



ARTICLE

Nudges can be both autonomy-preserving and effective: evidence from a survey and quasi-field experiment

Henrico van Roekel¹ , Laura M. Giurge², Carina Schott³ and Lars Tummers⁴ 

¹Utrecht University, Utrecht, Netherlands, ²London School of Economics, London, UK, ³Utrecht University, Utrecht, Netherlands and ⁴Utrecht University, Utrecht, Netherlands

Corresponding author: Henrico van Roekel; Email: h.vanroekel@uu.nl

(Received 13 October 2022; revised 2 March 2023; accepted 21 April 2023)

Abstract

Nudges are widely employed tools within organizations, but they are often criticized for harming autonomy and for being ineffective. We assess these two criticisms simultaneously: can nudges be both autonomy-preserving and effective in changing behavior? We developed three nudges – an opinion leader nudge, a rule-of-thumb and self-nudges – to reduce a particularly sticky behavior: email use. In a survey experiment of 4,112 healthcare employees, we tested their effect on perceived autonomy and subjective effectiveness. We also tested traditional policy instruments for comparison. Next, to assess objective effectiveness, we conducted a quasi-field experiment in a large healthcare organization with an estimate of 1,189 active email users. We found that each nudge in isolation, but especially when combined, was perceived to be both autonomy-preserving and effective, and more so than traditional policy instruments like an access limit or a monetary reward. We also found some evidence that the combination of all nudges decreased actual email use. This paper advances the literature by showing how innovations in nudge design improve nudges' ability to be autonomy-preserving and effective.

Keywords: nudge; autonomy; effectiveness; email use; experiment

Introduction

Nudges have become widely employed tools within organizations. The popularity of nudges is largely attributed to their advantages: they are simple and low-cost to implement. A nudge is an intervention that alters people's behavior by changing the choice architecture in which a decision is made. Importantly, it does not forbid anything or change incentives (Thaler & Sunstein, 2021). Nudges work by altering information, changing the structure of a decision or assisting with decision-making (Münscher *et al.*, 2016). They have been shown to increase healthy life choices (Lin *et al.*, 2017),

stimulate evidence-based medicine (Nagtegaal *et al.*, 2019) and improve human-computer interaction (Caraban *et al.*, 2019).

However, nudges have also attracted criticism. Two critiques are particularly salient in the scholarly literature and in the media (Tummers, 2022). First, scholars have argued that nudges reduce autonomy (Hausman & Welch, 2010; Wilkinson, 2013). Although some studies have addressed this criticism (e.g., Wachner *et al.*, 2020, 2021), the debate continues in part because scholars use varying conceptualizations of autonomy (Vugts *et al.*, 2020). In this paper, autonomy is understood as the extent to which employees experience agency: to what extent does a nudge allow employees to make independent decisions in their work? This approach fits the findings of Vugts *et al.* (2020), who show most scholars, albeit implicitly, understand nudge autonomy in terms of agency. Autonomy is a fundamental human need that drives motivation (Ryan & Deci, 2017). If nudges decrease autonomy, their presence may not be desirable. The second criticism is that nudges are ineffective in changing behavior. In defense of nudges, a recent meta-analysis conducted on more than 200 studies found that nudges are effective with small to medium effect sizes (Mertens *et al.*, 2022). However, other scholars find no evidence for the effectiveness of nudges after correcting for publication bias (Maier *et al.*, 2022). Relatedly, as Mertens *et al.* (2022) admit, a nudge's effectiveness often depends on the type of nudge. Finally, there is scant knowledge about why nudges are effective (Szasz *et al.*, 2022).

These two criticisms should urge scholars to study whether nudges are autonomy-preserving and effective as well as why, how and under what conditions they work. Autonomy and effectiveness of nudges may present a tension: effective nudges could be less autonomy-preserving and vice versa. For example, defaults are more effective than other types of nudges (Mertens *et al.*, 2022), but respondents also expected default nudges to be particularly detrimental to autonomy (Wachner *et al.*, 2020, 2021). However, this tension between autonomy and effectiveness is not a given. An effective nudge can also increase autonomy by helping people make the choices they want to make (De Ridder *et al.*, 2020). We argue that whether nudges can preserve autonomy and be effective at the same time depends on the nudge design. Recent innovations in nudge theory, like those on nudge+, nudge vs think, boosting and self-nudges, can inform the development of nudges that are autonomy-preserving and effective (Hertwig & Grüne-Yanoff, 2017; Reijula & Hertwig, 2022; Banerjee & John, 2023).

Building on nudge theory innovations, we developed three nudges – an opinion leader nudge, a rule-of-thumb and multiple self-nudges – that target a sticky behavior: email use. Prior research has shown that despite its promised benefits, email has become a source and symbol of stress at work (Barley *et al.*, 2011; Brown *et al.*, 2014). As a result, email has been associated with a host of negative outcomes, including lower work quality (Rosen *et al.*, 2019), increased burnout threat (Belkin *et al.*, 2020) and decreased life satisfaction (Kushlev & Dunn, 2015). Reducing email use has therefore become a topic of increasing attention among scholars and practitioners (Cecchinato *et al.*, 2014; Bozeman & Youtie, 2020).

To what extent can nudges preserve autonomy and be effective in decreasing email use? First, in a pilot study of 435 employees, we tested whether the three nudges are perceived as autonomy-preserving and effective. Next, in a large-scale survey

experiment among 4,112 healthcare employees, we measured perceived autonomy and subjective nudge effectiveness in comparison to traditional email interventions based on policy instruments like a monetary reward. Because social desirability bias can threaten the validity of a survey, we added a modified version of the Bayesian truth serum to illicit more truthful responses (Prelec, 2004). Finally, to test for objective nudge effectiveness, we implemented a quasi-field experiment in a large healthcare organization with an estimate of 1,189 active email users.

Overall, by showing that we can design nudges that are perceived as autonomy-preserving and effective, our paper provides much-needed nuance to the debate surrounding nudge development (Wilkinson, 2013; Wachner *et al.*, 2021; Mertens *et al.*, 2022). We also contribute by showing how nudges could help reduce email use. Email communication has long posed a threat to employee productivity and well-being and existing research has failed to provide a solution (for an exception, see Giurge & Bohns, 2021). Finally, our paper makes a concrete methodological contribution by showing how nudges can be tested using both perceptions and behavioral outcomes. We also include multiple insights that enrich the results, by, for example, using a Bayesian truth serum to counter social desirability bias and comparing nudges to traditional policy instruments (Prelec, 2004; Tummers, 2019).

Theory

The nudge debate

In their influential book *Nudge* (2008), Thaler and Sunstein describe how organizations can use nudges to cope with biases in human decision-making. Rooted in behavioral economics, a nudge is an intervention that aims to influence people's behavior based on insights about the bounded rationality of people (Hansen, 2016). Bounded rationality refers to the notion that people are imperfect decision-makers that do not have access to all information and computational capacities that are required to evaluate the costs and benefits of potential actions. It contrasts the rational agent model prevalent in neoclassical economics, in which people are seen as rational agents that maximize their utility (Simon, 1955). Building on the work of Simon, psychologists aimed to develop maps of bounded rationality (Tversky & Kahneman, 1974, 1981; Kahneman, 2011). They analyzed the systematic errors that distinguish the actions people take from the optimal actions assumed in the rational agent model. Tversky & Kahneman (1974) show that heuristics, though useful, can lead to predictable and systematic errors. For example, anchoring bias refers to the tendency of people to overvalue the first piece of information they receive (e.g., Nagtegaal *et al.*, 2020). Such biases explain why people sometimes do not respond to traditional managerial instruments, like a bonus or a ban (Tummers, 2019). Instead, nudges aim to change the choice architecture without changing economic incentives or forbidding any options (Münscher *et al.*, 2016; Thaler & Sunstein, 2021).

Despite their popularity, nudges are not without criticism, two of which are particularly salient (see e.g., Tummers, 2022). First, nudges are said to reduce autonomy. Indeed, some scholars argue that because nudges work via unconscious processes, they can exploit weaknesses, manipulate and reduce choice (Hausman & Welch,

2010; Hansen & Jespersen, 2013; Wilkinson, 2013), and as a result harm autonomy. Nudges that are not autonomy-preserving are problematic because autonomy presents one of three basic and universal psychological human needs (next to the need for relatedness and the need for competence) that drives human behavior and motivation (Ryan & Deci, 2017). Notably, there are also scholars who argue choice architecture is always present, regardless of whether one actively influences it (Sunstein, 2016). Empirical evidence on this issue is inconclusive. Studies suggest some nudges can harm autonomy while others do not (e.g., Wachner *et al.*, 2020, 2021, Michaelsen *et al.*, 2021). Similarly, research on the public acceptance of nudges also shows mixed findings (Davidai & Shafir, 2020; Hagman *et al.*, 2022).

Besides the criticism on autonomy, scholars have disputed whether nudges are effective. In a recent meta-analysis, with over 200 studies, Mertens *et al.* (2022) found that nudges are, on average, effective in changing behaviors with small to medium effect sizes. There are, however, several counterarguments to this claim. First, the authors indicate that the effectiveness of a nudge depends on the type of nudge: nudges focused on decision structure (e.g., default nudges), outperform nudges focused on decision information or decision assistance. Second, Szaszi *et al.* (2022) note that context matters: whether nudges are effective varies and the conditions under which they work are barely identified. Third, Maier *et al.* (2022), in a response to Mertens *et al.* (2022), point out that after correcting for publication bias, there is no evidence that nudges are effective. Related to this this debate, Bryan *et al.* (2021) note that instead of focusing on replication in behavioral science, we need a heterogeneity revolution by analyzing which particular nudge works for what situation.

Autonomy-preserving and effective nudges

One of the reasons for the different views on autonomy and nudging depends on one's definition of autonomy, or lack thereof. Based on a systematic review, Vugts *et al.* (2020) show that the discussion surrounding nudge autonomy is clouded by different conceptualizations of autonomy. They identify three conceptualizations of autonomy (p. 108), namely freedom of choice (i.e., 'the availability of options and the environment in which individuals have to make choices'), agency (i.e., 'an individual's capacity to deliberate and determine what to choose'), and self-constitution (i.e., 'someone's identity and self-chosen goals'). A nudge could simultaneously decrease one's autonomy in one conceptualization and increase one's autonomy for another (Vugts *et al.*, 2020). For example, by limiting one's freedom of choice you could help people reach implicit goals and improve self-constitution. Similarly, a nudge can help someone think about the right choice but also limit the range of available choices to pick from – this would promote agency but limit freedom of choice.

In this paper, we adopt the conceptualization of autonomy as agency because it presents a higher threshold for nudge autonomy than the initial definitions of nudging and libertarian paternalism (see Thaler & Sunstein, 2008). The initial understanding of autonomy in the nudge literature relied heavily on freedom of choice, but 'apart from a context that allows choice, autonomy also requires a capacity to choose and decide' (Vugts *et al.*, 2020: 116).

At first glance, autonomy and effectiveness may appear to present a tension. For a nudge to be autonomy-preserving, the assumption is that the nudge guarantees agency (Vugt *et al.*, 2020) – meaning that it allows someone to execute their personal judgement (Morgeson & Humphrey, 2006; Gorgievski *et al.*, 2016). In contrast, an effective nudge assumes someone's personal judgement is flawed, because nudges are effective by being based upon – and making use of – biases in human decision-making (Hansen, 2016). While not removing any option, a nudge is effective by actively changing the choice architecture (Thaler & Sunstein, 2021) and may be considered manipulative (Wilkinson, 2013). For example, decision-structure nudges like defaults are more effective than decision information nudges like social norms (Mertens *et al.*, 2022). At the same time, people also expect default nudges to lower autonomy more so than social norm nudges (Wachner *et al.*, 2020; Wachner *et al.*, 2021). In contrast, people find social norm nudges to be more autonomy-preserving, yet these are more often ineffective (Wachner *et al.*, 2021). Although in some cases the autonomy-effectiveness tension may emerge, it is not a given. In fact, scholars have developed arguments about how autonomy and effectiveness go hand in hand. For example, De Ridder *et al.* (2020) argue an effective nudge can increase autonomy by helping people make the choices they want to make.

In Figure 1, we consider four scenarios for the ability of nudges to preserve autonomy and be effective. These scenarios are theoretical extremes and do not suggest that it is a yes/no question. First, when a nudge is effective but not autonomy-preserving, it is *unavoidable* and may well be a nudge that appears manipulative in the sense that it does not offer 'an escape clause' (Wilkinson, 2013: 354). Second, when a nudge is neither autonomy-preserving nor effective, it can decrease autonomy, while failing to do what it intended to do. Other unintended consequences may include when a nudge backfires or triggers reactance (Osman, 2020). Third, when a nudge is not effective but preserves autonomy, the nudge is meaningless and might be met with indifference (Wachner *et al.*, 2021). Finally, nudges could preserve autonomy and be effective (De Ridder *et al.*, 2020).

It is likely that the specific design of nudges impacts their autonomy and effectiveness. As such, when designing nudges, scholars should make use of theoretical concepts that are expected to increase nudges' ability to preserve autonomy and be effective. Below, we discuss nudge vs think, boosting and self-nudges (Sunstein, 2016; Hertwig & Grüne-Yanoff, 2017; Reijula & Hertwig, 2022).

First, scholars have discussed how nudging relates to thinking. Sunstein (2016) distinguishes System 1 (based on fast, intuitive thinking) and System 2 (using slow deliberation) nudges (based on Kahneman, 2011). People seem to prefer System 2 nudges (Sunstein, 2016). Furthermore, Lin *et al.* (2017) argue that nudges that promote reevaluation are more effective. Therefore, nudges that rely on conscious (System 2) decision-making, for example through employing information, may combine autonomy and effectiveness, more so than nudges that rely on unconscious (system 1) processes, for example by changing defaults. Additional arguments about how nudging can include deliberation have been developed by John *et al.* (2009) who discuss how 'nudge' and 'think' as behavioral change strategies may influence each other. Herein, a deliberative nudge refers to the combination of both (John *et al.*, 2022). In a similar vein, nudge+ refers to a Banerjee (Banerjee & John, 2021, 2023).

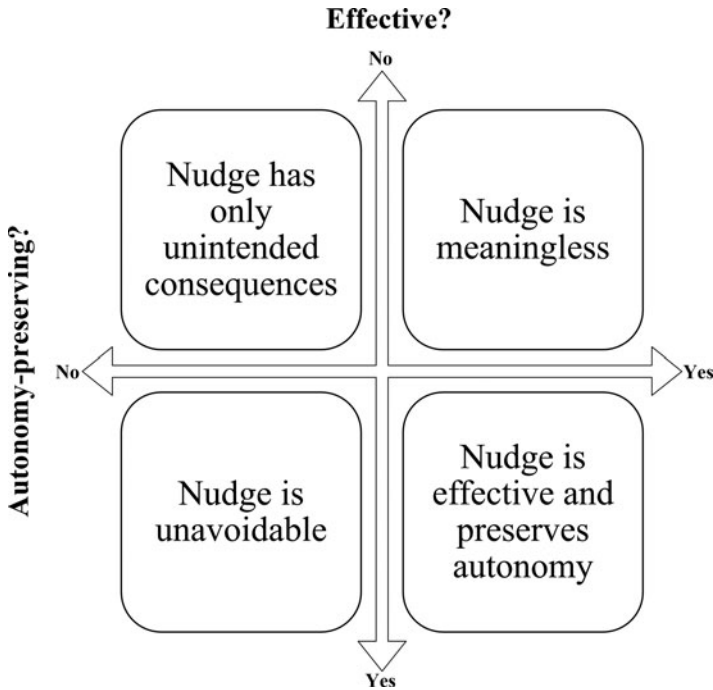


Figure 1. Four scenarios for autonomy and effectiveness of nudges.

Second, scholars have advocated to use boosting as a behavioral intervention (Hertwig & Grüne-Yanoff, 2017; Hertwig, 2017). A boost aims to enhance someone's decision-making by providing skills, knowledge or tools. Those advocating for boosting go even further than System 2 nudges by arguing that bounded rationality is malleable, and interventions should teach the decision-maker to change their behavior. An example of a boost is improving statistical reasoning with a brief training (Bradt, 2022). Like the distinction between System 1 and System 2 nudges, in practice, the distinction between boosts and nudges may be harder to uphold and interventions may carry characteristics from both nudging and boosting (Van Roekel *et al.*, 2022d). Yet, we can leverage this conceptual overlap by developing nudges that use insights from boosting. For example, to introduce a decision tree to guide decision-making is a type of boost (Hertwig & Grüne-Yanoff, 2017). A rule-of-thumb, which is a type of nudge, is effectively a simpler version of a decision tree (Münscher *et al.*, 2016).

Finally, scholars have been studying involvement of people in designing choice architecture. Involving employees in the process may make nudges more autonomy-preserving and effective. When it comes to development, this means being transparent and involving the target group in the trajectory of analyzing behavior and designing the nudges to get their support (Bruns *et al.*, 2018; Tummers, 2019). Employees could also be involved in the execution, an insight derived from the concept of self-nudging. Self-nudging suggests people can use nudges to self-regulate: 'nudger' and 'nudgee' become the same person. Self-nudges require

awareness of how one's environment affects one's behavior as well as knowledge of a nudge that can modify this relationship (Reijula & Hertwig, 2022: 123). In that sense, self-nudges can be regarded as a type of boost and may present a type of behavioral intervention that is both autonomy-preserving and effective (Reijula & Hertwig, 2022).

The role of nudges in reducing email use

We study nudges in the context of email use. Email has become a primary means of communication at work, because, in theory, it allows employees to decide when and where to work (Rosen *et al.*, 2019). However, email communication has become a unique job demand and source of stress at work because it facilitates non-stop sharing or requesting of input (Barley *et al.*, 2011). As such, many employees today feel compelled to read and respond to email in real-time, contributing to the development of workplace norms around continuous connectivity and instant responsiveness (Brown *et al.*, 2014; Giurge & Bohns, 2021).

Because of these norms, email communication has been associated with a host of negative consequences both at work and outside of work. For example, email undermines work quality because it fragments employees' attention (Jackson *et al.*, 2003). Email communication has also been associated with greater burnout and lower life satisfaction in part because it prohibits employees to disconnect from work and engage in non-work activities such as leisure that are beneficial for well-being (Kushlev & Dunn, 2015; Belkin *et al.*, 2020).

Ironically, although many studies tend to use email to distribute nudges (DellaVigna & Linos, 2022), nudges have rarely been used to directly alter email use. Rather, most studies that aim to alter email use are focused on changing the person rather than the environment. For example, Dabbish & Kraut (2006) found that several individual email management tactics (e.g., having less email folders) were associated with lower email overload. Relatedly, Gupta *et al.* (2011) found that limiting the moments when one checks their email decreased stress, which in turn predicted greater well-being. In terms of nudges, we only found one paper that employed nudges to increase awareness of phishing (Vitek & Syed Shah, 2019), which arguably focuses more on changing how employees interact with the content of emails than with email use. However, outside the academic literature, many nudge-like software is available, like reminders when an email is written poorly or a simple cognitive test to assess whether the user is fit to send emails at certain times (Balebako *et al.*, 2011). In line with this evidence, we expect that nudges can be used not only to inform people about how to improve and engage with email content but also how to address email use altogether (Cecchinato *et al.*, 2014; Bozeman & Youtie, 2020). Hence, our main hypothesis¹ is:

H1: Nudges will be both autonomy-preserving and effective in decreasing email use.

¹We preregistered four hypotheses: one or two hypotheses per study. In the general paper, we focus on one overarching hypothesis. Supplementary Appendix G discusses the preregistration and evaluation of all original hypotheses.

Methods

We study the autonomy-preservation and effectiveness of nudges in the context of email use among healthcare workers, which is a group of employees that are particularly prone to burnout and high email use (Reith, 2018; Van Roekel *et al.*, 2020). Table 1 presents an overview of our studies. After the pre-study to develop the nudges and the pilot to test the nudges, the empirical studies (the survey experiment and the quasi-field experiment) allow to evaluate our main hypothesis. The main empirical studies underwent ethical review were preregistered (Van Roekel *et al.*, 2022a, 2022b, 2022c; see also Supplementary Appendix G) and present open data (more information below).

Pre-study: developing nudges

In a pre-study, we developed the nudges and interviewed 11 employees in the organization where we would later conduct our quasi-field experiment (5 HR advisors, 1 program manager, 1 team manager, 1 nurse, 1 occupational physician, 1 occupational health nurse and 1 IT employee). Supplementary Appendix A contains the semi-structured interview guide. We used the interviews to develop three nudges: an opinion leader nudge, a rule-of-thumb and multiple self-nudges. An opinion leader nudge is a message that describes the behavior of a person of influence, assuming it will convince receivers to follow the social reference point (Valente & Pumpuang, 2007; Münscher *et al.*, 2016). A rule-of-thumb is an easy-to-follow guideline that works well in most situations and decreases the effort of a decision (Münscher *et al.*, 2016; Hertwig & Grüne-Yanoff, 2017). Self-nudges are nudges redesigned to be used by employees to nudge themselves. They help boost self-control (Hertwig & Grüne-Yanoff, 2017). Supplementary Appendix B presents the nudges that we developed to decrease email use in detail.

Pilot: testing nudges

After developing the nudges, we piloted them in an online Prolific panel ($N = 435$). We used the panel to assess the perceived autonomy and subjective nudge effectiveness from the perspective of the general working population (DellaVigna *et al.*, 2019). For the measurement of subjective nudge effectiveness, we asked the panel to predict

Table 1. Overview of studies

Study	Method	<i>N</i>	Goal
Pre-study	Interviews	11 respondents	Develop nudges
Pilot	Survey experiment	435 respondents (general population)	Test nudges
1	Survey experiment	4,112 healthcare employees	Test H1 (perceived autonomy and subjective nudge effectiveness)
2	Quasi-field experiment	Healthcare organization with an estimate of 1,189 active email users	Test H1 (objective nudge effectiveness)

the feasibility, appropriateness, meaningfulness, and effectiveness of the nudges in their organization (following the FAME-approach for evidence-based practice, Jordan *et al.*, 2019). Supplementary Appendix C presents all survey measures used in this paper and Supplementary Appendix D describes the methods and results of the pilot in detail.

All nudges were assessed as autonomy-preserving. Respondents indicated that on average they ‘somewhat agree’ to ‘agree’ with the nudges being autonomy-preserving; the means are all above 5 on a 1–7 scale. The results also indicate respondents thought the nudges would be ‘somewhat effective’. Scores were highest for appropriateness (4.55–5.28 on a 7-point scale) and lowest for effectiveness (3.97–4.35). The only score just below the midpoint (<4) was for the effectiveness of the rule-of-thumb (3.97), indicating that this nudge was perceived as least effective.

Study 1: Survey experiment

The goal of Study 1 was to test perceived autonomy and subjective nudge effectiveness in a large-scale survey experiment among healthcare employees in the Netherlands ($N = 4,112$). Employees assessed the nudges individually and combined. To compare these nudges with alternative organizational interventions, we also asked employees to assess traditional policy instruments (i.e., an email access limit to limit emailing to only 2 h per day, a monetary reward for emailing less than before or public praise for emailing less than your colleagues). Supplementary Appendix E details the specific text of the traditional interventions.

The large-scale survey experiment was part of a longitudinal survey for which ethical approval was granted (Faculty’s Ethical Review Committee of the Faculty of Law, Economics and Governance, Utrecht University; no. 2019-004). Respondents provided informed consent, including allowing for the publication of anonymized data. The main data for this study is available at https://osf.io/6n2g4/?view_only=895be1c46d384867b52e22ff30892ba8.

Participants

We collected data between 18 May and 20 June 2022 via a Qualtrics survey. Respondents were required to use email in their job, list-wise deletion was applied.

The mean age of the respondents was 51.94 (SD = 9.65, Min. = 20, Max. = 74, 3 missing). Regarding gender, 3,506 were female (85.3%), 587 were male (14.3%) and 19 respondents indicated X or that they would rather not say (0.4%). Respondents worked in all healthcare sectors: 1,515 (36.8%) in hospitals, 1,059 (25.8%) in nursing or home care, 653 (15.9%) in mental healthcare, 620 (15.1%) in disabled care and 265 (6.4%) in other healthcare. A total of 2,116 (51.5%) respondents worked 29 or more hours a week, 1,914 (46.5%) of respondents worked 16–28 h a week, 69 respondents (1.7%) worked 15 h or less, and 13 respondents (0.3%) reported to have a zero-hours contract.

Procedure and measures

All survey measures mentioned below are included in Supplementary Appendix C. Respondents first passed an eligibility check (respondents had to use email at their

job). We assessed two measures of email use (email volume and email time) with open questions adapted from Sumecki *et al.* (2011). We assessed email overload with a seven-item scale ($\alpha = 0.80$) adapted from Dabbish & Kraut (2006), ranging from 'strongly disagree' (1) to 'strongly agree' (7). Next, respondents were exposed to one of the seven interventions randomly: one of the three nudges, the combination of all nudges or one of the three traditional email interventions (translated in Dutch). Supplementary Appendix F shows that randomization was successful across gender, age, healthcare sector and amount of working hours. The instruction accompanying the intervention read: 'Imagine the organization you work for sends you the following message about using email in your organization. Please read the message carefully'.

After the conditions, respondents were asked to evaluate the subjective nudge effectiveness. We used self-admission rates for our main analysis and added an adapted version of the Bayesian truth serum to increase the credibility of the given answers (Prelec, 2004; John *et al.*, 2012; Weaver & Prelec, 2013; Frank *et al.*, 2017; Van de Schoot *et al.*, 2021; Schoenegger, 2023). Supplementary Appendix C elaborates on the serum.

Next, like the pilot study, we assessed perceived autonomy with three items ($\alpha = 0.93$) on a 7-point Likert scale (Decision-Making Autonomy; WDQ; Morgeson & Humphrey, 2006; translation adapted from Gorgievski *et al.*, 2016). We also measured work engagement with three items ($\alpha = 0.80$) on a 5-point Likert scale ranging from 'Never' (1) to 'Always (daily)' (5) (Schaufeli *et al.*, 2019). All items/questions were translated in Dutch. In the survey, items within each measure were randomized. At the end of the survey, respondents provided background characteristics. We used one-way analysis of variance for our main analysis. Significance levels were set at $p = 0.05$. We report exact p -levels.

Study 2: Quasi-field experiment

In Study 2, we tested the nudges sequentially in a quasi-field experiment in a large Dutch healthcare organization. This quasi-experiment had a One-Group Pre-test-Post-test Design with multiple sequential treatments and post-tests (Shadish *et al.*, 2002).

The prefix 'quasi' is appropriate because the experiment did not include a control group (Shadish *et al.*, 2002). Within the organization, there were technical limitations so interventions could not be randomly distributed to a selection of employees (i.e., there was no option to randomly send text messages to employees or randomly show intranet messages to a selection of employees) nor could any treatment group be separated from a control group when measuring email use. Moreover, any alternative treatment distributions that were considered (e.g., physical posters) would risk spill-over effects within the organization. The main disadvantage of a design without a control group is that differences in email use between pre- and post-intervention periods may be caused by elements or events unrelated to the treatment. Nevertheless, there are two main reasons why quasi-experiments are valuable designs to assess causality in instances where randomized designs are not possible or desirable (Shadish *et al.*, 2002; Grant & Wall, 2009). First, a quasi-experiment like this one should be seen as a method of action research, an opportunity for collaboration

between researchers and practitioners to jointly improve, in this case, the use of email within the organization (Grant & Wall, 2009). Second, our study still provides an estimate of the nudge effectiveness and we have taken several measures to improve reliability, including only measuring full workdays (excluding weekends and single holidays), planning the experiment in a period where few natural fluctuations were expected (there were no long holidays right before, during or after the experiment and no major events happened) and tracking email use for 8 weeks. Such measures make the effects of external events (like history or maturation) less likely (Shadish *et al.*, 2002).

We tested a total of four interventions (the three nudges and their combination), added a post-test after each intervention, an additional post-test before the combination of nudges, and two additional post-tests at the end of the experiment. Email use was measured weekly for eight consecutive weeks. Consequently, our design was the following, whereby O_n refers to the n th test and X_n to the n th treatment:

$O_1 \quad X_1 \quad O_2 \quad X_2 \quad O_3 \quad X_3 \quad O_4 \quad O_5 \quad X_4 \quad O_6 \quad O_7 \quad O_8$

In this study, we measured email use only, and not perceived autonomy, for two reasons. First, the added value of the field experiment was to test the effectiveness of the nudges on real behavior, while the survey experiment establishes perceptions of effectiveness and autonomy. A measure of autonomy in the field experiment would, again, be a perception. Second, the organization in which the quasi-field experiment was conducted, did not allow for large-scale surveying of employees, making it impossible to collect employees' perceptions about autonomy. A drawback of this approach is that autonomy may be perceived differently in the field compared to the survey study. While we cannot rule this out, it would only be a problem if nudges are considered less autonomy-preserving in real-life settings compared to hypothetical settings. However, a recent study showed that when people expect a nudge to diminish their autonomy in a hypothetical setting, they do not report any differences in autonomy for that same nudge in a real-life setting (Wachner *et al.*, 2021). Hence, whereas there is mixed evidence on the autonomy of nudges in hypothetical settings (e.g., Michaelsen *et al.*, 2021), nudges can likely be considered autonomy-preserving in real-life settings if the same nudges are considered autonomy-preserving in a hypothetical setting.

This study received ethical approval (Faculty's Ethical Review Committee of the Faculty of Law, Economics and Governance, Utrecht University; no. 2022-001). Data were collected via the Microsoft Office 365 portal of the organization. The organization signed a formal agreement to share and enable publication of anonymized data. The main data for this study is available at https://osf.io/6n2g4/?view_only=895be1c46d384867b52e22ff30892ba8.

Participants

The quasi-field experiment was conducted at a large healthcare organization in the Netherlands. The organization has 22 locations in one city, employs around 2,300 employees (not including volunteers) and delivers care to more than 6,500 elderly clients.

Nudges

The nudges were identical to those in Study 1, but with a few minor changes for a better fit with the context. For the opinion leader nudge, rather than ‘your HR manager’, the name of the HR manager was included to increase the ecological validity of the nudge. Third, for the rule-of-thumb, the suggested ‘within a day’ communication option was a Teams message as this was the preferred mode of communication. The rule-of-thumb also included a brief note that any communication about clients should be done with a secure messaging tool.

Procedure

We measured email use (amount of sent emails) during eight consecutive weeks. All employees were subjected to the three nudges, distributed a week apart and starting in the second week. The nudges were distributed on three subsequent Mondays (28 March, 4 April and 11 April 2022) around 11:30 AM CEST to 3,038 work phones. The time was purposefully chosen because most employees experience a drop in daily workload after the morning duties. The SMS messages read ‘Do you also want an emptier mailbox? Click here for a message/the second message/the last tips about emailing within [organization]’ [Link]. Regards, [organization] Two weeks later (25 April, at 11:30 AM CEST), we posted the combination of all the nudges on the intranet of the organization.

To measure objective nudge effectiveness, the key dependent variable is the number of emails sent within the organization using administrative data available via Microsoft Office 365. In our main analysis, we excluded weekends and holidays as on these days, employees would email much less. We used linear mixed models to assess statistical significance, comparing each week to the week before in separate tests (Krueger & Tian, 2004) and using the Benjamin–Hochberg false discovery rate control to correct for multiple tests (Glickman *et al.*, 2014). False discovery rate control is a less conservative alternative to the Bonferroni correction. It involves (a) sorting p-values in ascending order, (b) calculating the corrected p threshold per test by dividing the test number (e.g., 1 for test 1) by the total amount of tests (7 in this case) and multiplying this by the maximum false discovery rate (set at 0.05) and (c) declaring those tests with p-values lower than the corrected p threshold significant.

Results

Results Study 1: Survey experiment

Table 2 presents the correlations. One notable finding is that the score for perceived autonomy and all non-compliance estimates have significant negative correlations. This indicates perceived autonomy and subjective nudge effectiveness are positively correlated.

Figure 2 presents the means and 95% confidence intervals for perceived autonomy for each intervention.

A one-way analysis of variance showed that the effects on perceived autonomy differed significantly, $F(6, 1,823.18) = 135.51$, $p < 0.001$ ($\omega^2 = 0.17$)². *Post hoc* analyses

²Welch ANOVA was reported and Games-Howell *post hoc* tests were used as equal variances could not be assumed, $F(6,4105) = 19.12$, $p < 0.001$.

Table 2. Correlations ($N = 4,112$)

	M (SD)	1	2	3	4	5	6	7	8
1. Email volume	63.52 (86.25)	–	–	–	–	–	–	–	–
2. Email time	58.88 (66.98)	0.54**	–	–	–	–	–	–	–
3. Email overload	3.06 (1.01)	0.24**	0.11**	–	–	–	–	–	–
4. Autonomy	4.05 (1.56)	–0.03	–0.03*	0.002	–	–	–	–	–
5. Work engagement	3.88 (0.71)	–0.01	0.01	–0.17**	0.07**	–	–	–	–
6. Self-admission rate	0.68 (0.47)	–0.03	–0.02	–0.12**	–0.34**	–0.03*	–	–	–
7. Prevalence estimate	65.91 (25.25)	0.11**	0.08**	–0.02	–0.24**	–0.08**	0.43**	–	–
8. Admission estimate	60.2 (27.55)	0.03	0.03*	–0.1**	–0.18**	0.02	0.30**	0.46**	–

* $p < 0.05$, ** $p < 0.001$ (two-tailed).

Notes: Correlations are Pearson except for those with email volume and email time, these are Spearman as for these variables the data indicated outliers. Variables 6–8 measure subjective nudge effectiveness. Note the self-admission rate is coded as a dummy (0: would comply, 1: would not comply). Supplementary Appendix C provides more information on the measurement on the self-admission rates, prevalence estimates and admission estimates.

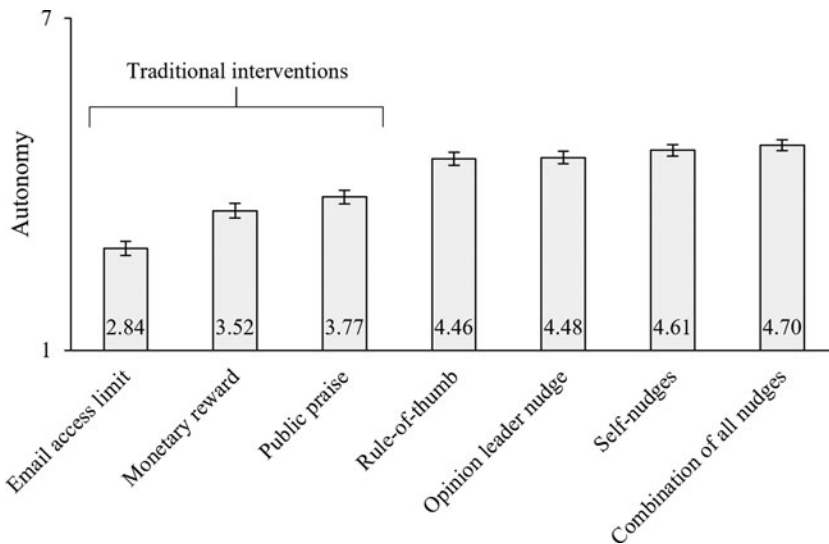


Figure 2. Nudges are seen as more autonomy-preserving than the midpoint (>4), and more autonomy-preserving than traditional interventions. *Note:* Perceived autonomy scores show 95% confidence intervals.

indicated that all traditional interventions scored significantly lower ($p < 0.001$) on perceived autonomy than any intervention with nudges. Besides, the email access limit scored significantly lower ($p < 0.001$) than the monetary reward and public praise. One significant difference between the nudges was found, the difference between all nudges and the rule-of-thumb is significant ($p = 0.028$). In sum, the results confirm our hypothesis that nudges are autonomy-preserving (scoring 4.46–4.7 on a 7-point scale), more so than traditional interventions (scoring 2.84–3.77).

Figure 3 presents the self-admission rates and the Bayesian Truth Serum results (recoded to self-admission rates of *compliance* to indicate how many respondents indicated they would send less emails) as a corrected conservative estimate of true compliance.

We conducted a one-way analysis of variance for the self-admission rates. This analysis showed that effects differed significantly, $F(6, 1,822.54) = 26.33$, $p < 0.001$ ($\omega^2 = 0.036$)³. *Post hoc* analyses indicated that all traditional interventions had significantly lower compliance than any nudge ($p < 0.001$), except for the rule-of-thumb. The rule-of-thumb had significantly higher compliance than the email access limit ($p = 0.017$) and public praise ($p = 0.005$), but not the monetary reward ($p = 0.465$). Besides, the rule-of-thumb had significantly lower compliance than the opinion leader nudge ($p = 0.007$) and all nudges ($p < 0.001$). Finally, the self-nudges had significantly lower compliance than the combination of all nudges ($p = 0.042$). The Bayesian truth serum indicates roughly the same, but more conservative, distribution,

³Welch ANOVA was reported and Games-Howell *post hoc* tests were used as equal variances could not be assumed, $F(6,4105) = 87.21$, $p < 0.001$.

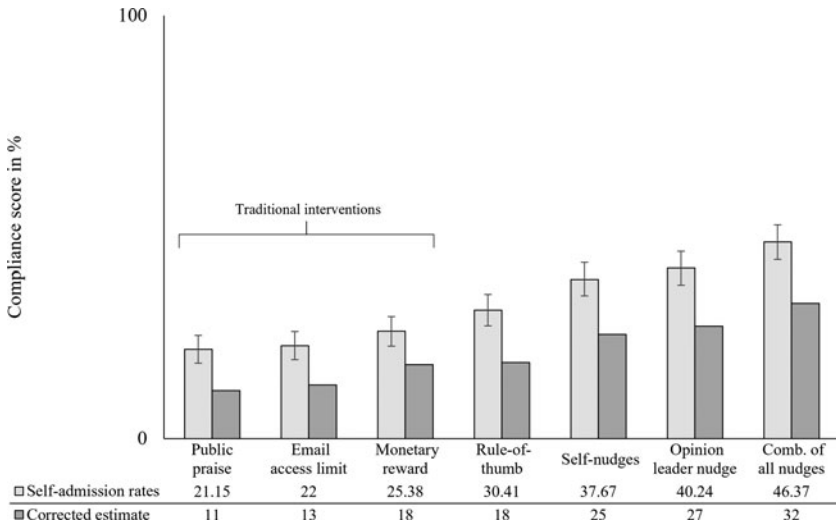


Figure 3. Nudges are perceived as more effective than the traditional interventions, but less than 50% of employees would comply with any intervention. *Note:* Self-admission rates show 95% confidence intervals.

with a notable exception for the already insignificant difference between the monetary reward and the rule-of-thumb.

The results are in line with our main hypothesis. For a notable group of employees (30–46%), the nudges would be effective in reducing email use. The nudges are more effective in reducing email use than an email access limit, monetary reward or public praise (the latter scored 21–25% compliance) (except the difference between the rule-of-thumb and the monetary reward is non-significant).

Results Study 2: quasi-field experiment

The first nudge was viewed 220 times, the second nudge 106 times and the third nudge 75 times. The combination of all nudges received 142 views in the first week (25 April–1 May) and 43 views in the second week (2 May–8 May), 185 views in total. This indicates many of the 3,038 recipients did not click on the link in the SMS. However, calculating a response rate on that total is misleading, as this does not equal the number of employees who use email. To estimate a more realistic response rate, we extracted email data on an individual level from the organization for the first week of the study (28 March–3 April). We first checked how many email IDs were in use that week ($N = 2,618$) and introduced the eligibility criterium of having sent at least one email that week, resulting in a total N of 1,189 active email users. This suggests that the estimated response rates for the first nudge was 18.50%, for the second nudge 8.92%, for the third nudge 6.30% and for the combination of all three nudges 15.56%.

During the 8-week intervention period, a total of 236,785 emails were sent within the organization (excluding weekends and holidays). This is an average of 6,400 per day with a standard deviation of 1,125.

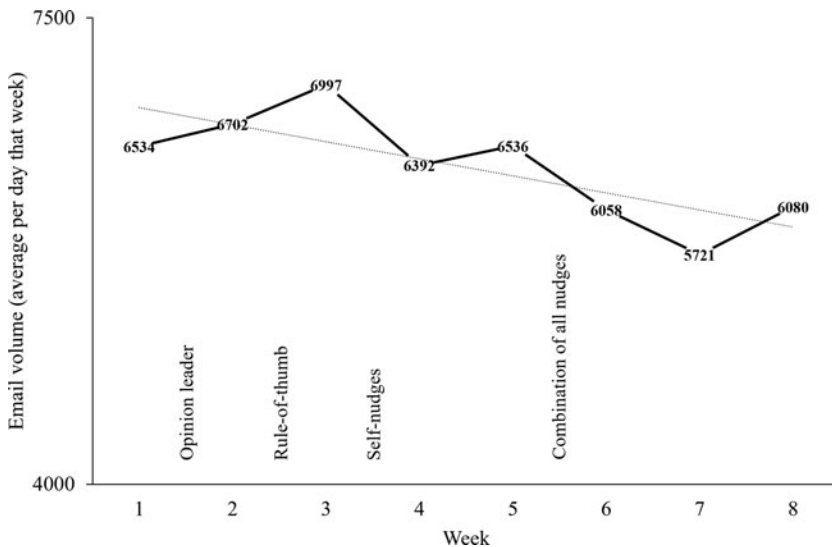


Figure 4. Email use decreases after the self-nudges and the combination of all nudges.

Figure 4 presents the average amount of email per day that was sent in the organization during each week of the intervention period, excluding holidays (Monday in week 5, Wednesday in week 6 and Thursday in week 7⁴) and weekends. We fit a linear trendline to indicate that, in general, email use decreased during the 8-week period ($y = -128.03x + 6953.6$; $R^2 = 0.593$). Between the first and last week of the quasi-experiment, average email use decreased by 6.95%. The biggest difference recorded was between week 3 and week 7 (−18.24%). Specifically, average email use decreased only during the week in which the self-nudges were distributed, and in the two weeks following the combination of all nudges.

To test for statistical significance, we used seven linear mixed models to compare each week to the week before, using the variable describing the week as a repeated measure fixed factor, and the unstructured repeated covariance type. Table 3 presents the results of the separate linear mixed models and presents the corrected results using the Benjamin–Hochberg false discovery rate control to correct for multiple tests (Glickman *et al.*, 2014).

The results indicate that evidence for our hypothesis in the quasi-field experiment is mixed: email use did not decrease after presenting the opinion leader nudge or the rule-of-thumb, but it did decrease after presenting the self-nudges and the combination of all nudges. Statistical tests indicate a significant decrease ($p < 0.05$) after the combination of nudges, but this result is not significant after controlling for multiple tests. Across two months, however, the linear trendline indicates email use did generally decrease.

⁴Monday, April 18 was Easter Monday; Wednesday, April 27 was King's Day; Thursday, May 5 was Liberation Day.

Table 3. Linear mixed models and false discovery rate control

Test	Weeks	Days analyzed	Results	p < 0.05?	FDR p	p < FDR p?
1	1 vs 2	M, Tu, W, Th, F	$t(4) = -1.67, p = 0.170$	No	0.04	No
2	2 vs 3	M, Tu, W, Th, F	$t(4) = -1.03, p = 0.360$	No	0.05	No
3	3 vs 4	M, Tu, W, Th, F	$t(4) = 2.76, p = 0.051$	No ^a	0.02	No
4	4 vs 5	Tu, W, Th, F	$t(4) = -5.93, p = 0.010$	Yes	0.007	No ^a
5	5 vs 6	Tu, Th, F	$t(4) = 6.30, p = 0.024$	Yes	0.01	No ^a
6	6 vs 7	M, Tu, F	$t(4) = 2.83, p = 0.106$	No	0.03	No
7	7 vs 8	M, Tu, W, F	$t(4) = -1.46, p = 0.240$	No	0.04	No

Notes: In case any of the values of any day was missing (due to holidays), this day was removed from analysis in both weeks of a particular test. M, Monday; Tu, Tuesday; W, Wednesday; Th, Thursday; F, Friday. FDR p refers to the corrected p threshold calculated with false discovery rate control.

^aIndicates p-values very close to significance.

Discussion

Main findings

Our analysis of the survey experiment and quasi-field experiment yields three main findings. First, the nudges we developed were perceived as autonomy-preserving, and significantly more so than traditional interventions (email access limit, monetary reward and public praise). Second, our nudges were perceived as significantly more effective than the traditional interventions (except for the rule-of-thumb vs the monetary reward), but in general less than 50% of employees would comply with any intervention. We observe combining multiple nudges increased employees' perceptions of autonomy and effectiveness. We also found a positive correlation between perceived autonomy and subjective nudge effectiveness. Third, further evidence for the objective effectiveness of the nudges is presented in the quasi-field experiment. Email use in the healthcare organization decreased generally during the 8 weeks of our quasi-field experiment. Specific decreases were observed after the self-nudges and the combination of all nudges, albeit, after controlling for multiple tests, these effects did not reach conventional statistical significance levels.

Implications

The findings contribute to three major scholarly debates: nudge design, email use and interventions in the field.

Our paper contributes to the nudge literature by bringing nuance to the debates about the ability of nudges (1) to preserve autonomy (Hausman & Welch, 2010; Hansen & Jespersen, 2013), (2) to be effective in changing behavior, (Mertens *et al.*, 2022; Szaszi *et al.*, 2022; Maier *et al.*, 2022) and (3) whether these two criticisms inherently present tensions (Wachner *et al.*, 2020; Wachner *et al.*, 2021; Mertens *et al.*, 2022). We developed four scenarios for the autonomy and effectiveness of nudges. Using innovations in nudge design like self-nudges, we show that multiple types of nudges can be perceived as autonomy-preserving and effective in a survey

setting. Another interesting finding is that combining nudges appears fruitful – suggesting that the sum may be perceived to be more than its parts. However, results are less pronounced in the field setting compared to the survey. There may be a variety of reasons, including that in the survey respondent's undivided attention is on a nudge, whereas in the field setting, employees may receive more messages simultaneously and choose not to engage (this is visible in the number of views the nudges received). Yet, ironically, the fact that a large share of employees chose not to engage does support the notion that nudges are autonomy-preserving even, and perhaps particularly, in field settings (Wachner *et al.*, 2021). These results further emphasize the importance of the heterogeneity revolution (Bryan *et al.*, 2021): rather than making statements of nudges in general, each nudge could have different consequences, different mechanisms and different effects depending on context.

Second, we contribute to the literature on email use by showing that nudges, and especially bundles of nudges, may help to reduce email use, which presents a serious threat to employee well-being (Brown *et al.*, 2014; Reinke & Chamorro-Premuzic, 2014). We find nudges do have potential: employees are quite positive about them when it comes to autonomy and effectiveness, more so than policy instruments that are more commonly studied in the literature and used in organizations (Aguinis *et al.*, 2013; Handgraaf *et al.*, 2013; Tummers, 2019).

Third, our paper contributes methodologically, specifically on testing (behavioral) interventions. We introduced and redeveloped multiple ways in which nudges can be evaluated, prior to their implementation. In our pilot study, we assessed respondents' granular opinion on the nudges by distinguishing between feasibility, appropriateness, meaningfulness and effectiveness (Jordan *et al.*, 2019). In our survey experiment, we used a modified version of the Bayesian truth serum to counter social desirability bias (Prelec, 2004; John *et al.*, 2012). The results show that respondents are likely to overestimate their own compliance. At the same time, the truth serum tends to be conservative (John *et al.*, 2012), meaning that the true value is likely to lie in between. This redeveloped serum could be a useful tool for scholars to evaluate nudges or other interventions to illicit more truthful responses. Finally, the comparative evaluation of nudges with traditional interventions has shed light on where nudges are positioned in the realm of policy and managerial interventions in general.

This research also has practical implications for managers and public policy. For managers, our research strengthens the argument that managers could turn to nudges as a valid and low-cost alternative to traditional policy instruments. Our findings suggest that unlike traditional policy instruments that might undermine employee autonomy (such as limiting email access), nudges can be autonomy-preserving and effective and can be used for concrete organizational challenges like email use. Our study also has implications for public policy. First, maintaining employee well-being in healthcare is an urgent public policy challenge in countries across the world (e.g., Rotenstein *et al.*, 2018). Well-being among healthcare employee has been increasingly put under pressure through, among others, the COVID-19 crisis (Spoorthy *et al.*, 2020) and the aging workforce (Van Dalen *et al.*, 2010). We have developed nudges to reduce a prevalent stressor in healthcare employees' jobs: email use. Studies have shown that email use can have very negative consequences for employees during and outside of work (Jackson *et al.*, 2003; Kushlev & Dunn, 2015; Belkin *et al.*, 2020;

Giurge & Bohns, 2021). While it is unlikely that nudges will solve everything, we show how nudges can be part of efforts to contribute to improve the well-being of healthcare employee. Second, autonomy and effectiveness are critical issues in public policy. Scholars and practitioners have extensively debated whether nudges are a suitable policy instrument, including whether they are autonomy-preserving and effective (e.g., De Ridder *et al.*, 2020; Tummers, 2022). Although our results may be context-dependent (e.g., Andersson & Almqvist, 2022; discussed below), our study suggests nudges can be autonomy-preserving and effective.

Limitations

We want to highlight several conceptual and methodological limitations. First, all nudges shared similarities: they were text-based, infographics, and sent via SMS or intranet. This allows for better comparison between the nudges because we minimize confounding variables. However, they do not fully represent the spectrum of what nudges can be (Thaler & Sunstein, 2021). We can therefore only draw conclusions from the nudges we tested. Compare, for example, the study by Andersson & Almqvist (2022), who find that the Swedish public prefers information and subsidies – both traditional policy instruments – above nudges. Following the logic of the heterogeneity revolution, future research should assess to what extent other types of nudges are able to preserve autonomy and be effective in other contexts (Bryan *et al.*, 2021). Future research may also explore whether similar nudges, or bundles of nudges, could also be of help with different organizational challenges that affect well-being, like limiting work hours and maintaining a work-life balance (Pak *et al.*, 2022). Finally, while we have compared nudges to traditional policy instruments across multiple studies, future research can explore potential causal mechanisms that explain why certain nudges have varying effects on autonomy and effectiveness. For example, in our survey experiment, the opinion leader nudge scored highest on individual effectiveness. It is possible that this nudge is more effective because it uses role modeling behaviors and fosters reciprocity between leaders and followers (e.g., Decuyper & Schaufeli, 2020).

Second, in this study, we conceptualized autonomy as the extent to whether nudges guarantee agency. However, Vugts *et al.* (2020) argue that nudges may also influence freedom of choice and self-constitution, which are the other two conceptualizations of autonomy. The question of whether a nudge is strengthening or empowering autonomy depends not only on the nudge itself, but also on the conceptualization of autonomy one focuses on (Vugts *et al.*, 2020). More research is needed to better understand how nudges shape autonomy.

Third, our study presents both survey and quasi-field experimental evidence. The survey experiment does not measure actual behavior but intent. While intentions match behavior to some extent (e.g., Armitage & Conner, 2001; Hassan & Wright, 2020), a field experiment would introduce many aspects that a survey experiment lacks (Fishbein & Ajzen, 2010). The most important limitations of the quasi-experiment we used are the lack of a control group and randomized treatment allocation. Therefore, results may be biased by confounding variables (Shadish *et al.*, 2002; Grant & Wall, 2009). Although we have taken several measures to deal with

this, the effects of external events or elements unrelated to the treatment cannot be ruled out. Together with the limited amount of evidence in the quasi-field experiment, this constitutes a serious limitation. Also, the experimental period of 8 weeks is a considerable amount of time and a common timeline for work interventions (see Grant *et al.*, 2014). Yet this does not warrant any claims about the true long-term effects of the nudges. In general, while some nudges, like defaults (e.g., Venema *et al.*, 2018), can cause long-term effects, the long-term effects of nudges are insufficiently researched (Marchiori *et al.*, 2017). Regarding the response rate, only a minority of employees viewed the nudges. Still, while future research could address these limitations by designing randomized controlled trials (Gerber & Green, 2012), our quasi-experimental approach does have benefits by testing nudges in the field and measuring effects on actual email use. We concur with Grant & Wall (2009), who argue a quasi-experiment can be a method of action research, providing an opportunity for collaboration between researchers and practitioners to jointly tackle organizational challenges.

Conclusion

This paper provides a nuanced perspective toward one of the most applied and debated behavioral interventions: nudges. Our theoretical approach and empirical substantiation indicate that nudges can be designed to be both autonomy-preserving and effective. Going forward, scholars and practitioners can leverage these insights to maximize the potential of what nudges can do.

Supplementary material. To view supplementary material for this article, please visit <https://doi.org/10.1017/bpp.2023.18>.

Acknowledgements. The authors wish to thank the editor and anonymous reviewers for their helpful feedback. We thank Ulrike Weske and Jeroen Swanen for their supervision in the field experiment, Karen Leclercq, Roy Vosters, Iris Habraken and Maddie Houtstra for their cooperation in the field experiment, Enno Wigger and Jost Sieweke for their advice on the analyses and Arnold Bakker for his mentorship in developing this study. Earlier versions of this article were presented at the annual conferences of the International Research Society for Public Management (IRSPM) in 2022 and 2023. We thank the participants for their helpful comments and suggestions. This research was funded by IZZ, a healthcare employee collective in the Netherlands. Additionally, L.T. acknowledges funding from NWO Grant 016.VIDI.185.017. Except for the cooperation with IZZ on collecting the data, the funders had no role in the research.

References

- Aguinis, H., H. Joo and R. K. Gottfredson (2013), 'What monetary rewards can and cannot do: how to show employees the money', *Business Horizons*, 56(2): 241–249.
- Andersson, P. and G. Almqvist (2022), 'Carrots, sticks, sermons or nudges? Survey evidence of the Swedish general public's attitude towards different public policy tools', *Behavioural Public Policy*, 1–26. doi:10.1017/bpp.2022.31.
- Armitage, C. J. and M. Conner (2001), 'Efficacy of the theory of planned behaviour: a meta-analytic review', *British Journal of Social Psychology*, 40(4): 471–499. doi:10.1348/014466601164939.
- Balebako, R., P. G. Leon, H. Almuhiemedi, P. G. Kelley, J. Mugan, A. Acquisti and N. Sadeh-Konieczpol (2011), Nudging users towards privacy on mobile devices. PINC2011.
- Banerjee, S. and P. John (2021), 'Nudge plus: incorporating reflection into behavioral public policy', *Behavioural Public Policy*, 5717(332): 1–16.

- Banerjee, S. and P. John (2023), 'Nudge Plus: Putting Citizens at the Heart of Behavioural Public Policy', in L. Reich and C. Sunstein (eds), *Forthcoming in the Research Handbook on Nudges and Society*, Cheltenham: Edward Elgar Publishing.
- Barley, S. R., D. E. Meyerson and S. Grodal (2011), 'E-mail as a source and symbol of stress', *Organization Science*, **22**(4): 887–906.
- Belkin, L. Y., W. J. Becker and S. A. Conroy (2020), 'The invisible leash: the impact of organizational expectations for email monitoring after-hours on employee resources, well-being, and turnover intentions', *Group and Organization Management*, **45**(5): 709–740.
- Bozeman, B. and J. Youtie (2020), 'Robotic bureaucracy: administrative burden and red tape in university research', *Public Administration Review*, **80**(1): 157–162.
- Bradt, J. (2022), 'Comparing the effects of behaviorally informed interventions on flood insurance demand: an experimental analysis of 'boosts' and 'nudges'', *Behavioural Public Policy*, **6**(3): 485–515.
- Brown, R., J. Duck and N. Jimmieson (2014), 'E-mail in the workplace: the role of stress appraisals and normative response pressure in the relationship between e-mail stressors and employee strain', *International Journal of Stress Management*, **21**(4): 325–347.
- Bruns, H., E. Kantorowicz-Reznichenko, K. Klement, M. L. Jonsson and B. Rahali (2018), 'Can nudges be transparent and yet effective?' *Journal of Economic Psychology*, **65**: 41–59.
- Bryan, C. J., E. Tipton and D. S. Yeager (2021), 'Behavioural science is unlikely to change the world without a heterogeneity revolution', *Nature Human Behaviour*, **5**(8): 980–989. doi:10.1038/s41562-021-01143-3.
- Caraban, A., E. Karapanos, D. Gonçalves and P. Campos (2019), '23 ways to nudge: a review of technology-mediated nudging in human-computer interaction', In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems, pp. 1–15.
- Cecchinato, M. E., J. Bird and A. L. Cox (2014), 'Personalised email tools: a solution to email overload?' In CHI'14 Workshop: Personalised Behaviour Change Technologies. ACM Conference on Human Factors in Computing Systems (CHI).
- Dabbish, L. A. and R. E. Kraut (2006), 'Email overload at work: an analysis of factors associated with email strain', In Proceedings of the 2006 20th Anniversary Conference on Computer Supported Cooperative Work, pp. 431–440.
- Davidai, S. and E. Shafir (2020), 'Are 'nudges' getting a fair shot? Joint versus separate evaluation', *Behavioural Public Policy*, **4**(3): 273–291. doi:10.1017/bpp.2018.9.
- Decuyper, A. and W. Schaufeli (2020), 'Leadership and work engagement: exploring explanatory mechanisms', *German Journal of Human Resource Management*, **34**(1): 69–95.
- Dellavigna, S. and E. Linos (2022), 'RCTs to scale: comprehensive evidence from two nudge units', *Econometrica*, **90**(1): 81–116.
- DellaVigna, S., D. Pope and E. Vivalt (2019), 'Predict science to improve science', *Science*, **366**(6464): 428–429.
- De Ridder, D., J. Feitsma, M. Van den Hoven, F. Kroese, T. Schillemans, M. Verweij and E. De Vet (2020), 'Simple nudges that are not so easy', *Behavioural Public Policy*, 1–19.
- Fishbein, M. and I. Ajzen (2010), *Predicting and Changing Behavior: The Reasoned Action Approach*, New York: Psychology Press.
- Frank, M. R., M. Cebrian, G. Pickard and I. Rahwan (2017), 'Validating Bayesian truth serum in large-scale online human experiments', *PLoS One*, **12**(5): e0177385.
- Gerber, A. S. and D. P. Green (2012), *Field Experiments: Design, Analysis, and Interpretation*. New York City: WW Norton.
- Giurge, L. M. and V. K. Bohns (2021), 'You don't need to answer right away! Receivers overestimate how quickly senders expect responses to non-urgent work emails', *Organizational Behavior and Human Decision Processes*, **167**: 114–128.
- Glickman, M. E., S. R. Rao and M. R. Schultz (2014), 'False discovery rate control is a recommended alternative to Bonferroni-type adjustments in health studies', *Journal of Clinical Epidemiology*, **67**(8): 850–857.
- Gorgievski, M. J., P. Peeters, E. F. Rietzschel and T. Bipp (2016), 'Betrouwbaarheid en Validiteit van de Nederlandse vertaling van de Work Design Questionnaire', *Gedrag en Organisatie*, **29**(3): 273–301.
- Grant, A. M. and T. D. Wall (2009), 'The neglected science and art of quasi-experimentation: why-to, when-to, and how-to advice for organizational researchers', *Organizational Research Methods*, **12**(4): 653–686.

- Grant, A. M., J. M. Berg and D. M. Cable (2014), 'Job titles as identity badges: how self-reflective titles can reduce emotional exhaustion', *Academy of Management Journal*, **57**(4): 1201–1225.
- Gupta, A., R. Sharda and R. A. Greve (2011), 'You've got email! Does it really matter to process emails now or later?' *Information Systems Frontiers*, **13**(5): 637–653.
- Hagman, W., A. Erlansddon, S. Dickert, G. Tinghög and D. Västfjäll (2022), 'The effect of paternalistic alternatives on attitudes toward default nudges', *Behavioural Public Policy*, **6**(1): 95–118. doi:10.1017/bpp.2019.17.
- Handgraaf, M. J., M. A. V. L. De Jeude and K. C. Appelt (2013), 'Public praise vs. private pay: effects of rewards on energy conservation in the workplace', *Ecological Economics*, **86**: 86–92.
- Hansen, P. G. (2016), 'The definition of nudge and libertarian paternalism: does the hand fit the glove?' *European Journal of Risk Regulation*, **7**(1): 155–174.
- Hansen, P. G. and A. M. Jespersen (2013), 'Nudge and the manipulation of choice', *European Journal of Risk Regulation*, **4**(1): 3–28.
- Hassan, S. and B. E. Wright (2020), 'The behavioral public administration movement: a critical reflection', *Public Administration Review*, **80**(1): 163–167. doi:10.1111/puar.13130.
- Hausman, D. M. and B. Welch (2010), 'Debate: to nudge or not to nudge', *Journal of Political Philosophy*, **18**(1): 123–36.
- Hertwig, R. (2017), 'When to consider boosting: some rules for policy-makers', *Behavioural Public Policy*, **1**(2): 143–61.
- Hertwig, R. and T. Grüne-Yanoff (2017), 'Nudging and boosting: steering or empowering good decisions', *Perspectives on Psychological Science*, **12**(6): 973–86.
- Jackson, T., R. Dawson and D. Wilson (2003), 'Reducing the effect of email interruptions on employees', *International Journal of Information Management*, **23**(1): 55–65.
- John, P., G. Smith and G. Stoker (2009), 'Nudge nudge, think think: two strategies for changing civic behaviour', *The Political Quarterly*, **80**(3): 361–370.
- John, L. K., G. Loewenstein and D. Prelec (2012), 'Measuring the prevalence of questionable research practices with incentives for truth telling', *Psychological Science*, **23**(5): 524–532.
- John, P., A. Martin and G. Mikołajczak (2022), 'Support for behavioral nudges versus alternative policy instruments and their perceived fairness and efficacy', *Regulation & Governance*. doi:10.1111/rego.12460.
- Jordan, Z., C. Lockwood, Z. Munn and E. Aromataris (2019), 'The updated Joanna Briggs institute model of evidence-based healthcare', *JBI Evidence Implementation*, **17**(1): 58–71.
- Kahneman, D. (2011), *Thinking, Fast and Slow*. New York City: Macmillan.
- Krueger, C. and L. Tian (2004), 'A comparison of the general linear mixed model and repeated measures ANOVA using a dataset with multiple missing data points', *Biological Research for Nursing*, (2): 151–157. doi:10.1177/1099800404267682. PMID: 15388912.
- Kushlev, K. and E. W. Dunn (2015), 'Checking email less frequently reduces stress', *Computers in Human Behavior*, **43**: 220–228.
- Lin, Y., M. Osman and R. Ashcroft (2017), 'Nudge: concept, effectiveness, and ethics', *Basic and Applied Social Psychology*, **39**(6): 293–306.
- Maier, M., F. Bartoš, T. D. Stanley, D. R. Shanks, A. J. Harris and E. J. Wagenmakers (2022), 'No evidence for nudging after adjusting for publication bias', *Proceedings of the National Academy of Sciences*, **119**(31): e2200300119.
- Marchiori, D. R., M. A. Adriaanse and D. T. De Ridder (2017), 'Unresolved questions in nudging research: putting the psychology back in nudging', *Social and Personality Psychology Compass*, **11**(1): e12297.
- Mertens, S., M. Herberz, U. J. Hahnel and T. Brosch (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.
- Michaelsen, P., L. O. Johansson and M. Hedesström (2021), 'Experiencing default nudges: autonomy, manipulation, and choice-satisfaction as judged by people themselves', *Behavioural Public Policy*, 1–22. doi:10.1017/bpp.2021.5.
- Morgeson, F. P. and S. E. Humphrey (2006), 'The Work Design Questionnaire (WDQ): developing and validating a comprehensive measure for assessing job design and the nature of work', *Journal of Applied Psychology*, **91**(6): 1321.
- Münscher, R., M. Vetter and T. Scheuerle (2016), 'A review and taxonomy of choice architecture techniques', *Journal of Behavioral Decision Making*, **29**(5): 511–524.

- Nagtegaal, R., L. Tummers, M. Noordegraaf and V. Bekkers (2019), 'Nudging healthcare professionals towards evidence-based medicine: a systematic scoping review', *Journal of Behavioral Public Administration*, **2**: 2.
- Nagtegaal, R., L. Tummers, M. Noordegraaf and V. Bekkers (2020), 'Designing to debias: measuring and reducing public managers' anchoring bias', *Public Administration Review*, **80**(4): 565–576.
- Osman, M. (2020), '*Backfiring, Reactance, Boomerang, Spillovers, and Rebound Effects: Can We Learn Anything from Examples Where Nudges Do the Opposite of What They Intended?*' Retrieved from: <https://psyarxiv.com/ae756/download?format=pdf>
- Pak, S., A. Kramer, Y. Lee and K. J. Kim (2022), 'The impact of work hours on work-to-family enrichment and conflict through energy processes: a meta-analysis', *Journal of Organizational Behavior*, **43**(4): 709–743.
- Prelec, D. (2004), 'A Bayesian truth serum for subjective data', *Science*, **306**: 462–466.
- Reijula, S. and R. Hertwig (2022), 'Self-nudging and the citizen choice architect', *Behavioural Public Policy*, **6**(1): 119–149.
- Reinke, K. and T. Chamorro-Premuzic (2014), 'When email use gets out of control: understanding the relationship between personality and email overload and their impact on burnout and work engagement', *Computers in Human Behavior*, **36**: 502–509.
- Reith, T. P. (2018), 'Burnout in United States healthcare professionals: a narrative review', *Cureus*, **10**(12): e3681.
- Rosen, C. C., L. S. Simon, R. S. Gajendran, R. E. Johnson, H. W. Lee and S.-H.-J. Lin (2019), 'Boxed in by your inbox: implications of daily e-mail demands for managers' leadership behaviors', *Journal of Applied Psychology*, **104**(1): 19.
- Rotenstein, L. S., M. Torre, M. A. Ramos, R. C. Rosales, C. Guille, S. Sen and D. A. Mata (2018), 'Prevalence of burnout among physicians: a systematic review', *Jama*, **320**(11): 1131–1150.
- Ryan, R. M. and E. L. Deci (2017), *Self-determination Theory: Basic Psychological Needs in Motivation Development and Wellness*, New York: Guilford Publishing.
- Schaufeli, W. B., A. Shimazu, J. Hakanen, M. Salanova and H. De Witte (2019), 'An ultra-short measure for work engagement: the UWES-3 validation across five countries', *European Journal of Psychological Assessment*, **35**(4): 577.
- Schoenegger, P. (2023), 'Experimental philosophy and the incentivisation challenge: a proposed application of the Bayesian Truth Serum', *Review of Philosophy and Psychology*, **14**(1): 295–320.
- Shadish, W. R., T. D. Cook and D. T. Campbell (2002), *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*, Houghton: Mifflin and Company.
- Simon, H. A. (1955), 'A behavioral model of rational choice', *Quarterly Journal of Economics*, **69**(1): 99–118.
- Spoorthy, M. S., S. K. Pratapa and S. Mahant (2020), 'Mental health problems faced by healthcare workers due to the COVID-19 pandemic – a review', *Asian Journal of Psychiatry*, **51**: 102119.
- Sumecki, D., M. Chipulu and U. Ojiako (2011), 'Email overload: exploring the moderating role of the perception of email as a 'business critical' tool', *International Journal of Information Management*, **31**(5): 407–414.
- Sunstein, C. R. (2016), 'People prefer system 2 nudges (kind of)', *Duke LJ*, **66**: 121.
- Szaszi, B., A. Higney, A. Charlton, A. Gelman, I. Ziano, B. Aczel and E. Tipton (2022), 'No reason to expect large and consistent effects of nudge interventions', *Proceedings of the National Academy of Sciences*, **119** (31): e2200732119.
- Thaler, R. H. and C. R. Sunstein (2008), *Nudge: Improving Decisions about Health, Wealth, and Happiness*, New Haven, CT: Yale University Press.
- Thaler, R. H. and C. R. Sunstein (2021), *Nudge: The Final Edition*, London: Penguin Books.
- Tummers, L. (2019), 'Public policy and behavior change', *Public Administration Review*, **79**(6): 925–930.
- Tummers, L. (2022), 'Nudge in the news: ethics, effects, and support of nudges', *Public Administration Review*, <https://doi.org/10.1111/puar.13584>.
- Tversky, A. and D. Kahneman (1974), 'Judgment under uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty', *Science*, **185**(4157): 1124–1131.
- Tversky, A. and D. Kahneman (1981), 'The framing of decisions and the psychology of choice', *Science*, **211** (4481): 453–458.
- Valente, T. W. and P. Pumpuang (2007), 'Identifying opinion leaders to promote behavior change', *Health Education and Behavior*, **34**: 881–896. doi:10.1177/1090198106297855.

- Van Dalen, H. P., K. Henkens and J. Schippers (2010), 'How do employers cope with an ageing workforce? Views from employers and employees', *Demographic Research*, **22**: 1015–1036.
- Van de Schoot, R., S. D. Winter, E. Griffioen, S. Grimmelikhuijsen, I. Arts, D. Veen and L. G. Tummers (2021), 'The use of questionable research practices to survive in academia examined with expert elicitation, prior-data conflicts, Bayes factors for replication effects, and the Bayes truth serum', *Frontiers in Psychology*, **5241**, <https://doi.org/10.3389/fpsyg.2021.621547>.
- Van Roekel, H., I. van der Fels, A. B. Bakker and L. G. Tummers (2020), 'Healthcare workers who work with COVID-19 patients are more physically exhausted and have more sleep problems', *Frontiers in Psychology*, **11**: 3843.
- Van Roekel, H., L. Giurge, C. Schott and L. Tummers (2022a). Decreasing email overload with autonomy-preserving nudges. *Preregistration of the Pilot Study*. doi:10.17605/OSF.IO/YSGJ2
- Van Roekel, H., L. Giurge, C. Schott and L. Tummers (2022b), 'Decreasing email overload with autonomy-preserving nudges', *Preregistration of the Survey Study*. doi:10.17605/OSF.IO/89GBX
- Van Roekel, H., L. Giurge, C. Schott and L. Tummers (2022c), 'Decreasing email overload with autonomy-preserving nudges', *Preregistration of the Field Study*. doi:10.17605/OSF.IO/V5YC4
- Van Roekel, H., J. Reinhard and S. Grimmelikhuijsen (2022d), 'Improving hand hygiene in hospitals: comparing the effect of a nudge and a boost on protocol compliance', *Behavioural Public Policy*, **6**(1): 52–74.
- Venema, T. A., F. M. Kroese and D. T. De Ridder (2018), 'I'm still standing: a longitudinal study on the effect of a default nudge', *Psychology & Health*, **33**(5): 669–681.
- Vitek, V. and T. Syed Shah (2019) 'Implementing a Nudge to Prevent Email Phishing'. Retrieved from: <https://www.diva-portal.org/smash/get/diva2:1351517/FULLTEXT01.pdf>.
- Vugts, A., M. Van Den Hoven, E. De Vet and M. Verweij (2020), 'How autonomy is understood in discussions on the ethics of nudging', *Behavioural Public Policy*, **4**(1): 108–123.
- Wachner, J., M. A. Adriaanse and D. T. De Ridder (2020), 'And how would that make you feel? How people expect nudges to influence their sense of autonomy', *Frontiers in Psychology*, **11**: 607894.
- Wachner, J., M. A. Adriaanse and D. T. De Ridder (2021), 'The effect of nudges on autonomy in hypothetical and real life settings', *PLoS ONE*, **16**(8): e0256124.
- Weaver, R. and D. Prelec (2013), 'Creating truth-telling incentives with the Bayesian truth serum', *Journal of Marketing Research*, **50**(3): 289–302.
- Wilkinson, T. M. (2013), 'Nudging and manipulation', *Political Studies*, **61**(2): 341–355.

Cite this article: van Roekel H, Giurge LM, Schott C, Tummers L (2023). Nudges can be both autonomy-preserving and effective: evidence from a survey and quasi-field experiment. *Behavioural Public Policy* 1–24. <https://doi.org/10.1017/bpp.2023.18>