy eating, but, in fact, their **only objective** was to **reduce operating costs.**	their attempts to **reduce operating costs.**

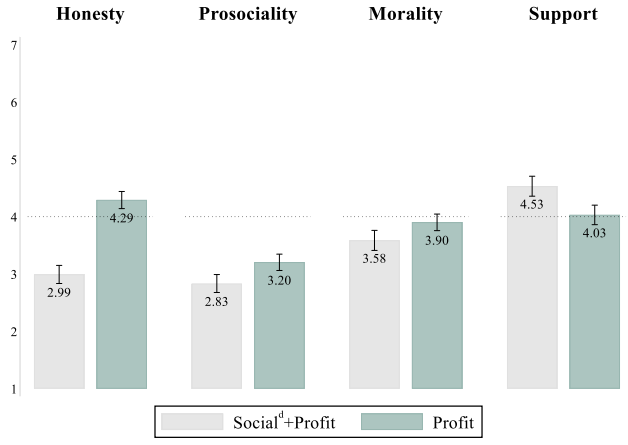


Fig. 5. Mean honesty, prosociality, morality and support per condition, Experiment 2a.
Notes. Error bars indicate standard errors of the mean. Dashed line indicates scale midpoint.

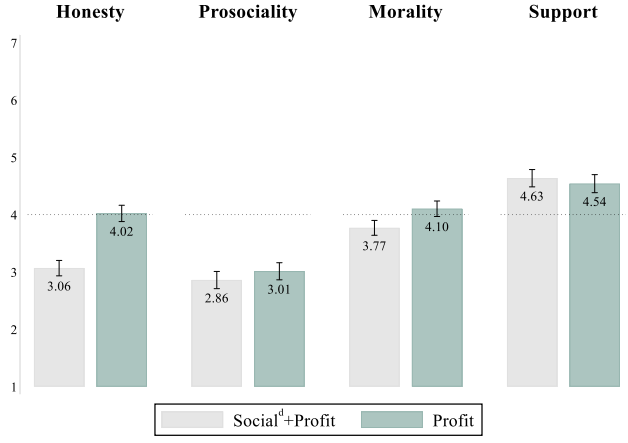


Fig. 6. Mean honesty, prosociality, morality and support, Experiment 2b.
Notes. Error bars indicate standard errors of the mean. Dashed line indicates scale midpoint.

6. Experiment 3 – For-profit versus non-profit organizations

Taken all together, the previous experiments suggest that strong tainting of nudges, if it occurs, will occur to a limited degree and will probably require a boost, such as deception on the part of the organization, beyond merely earning profit from one's good deeds. In

Table 4
Scenarios used in Experiment 3.

Wessex Evening Standard

Non-profit sector		*For-profit sector*	
New "default" meals at South Wessex Health Trust		**New "default" meals at McCarthy Healthcare Ltd.**	
South Wessex Health Trust is a public non-profit organization that manages a number of taxpayer-funded hospitals in Wessex County. The trust offers its employees a pre-paid canteen service, with hot meals served during the afternoon and evening breaks. Employees pay monthly in advance via a salary deduction. Each day there is a 'default' meal, which they are served unless they fill out a form requesting something different one week in advance. Last September, the trust launched an initiative to **increase the salaries of their Senior Executives**. Part of this initiative is to replace the default meals with ones that **cost them significantly less to provide, but are just as tasty and nutritious**.		McCarthy Healthcare Ltd. is a private for-profit organization that owns a number of private hospitals in Wessex. The company offers its employees a pre-paid canteen service, with hot meals served during the afternoon and evening breaks. Employees pay monthly in advance via a salary deduction. Each day there is a 'default' meal, which they are served unless they fill out a form requesting something different one week in advance. Last September, the company launched an initiative to **increase the salaries of their Senior Executives**. Part of this initiative is to replace the default meals with ones that **cost them significantly less to provide, but are just as tasty and nutritious**.	
The annual amount paid by employees for their meals remained **unchanged**.			
Social[d] + *Profit:* The new default meals happen to be **healthier, lower in fat, sugar and calories**. Because of this, Elliot McCarthy, the trust's Director, said: "We have an obligation to ensure our employees eat healthily. We were inspired by 'nudge theory', according to which people usually stick with default options." The trust was widely praised for caring about healthy eating, but, in fact, their only objective was to **offer a generous remuneration package to their Senior Executives.**	*Profit:* Elliot McCarthy, the trust's Director, said: "We have an obligation to ensure that our resources are used efficiently. We were inspired by 'nudge theory', according to which people usually stick with default options." The trust was widely praised for the **generous remuneration package offered to their Senior Executives.**	*Social*[d] + *Profit:* The new default meals happen to be **healthier, lower in fat, sugar and calories**. Because of this, Elliot McCarthy, the company's CEO, said: "We have an obligation to ensure our employees eat healthily. We were inspired by 'nudge theory', according to which people usually stick with default options." The company was widely praised for caring about healthy eating, but, in fact, their only objective was to **offer a generous remuneration package to their Senior Executives.**	*Profit:* Elliot McCarthy, the company's CEO, said: "We have an obligation to ensure that our resources are used efficiently. We were inspired by 'nudge theory', according to which people usually stick with default options." The company was widely praised for the **generous remuneration package offered to their Senior Executives**.

Experiment 3, we investigated the possibility that this conclusion depends on the profit orientation of the nudging organization. In the UK, the country of residence of all our experimental participants, universities have both a public and a private face (e.g., Longden and Bélanger, 2013). They receive significant public funding and are subject to strict regulation. Yet, they are also revenue earning and seek to earn a profit. But this profit does not go to shareholders; it is used to run the university and invest in its future. Because we selected practices that could plausibly apply to both non-profit and for-profit organizations, we do not know whether participants treated the university as a non-profit or a commercial enterprise, and if that would make a difference to the possible tainting of nudges.

Some evidence provided by Greitemeyer and Sagioglou (2018) suggests that the profit orientation of the nudging organisation might indeed matter. They found that the use of a psychological persuasion technique (the "foot in the door" technique) to increase charitable donations is judged as permissible for a for-profit organization but not for a non-profit one. They explain this using expectancy violations theory (Burgoon, 1993, 2009), arguing that people are harsher in their judgements of organizations in the non-profit sector because the use of manipulative techniques violates their expectations of how such organizations should act. A similar suggestion comes from Hornsey et al. (2021), who show that non-profits are punished more severely when they violate consumer trust, because the public holds them to higher standards than for-profit entities (a "moral disillusionment effect").

This devaluation is also in line with the moral hypocrisy literature (e.g., Batson et al., 2007) showing that people dislike and punish those who betray moral standards (e.g., Barden et al., 2005; Laurent et al., 2014). Since people believe that, relative to their for-profit counterparts, non-profit organizations are in principle more moral (e.g., Smith and Richmond, 2007), more strongly associated with public-serving than self-serving behavior (e.g., Blodgett and Melconian, 2012; Handy et al., 2010; Szykman et al., 2004), and more trusted on social issues (e.g., Wootliff and Deri, 2001), they might judge non-profit organizations more harshly if they fail to live up to those higher moral standards. One such asymmetry is reported by Lin-Hi et al. (2015), who find that bad actions (negative CSR) decrease the judged trustworthiness of non-profit organizations much more than that of for-profit ones. There could be similar effects in the case of profitable prosocial nudges, in that a strong tainted nudge effect might arise primarily or only for non-profit organizations that use deception to disguise selfish motives.

An alternative prediction comes from moral licensing theory (e.g., Effron and Monin, 2010; Merritt et al., 2010), according to which individuals are sometimes willing to tolerate moral transgressions because of individuals' prior good deeds (i.e., prior good deeds are a license for subsequent wrongdoing). If non-profits typically possess moral credits because of good deeds that benefit society, people might be forgiving about their misdeeds, even in cases that involve deception. In line with this, Hornsey et al. (2021) suggested the better reputation of for-profits might serve as a "trust bank" that offers an insurance-like protection during difficult times. For example, if an organization has a good prior reputation, it will be less likely to suffer from negative publicity after a scandal (Decker, 2012). In that case, non-profit organizations that deceive may maintain their moral reputation, while for-profit organizations more often associated with greedy behavior may incur a strong tainted nudge effect.

We investigated this issue in Experiment 3, in which we contrasted nudges applied by organizations in the for-profit and non-profit sector. Given that moral credentials are more likely to be questioned when an organization uses fraudulent practices, we tested for strong tainted nudge when the profit is combined with a prosocial outcome, and the organization deceptively claims this was their actual motivation. However, instead of a university nudging their students, our scenarios involve a hospital nudging their employees. Hospitals in the UK can be both public and non-profit or private and for-profit, and therefore we can hold constant the type of organization, while varying its profit orientation. We know people hold different sector-specific stereotypes and, in particular, consider non-profit hospitals as more trustworthy, fair, and humane (e.g., Schlesinger et al., 2004).

6.1. Method

Participants were 415 UK residents (145 men, 263 women, 7 gender nondisclosed, M_{age} = 35.5, SD = 12.3, 3 age nondisclosed) recruited

from Prolific in exchange for £0.50 (average duration slightly above four minutes).

Participants learned about a hospital that replaced the default meals their employees receive during the breaks. We contrasted a public/non-profit hospital with a private/for-profit one. The exact scenarios used are shown in Table 4.

After reading the article, participants rated the initiative using the same eight items as in Experiment 2a to assess honesty (*Spearman's* $\rho = 0.63$, $p < 0.001$), prosociality (*Spearman's* $\rho = 0.56$, $p < 0.001$), morality (*Spearman's* $\rho = 0.80$, $p < 0.001$), and support (*Spearman's* $\rho = 0.80$, $p < 0.001$). To assess whether differences in expectations might drive any domain specific tainted nudge effect, we added two extra questions, adapted from Greitemeyer and Sagioglou (2018), to assess whether this type of initiative was expected and whether it matched the image participants had of the respective organization. Responses to these two items were highly correlated and were combined to create a single measure of *expectancy* (*Spearman's* $\rho = 0.58$, $p < 0.001$).

6.2. Results

The results are reported in Fig. 7 (panel A for the non-profit and panel B for the for-profit sector).

The results are remarkably similar in the two settings. As in Experiments 2a and 2b, we start by looking at honesty evaluations: as expected, honesty was significantly lower in the Social^{d} + Profit compared to the Profit condition for both settings: non-profit sector, $t(203) = -1.75$, $p = 0.041$, $d = -0.24$, 95% CI [−0.52, 0.03], $BF_{-0} = 1.213$; for-profit sector, $t(208) = -3.86$, $p < 0.001$, $d = -0.53$, 95% CI [−0.81,

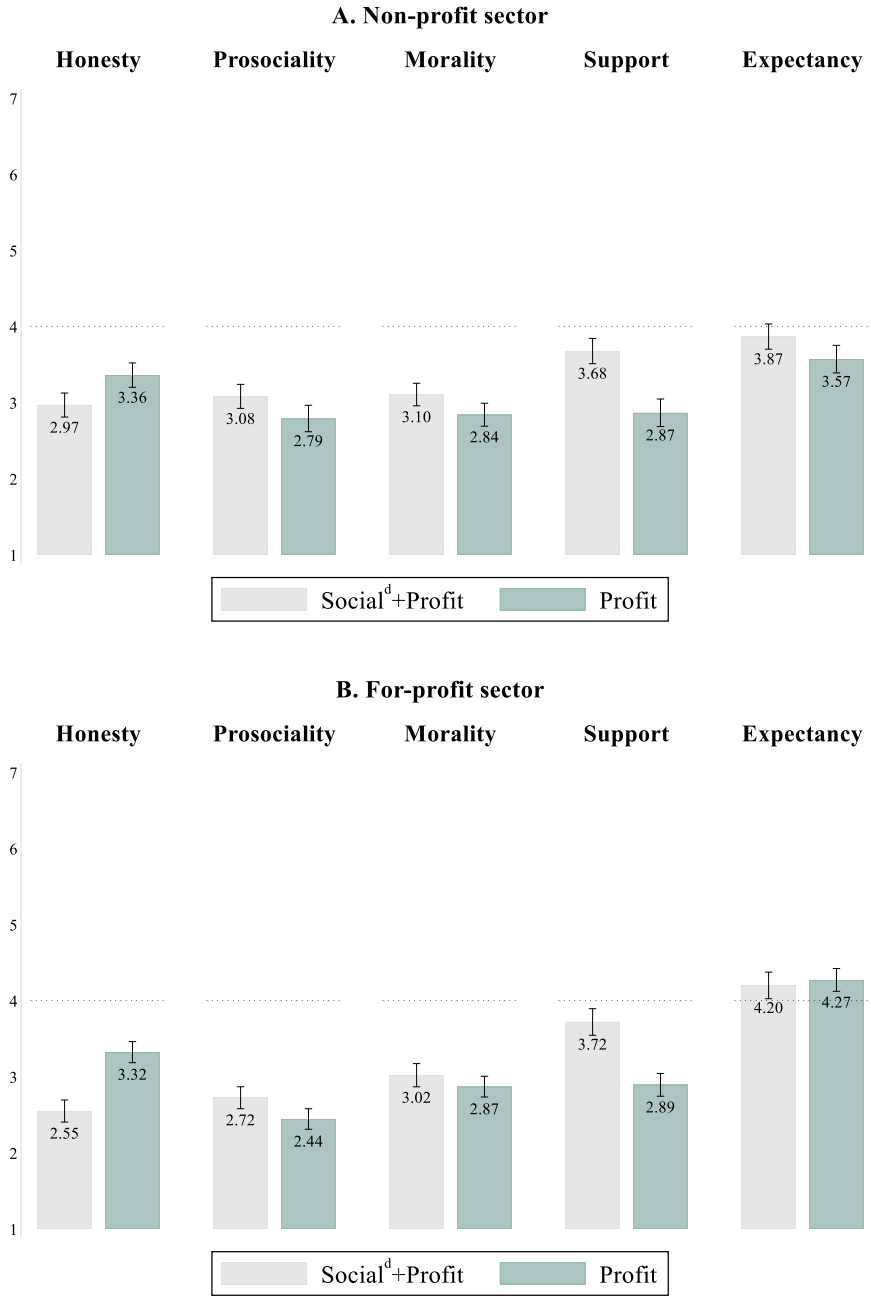


Fig. 7. Mean honesty, prosociality, morality, support, and expectancy per condition, Experiment 3.
Notes. Error bars indicate standard errors of the mean. Dashed line indicates scale midpoint.

−0.26], $BF_{-0} > 100$. We did not observe strong tainting in any of the other variables of interest. More specifically, participants provided similar evaluations of prosociality, morality and expectancy, independently of whether profit motives were disguised or not: non-profit sector, prosociality, $t(203) = 1.24$, $p = 0.892$, $d = 0.17$, 95% CI [−0.10, 0.45], $BF_{-0} = 0.072$; morality, $t(203) = 1.25$, $p = 0.894$, $d = 0.18$, 95% CI [−0.10, 0.45], $BF_{-0} = 0.072$; expectancy, $t(203) = 1.38$, $p = 0.916$, $d = 0.19$, 95% CI [−0.08, 0.47], $BF_{-0} = 0.068$; for-profit sector, prosociality, $t(208) = 1.42$, $p = 0.921$, $d = 0.20$, 95% CI [−0.08, 0.47], $BF_{-0} = 0.066$; morality, $t(208) = 0.74$, $p = 0.769$, $d = 0.10$, 95% CI [−0.17, 0.37], $BF_{-0} = 0.092$; expectancy, $t(208) = -0.33$, $p = 0.369$, $d = -0.05$, 95% CI [−0.32, 0.22], $BF_{-0} = 0.199$.

Finally, in both settings, in line with Experiment 2a, participants were significantly *more* willing to support the initiative in the Social[d] + Profit compared to the Profit condition: non-profit sector, support, $t(203) = 3.31$, $p = 0.001$, $d = 0.46$, 95% CI [0.18, 0.74], $BF_{10} = 23.413$; for-profit sector, support, $t(208) = 3.60$, $p < 0.001$, $d = 0.50$, 95% CI [0.22, 0.77], $BF_{10} = 58.452$, two-sided tests.

6.3. Discussion

Despite the change of setting and the different nature of the organizations involved, the results of Experiment 3 broadly align with those of Experiments 2a and 2b. There is little evidence of strong tainting and, when it comes to support, people appear to recognize the welfare benefits of win–win nudges and are forgiving even about deception when this is accompanied by a social benefit. This holds independently of the organization's profit orientation. Nonetheless, again as in Experiments 2a and 2b, the effect of deception appeared to eliminate many of the advantages of doing good.

7. Experiments 4a and 4b – Tainted altruism versus tainted nudge

Taken together, Experiments 1a-3 suggest that profit can taint prosocial nudges, but to a much lesser degree than previously demonstrated in the tainted altruism literature. Our work so far did not evaluate tainting in an altruism context, but rather focused exclusively on nudging. As such, our data do not allow us to address whether the tainting effect simply does not generalize to nudges, or is just not particularly robust. In Experiments 4a and 4b we investigated whether organizations are more susceptible to tainting when they combine profit with altruistic acts compared to combining profit with prosocial nudging.

Furthermore, Experiments 1a-3 focused exclusively on the health domain, where the Social + Profit condition involved prosocial benefits that improved the nudgees' health, while improving the organization's bottom line. The profit orientation of the organization is less salient when they design an intervention to improve, for instance, their customers' health or their environmental impact. Participants' evaluations, therefore, might have been affected by the comparison of benefits and the presence of economic incentives only on the nudger's side. It could be the case that tainting is weaker, and people are more forgiving, if both the nudger and the nudgee make money. To investigate this possibility, in Experiments 4a and 4b, we examined whether similar tainting effects occur in the financial domain, where both the nudger's and the nudgees' benefits are monetary.

In both new experiments, we searched for evidence of tainted nudge in relation to the initiative of a financial management company, a clearly profit-oriented form of enterprise in which both parties stand to benefit financially. Since investors make many costly and predictable mistakes (e.g., Barber et al., 2021; Subrahmanyam, 2008), a financial management company can use nudging to improve their clients' financial outcomes, and these improvements may or may not lead to increased profits for the company. This setting has also the advantage of admitting altruistic initiatives. We were, therefore, able to test for tainted altruism as well, by considering how people respond to a charitable initiative made by the same financial management company.

The materials were designed to parallel the RED initiative by the GAP clothing chain that featured in a study described by Newman and Cain (2014, Experiment 4), but we used a fictional financial management company instead. As in the GAP study, there was no bolding in the text.

Given the change of setting, we designed Experiments 4a and 4b so that we could test for both tainting and strong tainting (Hypotheses 2 and 3). The two experiments, which did not include deception, differed in the salience of the profit information and the content of the comprehension checks.

7.1. Experiment 4a

7.1.1. Method

Participants were 1223 UK residents recruited from Prolific in exchange for £0.50. We excluded 8 participants who failed comprehension and attention checks, leaving us with 1215 valid responses (599 men, 600 women, 13 non-binary, 3 gender nondisclosed, $M_{age} = 41.1$, SD = 13.7, 5 age nondisclosed). The experiment took a little over four minutes on average.

Participants were randomly assigned to one of six conditions in a 2 (Altruism versus Nudging) by 3 (Social versus Profit versus Social + Profit) design. In all conditions, they learned about an investment company that launched a new retirement planning initiative. As in Experiments 1a and 1b, the Social condition described a successful prosocially motivated initiative, the Profit condition described a successful profit motivated initiative, and the Social + Profit condition described a successful prosocially motivated initiative that resulted in profits. To contrast the tainting effect of profit in the charitable and the nudging context, we varied whether the initiative involved giving to charity or assigning clients to default investment plans. The complete scenarios are shown in Table 5.

After reading the scenario, participants evaluated the initiative with the same items used in previous experiments to assess morality (*Spearman's* $\rho = 0.83$, $p < 0.001$) and support (*Spearman's* $\rho = 0.83$, $p < 0.001$). We modified the prosociality questions to ask whether the company launched the initiative for prosocial reasons and to benefit society (in the altruism context) or their clients (in the nudge context). These answers were averaged to create a measure of prosociality (*Spearman's* $\rho = 0.62$, $p < 0.001$).

In addition to these core measures, we introduced a measure of *recommendation*. This consisted of two questions asking if participants would recommend the initiative to a friend or relative and whether they would buy an investment product from the company. Responses to these two items were highly correlated, and were combined to create a single recommendation measure (*Spearman's* $\rho = 0.74$, $p < 0.001$). In the context of the hypothetical scenarios used in our experiments, recommendation reflects the extent to which judgements of the initiative translate into behavioral intentions towards the organization, taking us a step closer to identifying potential downstream consequences of tainting. This was a purely exploratory measure, and we advanced no hypotheses about it, since we recognised an organization might be recommended because it earns a profit, because it is a nice prosocial company, or for other reasons.

There were two attention checks and one comprehension check. To identify participants who read the questions carefully, we asked them to select the end point of the Likert-scale (either "Not at all" or "Very much so") in two attention check questions interspersed at random among the other items. To check if participants read the scenario, at the end of the experiment they had to identify the colour that featured in the name of the initiative among a list of four options (BLUE, RED, YELLOW, GREEN) presented in a randomised order.

7.1.2. Results

The results are reported in Fig. 8 (panel A for the charitable context

Table 5
Scenarios used in Experiments 4a and 4b.

Horizon is a UK/French investment company specializing in retirement planning. It was founded in 1974 by Samuel James and Marion Martin. Horizon has branches in ten European countries, with three cities in the UK.					
	Altruism			*Nudge*	
Social: To increase their contribution to society, in 2021 Horizon launched a new retirement planning initiative, which they called BLUE.	*Social + Profit:* To increase their contribution to society, in 2021 Horizon launched a new retirement planning initiative, which they called BLUE.	*Profit:* To increase their profitability, in 2021 Horizon launched a new retirement planning initiative, which they called BLUE.	*Social:* To increase their clients' well-being, in 2021 Horizon launched a new retirement planning initiative, which they called BLUE.	*Social + Profit:* To increase their clients' well-being, in 2021 Horizon launched a new retirement planning initiative, which they called BLUE.	*Profit:* To increase their profitability, in 2021 Horizon launched a new retirement planning initiative, which they called BLUE.
			BLUE was inspired by "nudge theory", which shows that people usually stick with default options. BLUE assigns all new clients to a default plan which will guide their investments. Clients can easily switch to a different plan if they wish.		
50% of the profits they earn from BLUE are donated to charities operating in low-income countries, especially in sub-Saharan Africa. The selected charities are dedicated to promoting sustainable energy, reducing disease, and increasing the availability of education. To date, Horizon have given £130 million to these charities.			BLUE helps clients avoid common investment mistakes, such as making too many trades, or being tempted by "flavour of the week" opportunities. To date, Horizon's clients have collectively earned an estimated £130 million more for their retirements due to the new approach.		
Horizon's BLUE initiative has [also: only for the Social + Profit conditions in both domains] been very profitable for the company, earning them approximately £130 million in additional profits.					

and panel B for the nudge context).

We start by exploring tainting effects in the charitable context. As shown in Fig. 8A, we find significant evidence for Hypothesis 1, as the initiative was evaluated less positively in the Profit condition than in the Social condition on the core dimensions of prosociality, morality and support: prosociality, $t(406) = -17.56$, $p < 0.001$, $d = -1.74$, 95% CI [−1.97, −1.51], $BF_{-0} > 100$; morality, $t(406) = -14.26$, $p < 0.001$, $d = -1.41$, 95% CI [−1.63, −1.19], $BF_{-0} > 100$; support, $t(406) = -10.12$, $p < 0.001$, $d = -1.00$, 95% CI [−1.21, −0.80], $BF_{-0} > 100$.

There was partial evidence for tainted altruism (H2) as the company's initiative was judged as less moral in the Social + Profit condition compared to the Social condition, $t(399) = -1.69$, $p = 0.046$, $d = -0.17$, 95% CI [−0.37, 0.03], $BF_{-0} = 0.837$. However, we did not find evidence for tainted altruism in prosociality, $t(399) = -1.26$, $p = 0.103$, $d = -0.13$, 95% CI [−0.32, 0.07], $BF_{-0} = 0.427$ or support, $t(399) = -0.415$, $p = 0.339$, $d = -0.04$, 95% CI [−0.24, 0.15], $BF_{-0} = 0.158$.

Crucially, there was no evidence of strong tainting (H3). For every measure, the initiative was judged *more* positively in the Social + Profit condition than in the Profit condition: prosociality, $t(409) = 16.79$, $p < 0.001$, $d = 1.66$, 95% CI [1.43, 1.88], $BF_{10} > 100$; morality, $t(409) = 13.17$, $p < 0.001$, $d = 1.30$, 95% CI [1.09, 1.51], $BF_{10} > 100$; support, $t(409) = 9.96$, $p < 0.001$, $d = 0.98$, 95% CI [0.78, 1.19], $BF_{10} > 100$, all two-sided tests. Participants evaluated charitable initiatives that also brought profit more positively than initiatives that brought only profit.

We found evidence that judgements of the initiative translate into behavioral intentions towards the company: the recommendation measure was higher in both the Social and the Social + Profit conditions compared to the Profit condition: Social versus Profit, $t(406) = -5.38$, $p < 0.001$, $d = -0.53$, 95% CI [−0.73, −0.33], $BF_{10} = 86.407$; Social + Profit versus Profit, $t(409) = 7.09$, $p < 0.001$, $d = 0.70$, 95% CI [0.50, 0.90], $BF_{10} > 100$. The Social and Social + Profit conditions did not differ, $t(399) = 1.63$, $p = 0.105$, $d = 0.16$, 95% CI [−0.03, 0.36], $BF_{10} = 0.396$, all two-sided tests.

Eyeballing Fig. 8B, it is evident that the results in the nudge context are very similar to those in the charitable context. In line with Hypothesis 1, the Profit nudge was evaluated less favourably than the Social nudge on all dimensions: prosociality, $t(403) = -14.03$, $p < 0.001$, $d = -1.39$, 95% CI [−1.61, −1.18], $BF_{-0} > 100$; morality, $t(403) = -12.12$, $p < 0.001$, $d = -1.20$, 95% CI [−1.42, −0.99], $BF_{-0} > 100$; and support, $t(403) = -10.05$, $p < 0.001$, $d = -1.00$, 95% CI [−1.20, −0.79], $BF_{-0} > 100$.

When it comes to tainting, only prosociality showed a significant difference between the Social + Profit and the Social conditions in the direction predicted by Hypothesis 2, $t(401) = -2.40$, $p = 0.008$, $d = -0.24$, 95% CI [−0.44, −0.04], $BF_{-0} = 3.499$. There were no differences for morality, $t(401) = -0.29$, $p = 0.385$, $d = -0.03$, 95% CI [−0.22, 0.17], $BF_{-0} = 0.141$; or support, $t(401) = 0.92$, $p = 0.821$, $d = 0.09$, 95% CI [−0.10, 0.29], $BF_{-0} = 0.061$.

There was no strong tainting (H3). For every dimension, the initiative was rated significantly *more* positively in the Social + Profit condition than in the Profit condition: prosociality, $t(400) = 11.75$, $p < 0.001$, $d = 1.17$, 95% CI [0.96, 1.38], $BF_{10} > 100$; morality, $t(400) = 12.09$, $p < 0.001$, $d = 1.21$, 95% CI [0.99, 1.42], $BF_{10} > 100$; support, $t(400) = 11.03$, $p < 0.001$, $d = 1.10$, 95% CI [0.89, 1.31], $BF_{10} > 100$, all two-sided tests.

As in the charitable context, participants were more likely to recommend the organization in the Social and the Social + Profit conditions compared to the Profit condition: Social versus Profit, $t(403) = -9.76$, $p < 0.001$, $d = -0.97$, 95% CI [−1.18, −0.76], $BF_{10} > 100$; Social + Profit versus Profit, $t(400) = 10.57$, $p < 0.001$, $d = 1.05$, 95% CI [0.85, 1.26], $BF_{10} > 100$. There was no difference in the recommendation measure between the Social + Profit and the Social conditions: $t(401) = 0.96$, $p = 0.337$, $d = 0.10$, 95% CI [−0.10, 0.29], $BF_{10} = 0.172$, all two-sided tests.

7.1.3. Discussion

Overall, the evidence for tainted effects in the financial domain was limited and restricted to a single evaluative dimension (morality for charitable actions and prosociality for nudges). It is possible that the lack of tainting was due to the comparatively low salience of the profit information since, to match the GAP study we were modelling, we did not highlight this information in bold as we did in all our previous experiments. While our attention checks ensured the questions were read by participants, our comprehension check focused on one aspect of our scenario – the name of the initiative – which appeared in most of the evaluation questions. Consequently, in Experiment 4b we increased the salience of the profit information by putting it on a separate page of the survey. We also included an additional comprehension check that allowed us to exclude participants who did not pay sufficient attention to the profit information. These exclusion criteria were part of the preregistered analysis plan.

7.2. Experiment 4b

7.2.1. Method

Participants were 1202 UK residents recruited from Prolific in

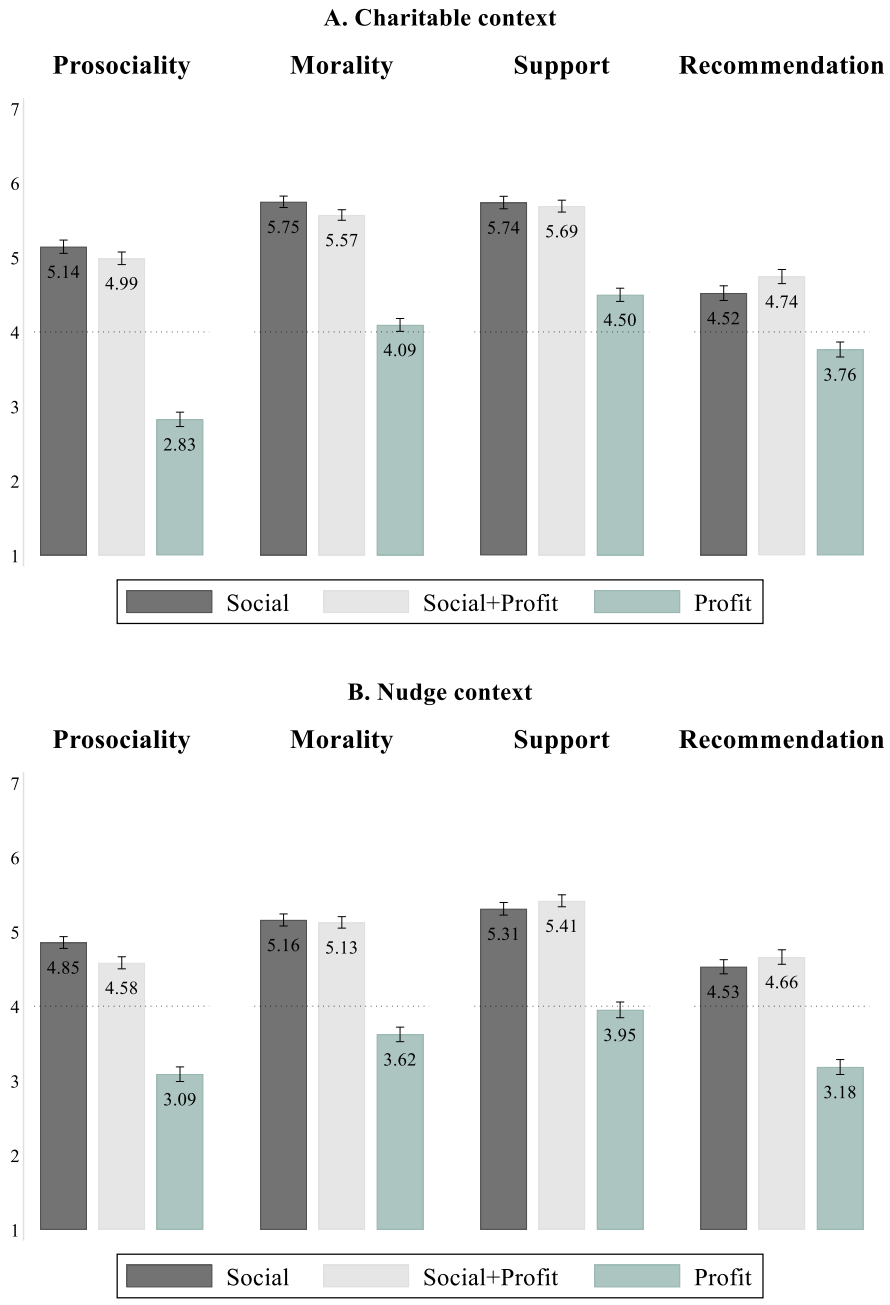


Fig. 8. Mean prosociality, morality, support, and recommendation per condition, Experiment 4a.
Notes. Error bars indicate standard errors of the mean. Dashed line indicates scale midpoint.

exchange for £0.50. We excluded 80 participants who failed any one of the attention or comprehension checks, leaving us with 1122 valid responses (559 men, 551 women, 10 non-binary, 2 gender nondisclosed, M_{age} = 41.2, SD = 13.7, 6 age nondisclosed). The experiment took slightly over four minutes on average.

Participants were randomly allocated to one of the six scenarios depicted in Table 5. The only difference between Experiments 4a and 4b is the presentation of the scenario. In the Profit and Social + Profit conditions of Experiment 4b, the profit information was shown on a separate screen after participants had read the information about the initiative. Experiment 4b also included a comprehension question asking participants to indicate, depending on the condition, the amount of money the investment company made for themselves, their clients or the charity. If participants entered the correct response (especially) in the Social + Profit condition, we could assume with confidence they understood the investment initiative brought profit to the company. We excluded participants who gave an incorrect answer.

After reading the scenario, participants rated the initiative using the same eight items as in Experiment 4a to assess prosociality (*Spearman's* $\rho = 0.60$, $p < 0.001$), morality (*Spearman's* $\rho = 0.78$, $p < 0.001$), support (*Spearman's* $\rho = 0.81$, $p < 0.001$) and recommendation (*Spearman's* $\rho = 0.72$, $p < 0.001$).

7.2.2. Results

The results are presented in Fig. 9 (panel A for the charitable context and panel B for the nudge context).

We start again with the charitable context. In line with Hypothesis 1, the initiative was evaluated less favourably in the Profit condition than

in the Social condition on the three core dimensions: prosociality, $t(362) = -18.69$, $p < 0.001$, $d = -1.96$, 95% CI [−2.21, −1.71], $BF_{-0} > 100$; morality, $t(362) = -16.41$, $p < 0.001$, $d = -1.72$, 95% CI [−1.96, −1.48], $BF_{-0} > 100$; support, $t(362) = -11.96$, $p < 0.001$, $d = -1.25$, 95% CI [−1.48, −1.03], $BF_{-0} > 100$.

The more salient profit information revealed stronger evidence of tainted altruism. The initiative was evaluated less favourably in the Social + Profit than in the Social condition on all three dimensions: prosociality, $t(378) = -2.55$, $p = 0.006$, $d = -0.26$, 95% CI [−0.46, −0.06], $BF_{-0} = 5.076$; morality, $t(378) = -3.37$, $p < 0.001$, $d = -0.35$, 95% CI [−0.55, −0.14], $BF_{-0} = 50.009$; support, $t(378) = -1.91$, $p = 0.028$, $d = -0.20$, 95% CI [−0.40, 0.01], $BF_{-0} = 1.281$.

In line with previous experiments, we found no strong tainting. Participants rated the initiative *more* favourably in the Social + Profit condition than in the Profit condition on all dimensions: prosociality, $t(378) = 16.32$, $p < 0.001$, $d = 1.68$, 95% CI [1.44, 1.91], $BF_{10} > 100$; morality, $t(378) = 12.61$, $p < 0.001$, $d = 1.29$, 95% CI [1.07, 1.52], $BF_{10} > 100$; support, $t(378) = 10.50$, $p < 0.001$, $d = 1.08$, 95% CI [0.86, 1.29], $BF_{10} > 100$, all two-sided tests.

Participants were more likely to recommend the organization in the Social and the Social + Profit conditions than in the Profit condition: Social versus Profit, $t(362) = -6.26$, $p < 0.001$, $d = -0.66$, 95% CI [−0.87, −0.44], $BF_{10} > 100$; Social + Profit versus Profit, $t(378) = 7.09$, $p < 0.001$, $d = 0.73$, 95% CI [0.52, 0.94], $BF_{10} > 100$. The Social + Profit and Social conditions did not differ significantly: $t(378) = 0.46$, $p = 0.645$, $d = 0.05$, 95% CI [−0.15, 0.25], $BF_{10} = 0.126$, all two-sided tests.

We now turn to the nudge context. As in the altruism context, participants evaluated the initiative less positively in the Profit condition

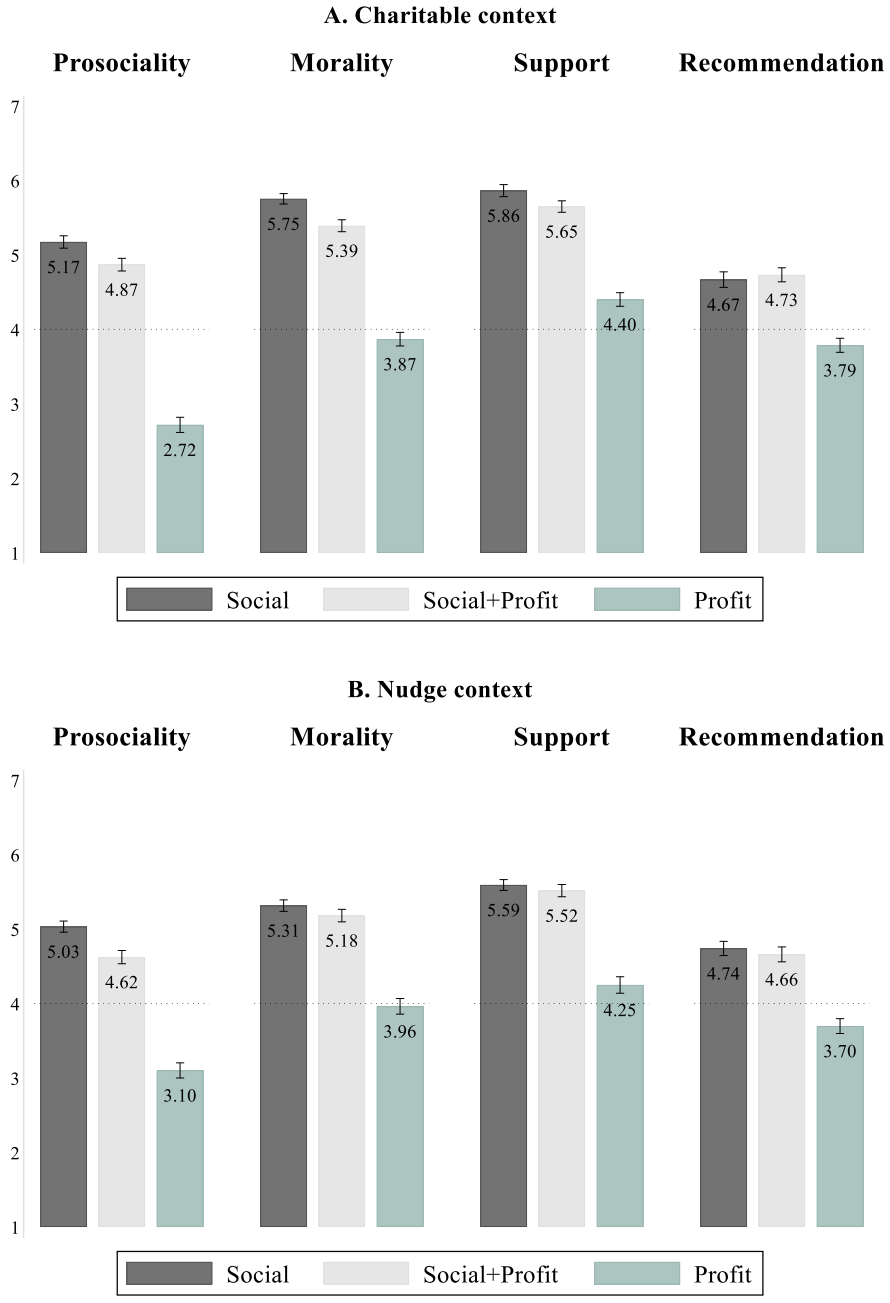


Fig. 9. Mean prosociality, morality, support, and recommendation per condition, Experiment 4b.
Notes. Error bars indicate standard errors of the mean. Dashed line indicates scale midpoint.

than in the Social condition on all dimensions: prosociality, $t(367) = -15.32$, $p < 0.001$, $d = -1.59$, 95% CI [−1.83, −1.36], $BF_{-0} > 100$; morality, $t(367) = -10.41$, $p < 0.001$, $d = -1.08$, 95% CI [−1.30, −0.86], $BF_{-0} > 100$; support, $t(367) = -10.17$, $p < 0.001$, $d = -1.06$, 95% CI [−1.28, −0.84], $BF_{-0} > 100$.

When it comes to tainted nudge, the results are very similar to Experiment 4a, suggesting again that profit did not strongly influence participants' evaluation of nudges. In particular, participants rated the Social + Profit initiative less favourably than the Social one only for prosociality, $t(375) = -3.51$, $p < 0.001$, $d = -0.36$, 95% CI [−0.56, −0.16], $BF_{-0} = 79.667$. Again, there were no differences for morality, $t(375) = -1.17$, $p = 0.121$, $d = -0.12$, 95% CI [−0.32, 0.08], $BF_{-0} = 0.386$; or support, $t(375) = -0.69$, $p = 0.246$, $d = -0.07$, 95% CI [−0.27, 0.13], $BF_{-0} = 0.215$. It is worth noting that, numerically, both morality and support showed differences in the predicted direction, but it appears the tainting effect of profit is smaller for nudging than for altruism.

We found clear evidence against strong tainted nudge, with the Social + Profit initiative rated *more* favourably than the Profit initiative on all dimensions, against Hypothesis 3: prosociality, $t(372) = 11.28$, $p < 0.001$, $d = 1.17$, 95% CI [0.95, 1.39], $BF_{10} > 100$; morality, $t(372) = 9.09$, $p < 0.001$, $d = 0.94$, 95% CI [0.73, 1.15], $BF_{10} > 100$; and support, $t(372) = 9.24$, $p < 0.001$, $d = 0.96$, 95% CI [0.74, 1.17], $BF_{10} > 100$, all two-sided tests.

Finally, participants were more likely to recommend the organization in the Social and the Social + Profit conditions than in the Profit condition: Social versus Profit, $t(367) = -7.56$, $p < 0.001$, $d = -0.79$, 95% CI [−1.00, −0.57], $BF_{10} > 100$; Social + Profit versus Profit, $t(372) = 6.82$, $p < 0.001$, $d = 0.71$, 95% CI [0.50, 0.91], $BF_{10} > 100$. There was no difference between the Social + Profit and the Social conditions: $t(375) = -0.58$, $p = 0.566$, $d = -0.06$, 95% CI [−0.26, 0.14], $BF_{10} = 0.134$, all two-sided tests.

7.3. Discussion

The results of Experiments 4a and 4b are broadly consistent with those of our earlier experiments. A noticeable difference is that tainting appears weaker in the financial domain. Unlike the health-related domain of the previous experiments, participants could have been less disturbed by the profit that accompanied the prosocial initiative, as the latter also resulted in monetary benefits for the nudgees as well as the nudger. Note that, in the scenarios we used, the profit from the initiative was always equally shared between the financial company and the charity or the clients. This design choice was dictated by our attempt to closely follow the GAP study described by Newman and Cain (2014, Experiment 4). But it is possible that the Social + Profit initiative could be evaluated as more or less acceptable depending on the allocation of profits between the organization and the beneficiaries of the initiative.

Although tainting was weaker overall than in previous experiments, when the profit information was made more salient in Experiment 4b, it appeared to be greater for charitable actions than for nudging. In neither case, however, did we find evidence for strong tainting. While Experiments 4a and 4b were partly inspired by the GAP study, our experimental design differed from it in several ways. First, to be consistent with our other experiments, we explicitly described the organization's motives, while in the GAP study inferences about motives were left to the participant. Second, the GAP study tested the effect of counterfactual information on tainted altruism, and therefore did not require the profit-only condition needed to test for strong tainted altruism. Third, we used a different subject pool (i.e., Prolific UK residents as opposed to US university students). Because of these important differences, we refrain from interpreting our results as a failure to replicate the previous evidence on tainted altruism, although they may suggest some limits to the generalisability of that evidence. We will come back to the difference between altruism and nudging in the general discussion. Before we turn to that, we present a composite analysis that jointly looks at all our experimental data.

8. Composite analysis

Overall, our experiments provide consistent evidence for tainting – albeit of different prevalence and strength for different settings – and very little evidence for strong tainting. Several researchers have argued for using *meta*-analytic estimates to obtain cumulative evidence for an underlying effect (e.g., Braver et al., 2014, Camerer et al., 2018, Open Science Collaboration, 2015). We followed this approach to obtain a best estimate of our overall effect sizes, and combined all the key results from the Social, Social + Profit, and Profit conditions from all our experiments in a composite analysis. To provide as broad and accurate a picture as possible, we included the results from the comparisons reported in the Supplementary Material (see S4 for the detailed results per experiment). We used standard random effects *meta*-analysis procedures to determine average effect sizes using the mean difference in prosociality, morality and support ratings between the relevant conditions, divided by the pooled standard deviation (Cohen's *d*). This composite analysis provides a good overview of the variability of the data and the overall estimate of the true effects.

The results for tainting are depicted in Fig. 10A, and for strong tainting in Fig. 10B. As can be seen, for all dependent variables the overall effect size favours tainting and does not favour strong tainting. Notably, the absolute effect size in Panel A is consistently smaller in magnitude than the effect size in Panel B for all three variables, suggesting that the negative effect of moving from a Social to a Social + Profit initiative is weaker than the negative effect of moving from a Social + Profit to a Profit initiative.

Fig. 10C and 10D break down the strong tainting comparisons further, distinguishing between experiments with and without deception. The importance of deception is very evident. When there is no deception, the Social + Profit judgements are comfortably above those for Profit. When there is deception, the Social[d] + Profit judgments are approximately the same as those for Profit, although on average still higher for the support dimension. As already suggested, deception is costly as it cancels out the effect of the prosocial benefit on judgments of prosociality and morality. This may be a result of prosociality and morality judgments being generally lower in the presence of deception. In Experiments 1a, 1b, 4a and 4b, when there was no suggestion that the organization deceived about its motives, participants seemed to have an overall positive view of the Social + Profit initiative, as revealed by average ratings above the midpoint of the scale. On the contrary, average ratings were below the midpoint in Experiments 2a, 2b and 3, suggesting that deception may be damaging.

9. General discussion

9.1. Key findings

Many organizations seek to improve the well-being of their customers and of society at large, providing this does not cut deeply into their profitability. Nudging provides a cost-effective way to achieve these goals. Indeed, organizations can often nudge their customers to take actions that are better both for the organization's finances and for the customers themselves. But there is a risk that potential customers will take a negative view of companies earning a profit in this way and consider them self-serving or cynical.

We investigated whether people do make such negative judgments when organizations use nudging to achieve social goals while earning a profit (tainted nudge) and, if so, whether they are judged even less favourably than when they aim at profit without benefitting society more broadly (strong tainted nudge). In broad-brush terms, we found that while organizations are judged more positively if they do not earn profit from doing good, they are still judged more positively than when they *only* earn profit. We conclude, therefore, that profitable prosocial nudges are likely to be judged less positively on key dimensions of morality, prosociality and support than nudges that contribute to society

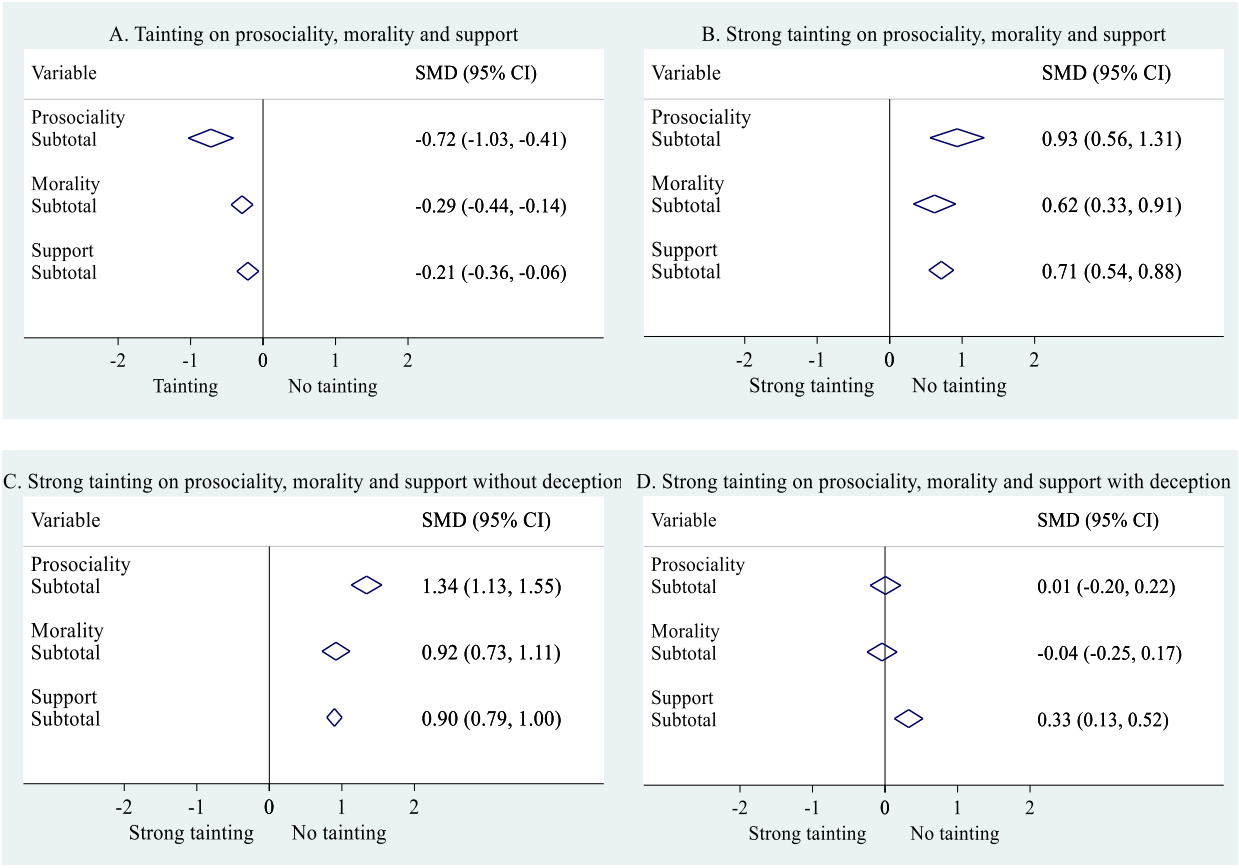


Fig. 10. Composite analysis for tainting (panel A), strong tainting (panel B), strong tainting without deception (panel C) and strong tainting with deception (panel D).

Notes. Diamond shapes indicate the average effect size and confidence intervals per dependent variable. The solid line represents an effect size of Cohen's d = 0. A negative effect size favours (strong) tainting, while a positive effect size favours no tainting.

alone.

However, as long as a nudging organization is fully transparent, prosocial nudges that result in profit are still judged more positively than otherwise identical nudges that produce only profit. It is largely when the organization deceives about its motives by falsely claiming to be socially motivated that it will be judged more harshly. Yet, even in this case, while the organization gets relatively little credit for its social good, it is not judged so harshly that the effects of deception universally outweigh those of the social good.

To reconcile our findings with earlier research on tainted altruism, we tested the effects of adding profit to an organization's purely charitable undertaking and found similar evidence as in the tainted nudge context. Profit (especially when made salient as in Experiment 4b) did taint the organization, but again this tainting did not approach the strong tainting previously suggested in the context of altruism. Organizations contributing to a charity while earning profit were still judged more favourably than organizations with profit as their only goal.

9.2. *Theoretical contributions*

We sought to learn whether earlier results on "tainted altruism" generalize to the domain of nudges. Previous research on attitudes toward nudges has primarily focused on the policy area, the choice architecture mechanism and the welfare benefits brought about by nudges (see e.g., Hagman et al., 2015; Reisch and Sunstein, 2016; Yan and Yates, 2019). Our work is based on the recognition that market pressures are likely to mean that for many, if not most, nudges profit and social benefits will be intertwined. This is simply because most organizations are concerned about their bottom line, and a sustainable nudge will ultimately have to translate into benefits for the organization that deploys it.

We showed that, although profits do taint prosocial nudges, they do so only to a limited extent. Even when it was very clear that the nudger had attempted to deceive about their motives by presenting a positive side effect as their goal all along, we did not observe strong tainting, but a general arosion of the positive effects of doing good. This is generally good news for honest nudgers, and an important addition to our understanding of the public acceptability of nudges. Our findings provide a new justification for nudge transparency, which has been called for in different contexts (e.g., Hansen and Jespersen, 2013; Loewenstein et al., 2015).

Our results also advance existing research on tainted altruism (e.g., Alcala et al., 2022; Carlson and Zaki, 2018; Newman and Cain, 2014) by testing its robustness and suggesting that, just as with nudges, deception is likely to play a key role. Some tainted altruism research has suggested that obtaining a personal benefit might routinely and significantly taint otherwise altruistic acts, and indeed taint them so much that an individual or organization would be better off not being altruistic at all. We have argued that these previous demonstrations often involve tainting not merely by profit, but also by deception (see also Alempaki et al., 2020). For example, in the large company scenario from Newman and Cain (2014, Experiment 3), in which the owner was judged more negatively when he donated to charity to improve his reputation than when he undertook an advertising campaign to achieve the same goal, there was no explicit suggestion that he was seeking to deceive. But that was a legitimate inference: "Mulberry [the owner] donated the money because he knew that the good publicity would boost the reputation of his company and get more people to come to his stores." In line with this

view, when Alcala et al. (2022) replicated the original study, they also assessed consumers' evaluations of Mulberry's deceptiveness, and found that he was judged as more deceptive when donating to charity than when advertising. Our results point in the same direction, and suggest that research on win–win initiatives would benefit from a deeper investigation of when and why organizational claims of prosociality might generate negative reactions, and in particular how large a role is played by perceived deception.

9.3. Individuals versus organizations

We focused on the actions of organizations as opposed to individuals. Since nudging is likely to be carried out by organizations (see Chapman et al., 2021 on the new frontiers for nudges and choice architecture in organizations), it is important to anticipate negative repercussions. Is it possible, for instance, that customers may dislike organizations that profit from doing good and this may have negative effects on sales? While research on tainted altruism might suggest this, it may be relevant that many of the most striking examples of tainting involve identifiable individuals taking altruistic or charitable actions, as in the case of Newman and Cain's (2014) man volunteering at a homeless shelter to get a date, or Carlson and Zaki's (2018) characters giving blood or returning a briefcase for a cash reward.

Our results suggest that conclusions that might be drawn from this earlier work may not apply to organizations seeking to nudge their clients, or indeed to organizations at all. It appears that people react differently to organizational interventions compared to individual behavior, and while they certainly respond less positively to organizations that profit while doing good compared to those that do only good, the magnitude of this effect is small. Even when we looked at whether organizations were tainted by charitable undertakings that earned a profit (Experiments 4a and 4b), we found limited tainting. Only when the company adopted a policy of deceiving about its motives were the positive effects of prosocial benefits largely eliminated. Moreover, when it came to whether a company could be recommended to consumers for its primary function (in this case, financial management) there was no suggestion at all of tainting, even in the charitable context.

Note that in the aforementioned examples of tainted altruism involving individuals, a morally neutral action is to do no good at all (e.g., "help at a coffeeshop to gain someone's affection" in Newman and Cain, 2014 or "play a game on one's phone" in Carlson and Zaki, 2018), while in the case of organizations, a morally neutral action is to continue doing business as usual. It could therefore be the case that the lack of strong tainting for organizations is partly due to people being negatively predisposed against organizations making money in the first place (see also e.g., Bhattacharjee et al., 2017 for supporting evidence of people holding anti-profit beliefs). Indeed, if we use the midpoint (4 out of 7, indicated with a dashed line in all result figures) of the morality and prosociality scales as a proxy for doing "no good at all", we find the average evaluations of profit-seeking activities are mostly below this midpoint. Nonetheless, the mean ratings of Social + Profit initiatives were consistently above the midpoint when the organization was transparent, and mostly below the midpoint when it was deceptive, suggesting that people were generally quite positively disposed towards prosocial initiatives that bring profits, provided these did not involve deception.

9.4. Practical implications

Our research offers important practical guidelines for the increasing number of organizations that use behavioral insights to influence their employees, their customers, or even people in general. The main takeaway is that win–win nudges will not fail to capitalize on the benefits of prosociality if organizations are fully transparent. Policy makers and companies can restrain the development of scepticism when engaging in win–win nudges by being open about their benefits. This is a positive finding, since although it is possible to promote interventions that separate prosocial benefits and profit, in some important real-world decisions the two types of benefits are inherently tied. For instance, this is often the case for environmental and health issues, which are among the most pressing issues currently faced by countries around the globe. Saving energy often saves money, even if people do it because they care about the environment. Adopting healthier lifestyles reduces the long-term cost of healthcare, even if people do it because they care about their own health. Even financial management companies are likely to earn more if they adopt policies that make their clients better off.

Our findings speak both to non-profit organizations, who focus more on doing good, and to for-profit organizations, who focus more on doing well. Win-win interventions will likely be supported independently of the profit orientation of the organizations who undertake them, a result which should eliminate one source of hesitancy with respect to the adoption of socially beneficial practices.

9.5. Limitations and directions for future research

Like most contributions in the closely related literature, our experiments rely on hypothetical scenarios. A strength of this methodology is its controlled experimental approach, which allowed us to make like-for-like comparisons of the same organization, in a setting within which several manipulations could be embedded with minimal changes to the baseline scenario. This gives us high confidence in the internal validity of our results. However, we recognize the trade-offs in terms of external validity. Although intentions are predictive of future behavior (see, e.g., the meta-analysis by Kraus, 1995), described events may cause different reactions than lived ones. Future work can examine how people evaluate real scenarios or examine actual willingness to support profitable nudges and organizations that implement them in the field, where the evaluators are active targets of the nudge.

Importantly, people's support for nudges may depend on who the nudge is targeting. For instance, Diepeveen et al. (2013) show that support for government policies to change health-related behaviors is highest for interventions targeting the behavior of others, while Avitzour et al. (2019) find that doctors are more supportive of nudging patients than are the patients themselves. Similarly, Cornwell and Krantz (2014) find that people are less likely to endorse nudges that affect them directly as opposed to nudges that affect people in general. Hence, future research should investigate if people are less forgiving of prosocial nudges that make profits if they themselves are the ones being nudged.

In addition, the motives and the benefits presented in our experiments were deliberately made salient and clearly identifiable. The true motive of the initiative, its consequences, and which of the latter were intended and which were accidental were unequivocally described to participants. In reality an organization's underlying motives may be ambiguous to observers, so discriminating between alternative motives and benefits can be more complex and leave room for multiple interpretations. When people become suspicious of an actor because they have difficulties in clearly deducing their motives, or when they encounter multiple motives (e.g., Szykman et al., 2004), they tend to weight negative information more heavily than positive information (e.g., Bhattacharya and Sen, 2004; Kim, 2011). Thus, it could be that even when the true motive of an organization is prosocial, people attribute self-serving or even manipulative intentions to their initiative because the prosocial motive is not clearly recognisable.

A final direction for future research concerns replication. Although some of our experiments were inspired by the studies reported in Newman and Cain (2014), they were not direct replications. Future research should examine the replicability and robustness of previous illustrations of strong tainted altruism. This replication would be an important step, both for theoretical and practical reasons, in delineating the circumstances under which strong tainted altruism emerges.

10. Conclusion

We examined whether economic considerations contaminate people's perception of prosocial nudges and other activities by organizations. This is an important question, given the growing interest of both policy makers and private sector entities in the implementation of nudges that achieve desirable social goals in an economically sustainable way. We demonstrated that profit might taint prosocial nudges, but to a limited extent. Perhaps most simply and reassuringly, our findings suggest that people recognize the welfare benefits of prosocial nudges. Nudging organizations can avoid any backlash by openly disclosing the profitable nature of their prosocial interventions.

Declarations

Funding The studies were funded by the Leverhulme Trust "Value" programme [grant number RP2012-V-022] and the Economic and Social Research Council [grant number ES/P008976/1] via the Network for Integrated Behavioural Science programme. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

Compliance with ethical standards

Conflict of interest All authors declare that they have no conflict of interest.

Ethical Approval All procedures performed in studies involving human participants were approved by the Humanities and Social Sciences Research Ethics Committee at the University of Warwick and were performed in accordance with the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Informed Consent Informed consent was obtained from all individual participants included in the studies.

Data, Materials and Code availability All data, materials and code underlying the results reported in the manuscript can be found on Open Science Framework at https://osf.io/4g83d/.

Author contributions

Author names are listed in alphabetical order. All authors contributed equally.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

We have made study materials and data fully available on Open Science Framework (https://osf.io/4g83d/).

Acknowledgements

We thank the Leverhulme Trust "Value" programme [grant number RP2012-V-022] and the Economic and Social Research Council [grant number ES/P008976/1] via the Network for Integrated Behavioural Science programme for financial support, Daniel Banki for his role in bringing this research question to our attention, and participants in various NIBS workshops and SPUDM 2019 for numerous helpful comments and suggestions.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.obhdp.2023.104244.

References

Adi, I. (2015). How Indra Nooyi turned design thinking into strategy. *Harvard Business Review, 93*(9), 80–85.

Alcala, V., Johnson, K., Steele, C., Wu, J., Zhang, D., & Pashler, H. (2022). The tainted altruism effect: A successful pre-registered replication. *Royal Society Open Science, 9* (1), 211152.

Alemanno, A., The Future of Behavioral Change: Balancing Public Nudging vs Private Nudging (January 4, 2016). 2nd AIM Lecture, May 6, 2015. Available at SSRN: https://ssrn.com/abstract=2710926 or http://dx.doi.org/10.2139/ssrn.2710926.

Alempaki, D., Isoni, A., & Read, D. (2020). Deception Aversion, Norm Violation and Consumer Responses to Prosocial Initiatives. *Available at SSRN*. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3640256.

Andersson, O., & Nelander, L. (2021). Nudge the Lunch: A Field Experiment Testing Menu-Primacy Effects on Lunch Choices. *Games, 12*(1), 2.

Arno, A., & Thomas, S. (2016). The efficacy of nudge theory strategies in influencing adult dietary behavior: A systematic review and meta-analysis. *BMC Public Health, 16* (1), 1–11.

Avitzour, D., Barnea, R., Avitzour, E., Cohen, H., & Nissan-Rozen, I. (2019). Nudging in the clinic: The ethical implications of differences in doctors' and patients' point of view. *Journal of Medical Ethics, 45*(3), 183–189.

Barber, B. M., Huang, X., Odean, T., & Schwarz, C. (2021). Attention induced trading and returns: Evidence from Robinhood users. *Journal of Finance*, forthcoming.

Barden, J., Rucker, D. D., & Petty, R. E. (2005). "Saying one thing and doing another": Examining the impact of event order on hypocrisy judgments of others. *Personality and Social Psychology Bulletin, 31*(11), 1463–1474.

Batson, C. D., Kennedy, C. L., Nord, L. A., Stocks, E. L., Fleming, D. Y. A., Marzette, C. M., … Zerger, T. (2007). Anger at unfairness: Is it moral outrage? *European Journal of Social Psychology, 37*(6), 1272–1285.

Becker-Olsen, K., & Potucek, S. (2013). Greenwashing. *Encyclopedia of Corporate Social Responsibility*, 1318–1323.

Benartzi, S., Beshears, J., Milkman, K. L., Sunstein, C. R., Thaler, R. H., Shankar, M., … Galing, S. (2017). Should governments invest more in nudging? *Psychological Science, 28*(8), 1041–1055.

Berman, J. Z., Levine, E. E., Barasch, A., & Small, D. A. (2015). The Braggart's dilemma: On the social rewards and penalties of advertising prosocial behavior. *Journal of Marketing Research, 52*(1), 90–104.

Berman, J. Z., & Silver, I. (2022). Prosocial behavior and reputation: When does doing good lead to looking good? *Current Opinion in Psychology, 43*, 102–107.

Bhattacharjee, A., Dana, J., & Baron, J. (2017). Anti-profit beliefs: How people neglect the societal benefits of profit. *Journal of Personality and Social Psychology, 113*(5), 671.

Bhattacharya, C. B., & Sen, S. (2004). Doing better at doing good: When, why, and how consumers respond to corporate social initiatives. *California Management Review, 47* (1), 9–24.

Blodgett, M. S., & Melconian, L. (2012). Health-care Nonprofits: Enhancing Governance and Public Trust. *Business and Society Review, 117*(2), 197–219.

Boruchowicz, C. (2021). *Public Policy, the Environment, and the Use of Green Nudges* (p. 175). Nudging Public Policy: Examining the Benefits and Limitations of Paternalistic Public Policies.

Braver, S. L., Thoemmes, F. J., & Rosenthal, R. (2014). Continuously cumulating meta-analysis and replicability. *Perspectives on Psychological Science, 9*(3), 333–342.

Burgoon, J. K. (1993). Interpersonal expectations, expectancy violations, and emotional communication. *Journal of Language and Social Psychology, 12*(1–2), 30–48.

Burgoon, J. K. (2009). Expectancy violations theory. In E. Griffin (Ed.), *A first look at communication theory* (pp. 84–97). New York: McGraw-Hill.

Cadario, R., & Chandon, P. (2019). Effectiveness or consumer acceptance? Tradeoffs in selecting healthy eating nudges. *Food Policy, 85*, 1–6.

Cadario, R., & Chandon, P. (2020). Which healthy eating nudges work best? A meta-analysis of field experiments. *Marketing Science, 39*(3), 465–486.

Camerer, C. F., Dreber, A., Holzmeister, F., Ho, T. H., Huber, J., Johannesson, M., … Wu, H. (2018). Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nature Human Behavior, 2*(9), 637–644.

Carlson, R. W., & Zaki, J. (2018). Good deeds gone bad: Lay theories of altruism and selfishness. *Journal of Experimental Social Psychology, 75*, 36–40.

Chance, Z., Dhar, R., Hatzis, M., & Bakker, M. (2016). How google optimized healthy office snacks. *Harvard Business Review*.

Chapman, G., Milkman, K. L., Rand, D., Rogers, T., & Thaler, R. H. (2021). Nudges and choice architecture in organizations: New frontiers. *Organizational Behavior and Human Decision Processes, 163*(C), 1–3.

Chen, H., Bernard, S., & Rahman, I. (2019). Greenwashing in hotels: A structural model of trust and behavioral intentions. *Journal of Cleaner Production, 206*, 326–335.

Choi, J. J., Laibson, D., Madrian, B. C., & Metrick, A. (2003). Optimal defaults. *American Economic Review, 93*(2), 180–185.

Clavien, C. (2018). Ethics of nudges: A general framework with a focus on shared preference justifications. *Journal of Moral Education, 47*(3), 366–382.

Cohen, I. G., Fernandez Lynch, H., & Robertson, C. T. (Eds.). (2016). *Nudging Health: Health Law and Behavioral Economics*. Baltimore: Johns Hopkins University Press.

Cornwell, J. F., & Krantz, D. H. (2014). Public policy for thee, but not for me: Varying the grammatical person of public policy justifications influences their support. *Judgment and Decision Making, 9*(5), 433.

Decker, W. H. (2012). A firm's image following alleged wrongdoing: Effects of the firm's prior reputation and response to the allegation. *Corporate Reputation Review, 15*(1), 20–34.
DellaVigna, S., & Linos, E. (2022). RCTs to scale: Comprehensive evidence from two nudge units. *Econometrica, 90*(1), 81–116.
Diepeveen, S., Ling, T., Suhrcke, M., Roland, M., & Marteau, T. M. (2013). Public acceptability of government intervention to change health-related behaviors: A systematic review and narrative synthesis. *BMC Public Health, 13*(1), 756.
Eccles, R. G., & Serafeim, G. (2013). A tale of two stories: Sustainability and the quarterly earnings call. *Journal of Applied Corporate Finance, 25*(3), 8–19.
Effron, D. A., & Monin, B. (2010). Letting people off the hook: When do good deeds excuse transgressions? *Personality and Social Psychology Bulletin, 36*(12), 1618–1634.
Faul, F., Erdfelder, E., Buchner, A., & Lang, A. G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149–1160.
Galinsky, A. D., & Moskowitz, G. B. (2000). Counterfactuals as behavioral primes: Priming the simulation heuristic and consideration of alternatives. *Journal of Experimental Social Psychology, 36*(4), 384–409.
Gosnell, G. K., List, J. A., & Metcalfe, R. D. (2020). The impact of management practices on employee productivity: A field experiment with airline captains. *Journal of Political Economy, 128*(4), 1195–1233.
Gravert, C., & Kurz, V. (2021). Nudging à la carte: A field experiment on climate-friendly food choice. *Behavioral Public Policy, 5*(3), 378–395.
Greitemeyer, T., & Sagioglou, C. (2018). When positive ends tarnish the means: The morality of nonprofit more than of for-profit organizations is tainted by the use of compliance techniques. *Journal of Experimental Social Psychology, 76*, 67–75.
Hagman, W., Andersson, D., Västfjäll, D., & Tinghög, G. (2015). Public views on policies involving nudges. *Review of Philosophy and Psychology, 6*(3), 439–453.
Handy, F., Seto, S., Wakaruk, A., Mersey, B., Mejia, A., & Copeland, L. (2010). The discerning consumer: Is nonprofit status a factor? *Nonprofit and Voluntary Sector Quarterly, 39*(5), 866–883.
Haran, U. (2013). A person–organization discontinuity in contract perception: Why corporations can get away with breaking contracts but individuals cannot. *Management Science, 59*(12), 2837–2853.
Hansen, P. G., & Jespersen, A. M. (2013). Nudge and the manipulation of choice: A framework for the responsible use of the nudge approach to behaviour change in public policy. *European Journal of Risk Regulation, 4*(1), 3–28.
Haugh, T. (2017). Nudging corporate compliance. *American Business Law Journal, 54*(4), 683–741.
Hornsey, M. J., Chapman, C. M., Mangan, H., La Macchia, S., & Gillespie, N. (2021). The moral disillusionment model of organizational transgressions: Ethical transgressions trigger more negative reactions from consumers when committed by nonprofits. *Journal of Business Ethics, 172*(4), 653–671.
Hotard, M., Lawrence, D., Laitin, D. D., & Hainmueller, J. (2019). A low-cost information nudge increases citizenship application rates among low-income immigrants. *Nature Human Behavior, 3*(7), 678–683.
Hummel, D., & Maedche, A. (2019). How effective is nudging? A quantitative review on the effect sizes and limits of empirical nudging studies. *Journal of Behavioral and Experimental Economics, 80*, 47–58.
Jachimowicz, J., Matz, S., & Polonski, V. (2017). *The Behavioral Scientist's Ethics Checklist.* Behavioral Scientist.
Jago, A. S., Kreps, T. A., & Laurin, K. (2019). Collectives in organizations appear less morally motivated than individuals. *Journal of Experimental Psychology: General, 148*(12), 2229.
John, P., Sanders, M., & Wang, J. (2014). The Use of Descriptive Norms in Public Administration: A Panacea for Improving Citizen Behaviors? *Available at SSRN*. https://ssrn.com/abstract=2514536.
Johnson, S. (2018). Dimensions of altruism: do evaluations of prosocial behavior track social good or personal sacrifice?. Available at SSRN 3277444.
Jung, J. Y., & Mellers, B. A. (2016). American attitudes toward nudges. *Judgment & Decision Making, 11*(1).
Kahneman, D. (2014). *Varieties of counterfactual thinking. In What might have been* (pp. 387–408). Psychology Press.
Kahneman, D., & Miller, D. T. (1986). Norm theory: Comparing reality to its alternatives. *Psychological Review, 93*(2), 136.
Kalil, A., Mayer, S. E., & Gallegos, S. (2021). Using behavioral insights to increase attendance at subsidized preschool programs: The Show Up to Grow Up intervention. *Organizational Behavior and Human Decision Processes, 163*(C), 65–79.
Kallbekken, S., & Sælen, H. (2013). 'Nudging' hotel guests to reduce food waste as a win–win environmental measure. *Economics Letters, 119*(3), 325–327.
Kim, H. S. (2011). A reputational approach examining publics' attributions on corporate social responsibility motives. *Asian Journal of Communication, 21*(1), 84–101.
Koh, H. K., Singer, S. J., & Edmondson, A. C. (2019). Health as a way of doing business. *Journal of American Medical Association, 321*(1), 33–34.
Kraus, S. J. (1995). Attitudes and the prediction of behavior: A meta-analysis of the empirical literature. *Personality and Social Psychology Bulletin, 21*(1), 58–75.
Kurz, V. (2018). Nudging to reduce meat consumption: Immediate and persistent effects of an intervention at a university restaurant. *Journal of Environmental Economics and Management, 90*, 317–341.
Lades, L. K., & Delaney, L. (2022). Nudge FORGOOD. *Behavioral. Public Policy, 6*(1), 75–94.
Laurent, S. M., Clark, B. A., Walker, S., & Wiseman, K. D. (2014). Punishing hypocrisy: The roles of hypocrisy and moral emotions in deciding culpability and punishment of criminal and civil moral transgressors. *Cognition & Emotion, 28*(1), 59–83.
Lin-Healy, F., & Small, D. A. (2012). Cheapened altruism: Discounting personally affected prosocial actors. *Organizational Behavior and Human Decision Processes, 117*(2), 269–274.
Lin-Healy, F., & Small, D. A. (2013). Nice guys finish last and guys in last are nice: The clash between doing well and doing good. *Social Psychological and Personality Science, 4*(6), 692–698.
Lin-Hi, N., Hörisch, J., & Blumberg, I. (2015). Does CSR matter for nonprofit organizations? Testing the link between CSR performance and trustworthiness in the nonprofit versus for-profit domain. *Voluntas: International Journal of Voluntary and Nonprofit Organizations, 26*(5), 1944–1974.
Longden, B., & Bélanger, C. (2013). Universities: Public good or private profit. *Journal of Higher Education Policy and Management, 35*(5), 501–522.
Loewenstein, G., Bryce, C., Hagmann, D., & Rajpal, S. (2015). Warning: You are about to be nudged. *Behavioral Science & Policy, 1*(1), 35–42.
Madrian, B. C., & Shea, D. F. (2001). The power of suggestion: Inertia in 401 (k) participation and savings behavior. *The Quarterly Journal of Economics, 116*(4), 1149–1187.
Makov, T., & Newman, G. E. (2016). Economic gains stimulate negative evaluations of corporate sustainability initiatives. *Nature Climate Change, 6*(9), 844.
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*(3), 101–104.
Merritt, A. C., Effron, D. A., & Monin, B. (2010). Moral self-licensing: When being good frees us to be bad. *Social and Personality Psychology Compass, 4*(5), 344–357.
Mertens, S., Herberz, M., Hahnel, U. J., & Brosch, T. (2022). The effectiveness of nudging: A meta-analysis of choice architecture interventions across behavioral domains. *Proceedings of the National Academy of Sciences, 119*(1), e2107346118.
Miller, D. T., & McFarland, C. (1986). Counterfactual thinking and victim compensation: A test of norm theory. *Personality and Social Psychology Bulletin, 12*(4), 513–519.
Newman, G. E., & Cain, D. M. (2014). Tainted altruism: When doing some good is evaluated as worse than doing no good at all. *Psychological Science, 25*(3), 648–655.
Oecd. (2019). Delivering Better Policies Through Behavioral Insights: New Approaches. *OECD Publishing, Paris,*. https://doi.org/10.1787/6c9291e2-en
Olivola, C. Y. (2011). *When noble means hinder noble ends: The benefits and costs of a preference for martyrdom in altruism* (pp. 49–62). The science of giving: Experimental approaches to the study of charity.
Olivola, C. Y., & Shafir, E. (2013). The martyrdom effect: When pain and effort increase prosocial contributions. *Journal of Behavioral Decision Making, 26*(1), 91–105.
Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science, 349*(6251), aac4716.
Oppenheimer, D. M., & Olivola, C. Y. (2011). *The science of giving: Experimental approaches to the study of charity*. Psychology Press.
Orange, E., & Cohen, A. M. (2010). From eco-friendly to eco-intelligent. *The Futurist, 44*(5), 28.
Pennycook, G., McPhetres, J., Zhang, Y., & Rand, D. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy nudge intervention. *PsyArXiv [Working Paper]*, 1–24. https://doi.org/10.31234/OSF.IO/UHBK9.
Raihani, N., & Power, E. (2021). No Good Deed Goes Unpunished: The social costs of prosocial behavior. *Evolutionary Human Sciences., 3*(e40), 1–21. https://doi.org/10.1017/ehs.2021.35
Reisch, L. A., & Sunstein, C. R. (2016). Do Europeans like nudges? *Judgment and Decision making, 11*(4), 310–325.
Roese, N. J. (1997). Counterfactual thinking. *Psychological Bulletin, 121*(1), 133.
Roese, N. J., & Olson, J. M. (Eds.). (2014). *What might have been: The social psychology of counterfactual thinking*. Psychology Press.
Schlesinger, M., Mitchell, S., & Gray, B. H. (2004). Public expectations of nonprofit and for-profit ownership in American medicine: Clarifications and implications. *Health Affairs, 23*(6), 181–191.
Sharma, V. K., Ingle, N. A., Kaur, N., Yadav, P., Ingle, E., & Charania, Z. (2016). Sugar Substitutes and Health: A Review. *Journal of Advanced Oral Research, 7*(2), 7–11.
Silver, I., & Silverman, J. (2022). Doing good for (maybe) nothing: How reward uncertainty shapes observer responses to prosocial behavior. *Organizational Behavior and Human Decision Processes, 168*, 104113.
Small, D. A., & Cryder, C. (2016). Prosocial consumer behavior. *Current Opinion in Psychology, 10*, 107–111.
Smith, P. C., & Richmond, K. A. (2007). Call for greater accountability within the US nonprofit sector. *Academy of Accounting and Financial Studies Journal, 11*(2), 75.
Strom, St. (2014, Sep 23). Soda Makers Coca-Cola, PepsiCo and Dr Pepper Join in Effort to Cut Americans' Drink Calories. New York Times. https://www.nytimes.com/2014/09/24/business/big-soda-companies-agree-on-effort-to-cut-americans-drink-calories.html?_r=2 (assessed 22-11-2021).
Subrahmanyam, A. (2008). Behavioral finance: A review and synthesis. *European Financial Management, 14*(1), 12–29.
Sunstein, C. R. (2015). The ethics of nudging. *Yale J. on Reg., 32*, 413.
Sunstein, C. R. (2019). *Which nudges do people like?* Edward Elgar Publishing: A national survey. In Handbook of Behavioral Change and Public Policy.
Sunstein, C. R. (2022). Sludge audits. *Behavioral. Public Policy, 6*(4), 654–673.
Sunstein, C. R., & Reisch, L. A. (Eds.). (2017). *The Economics of Nudge* (Vol. 4). London: Routledge. Critical Concepts of Economics.
Sunstein, C. R., & Reisch, L. A. (2019). A Bill of Rights for Nudging. *Journal of European Consumer and Market Law, 8*(3).
Szykman, L. R., Bloom, P. N., & Blazing, J. (2004). Does corporate sponsorship of a socially-oriented message make a difference? An investigation of the effects of sponsorship identity on responses to an anti-drinking and driving message. *Journal of Consumer Psychology, 14*(1–2), 13–20.

Tandel, K. R. (2011). Sugar substitutes: Health controversy over perceived benefits. *Journal of Pharmacology & Pharmacotherapeutics, 2*(4), 236.
Tannenbaum, D., Fox, C. R., & Rogers, T. (2017). On the misplaced politics of behavioral policy interventions. *Nature Human Behavior, 1*(7), 0130.
Thaler, R. H. (2018). Nudge, not sludge. *Science, 361*(6401), 431-431.
Thaler, R. H., & Benartzi, S. (2004). Save more tomorrow™: Using behavioral economics to increase employee saving. *Journal of Political Economy, 112*(S1). S164-S187.
Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions About Health, Wealth, and Happiness*. (New Haven, CT: Yale University Press).
Vandenbroele, J., Slabbinck, H., Van Kerckhove, A., & Vermeir, I. (2021). Mock meat in the butchery: Nudging consumers toward meat substitutes. *Organizational Behavior and Human Decision Processes, 163*(C), 105–116.
Vlaev, I., King, D., Dolan, P., & Darzi, A. (2016). The theory and practice of "nudging": Changing health behaviors. *Public Administration Review, 76*(4), 550–561.
Wootliff, J., & Deri, C. (2001). NGOs: The new super brands. *Corporate Reputation Review, 4*(2), 157–164.
Yan, H., & Yates, J. F. (2019). Improving acceptability of nudges: Learning from attitudes towards opt-in and opt-out policies. *Judgment and Decision Making, 14*(1), 26.