

Matthew Amengual  
University of Oxford  
matthew.amengual@sbs.ox.ac.uk

Evan P. Apfelbaum  
Boston University  
epa1@bu.edu

Accepted at *Management Science*, May 7, 2020

## ABSTRACT

When organizations want their employees to adopt behaviors that advance prosocial and instrumental aims, which motive should they express? A groundswell of recent work suggests that highlighting prosocial actions will inspire and motivate employees. Building on this work, we embed a field experiment in the context of an organizational change initiative (Study 1). A large university sought to change the behavior of administrative employees who purchase office supplies, encouraging them to place orders of at least \$50, referred to as “bundling.” We exploit the fact that the organization could justify the same behavior in contrasting ways. We randomly assign employees to view either a prosocial (“limiting pollution”), instrumental (“limiting costs”), or mixed motive (“limiting pollution and limiting costs”) for caring about bundling each time they access the organization’s procurement system. We then evaluate changes in employees’ behavior by comparing a six-month baseline to a six-month experimental period, covering 10,169 purchases in 556 offices. Contrary to expectations from related research, we observe that the instrumental motive led to significantly more bundling than the prosocial motive. Two follow-up vignette experiments probe theoretical mechanisms. They indicate that an instrumental motive seems more genuine (i.e., reflecting the organization’s true motive) than a prosocial motive (Study 2), and that seeming genuine increases individuals’ intention to bundle (Study 3). This research demonstrates that prosocial justifications can be less effective than instrumental ones and suggests that perceptions of genuineness may shape the effectiveness of behavioral change efforts in organizations.

## Acknowledgements

We thank Hannah Birnbaum, Greg Distelhorst, Valerie Karplus, Mike Norton, Todd Rodgers, John Sterman, Michael Toffel, and Maurizio Zollo, as well as colleagues at Boston University and University of Oxford for their helpful comments and suggestions. We also thank staff at the university where the field experiment took place for collaborating with us on this project. A previous version of this paper was presented at the Alliance for Research on Corporate Sustainability conference, held at MIT in 2018.

## 1. Introduction

Organizations of all kinds, from businesses and universities to governments and NGOs, call on their employees to change routine behaviors to streamline organizational functioning.

Oftentimes, the new behaviors that organizations want their employees to adopt advance goals that are both prosocial (e.g., environmental protection) and instrumental (e.g., cost reduction).

Organizations are keen to identify such “win-win” opportunities, but they face a predicament in terms of how to express their interest in them to their employees. Indeed, a fundamental question organizations increasingly face is how to communicate *why* they care about the behaviors they want their employees to adopt—through an expression of prosocial, instrumental, or mixed motives.

On one hand, burgeoning literatures on prosocial values and activities in psychology, economics, and management suggest that organizations should emphasize the prosocial dimension of actions in order to inspire employee engagement and effort. For example, research has shown that providing non-monetary, prosocial “incentives” (e.g., a donation received by a charity of an employees’ choice) or instilling a sense of meaning and purpose to work (e.g., highlighting an organizations’ philanthropic efforts) can increase motivation, performance, and retention, among other positive outcomes (Bode, Singh and Rogan 2015; Burbano 2016; Cassar 2018; Grant 2008; Ryan and Deci 2000; Wrzesniewski et al 2014). The implication of this research is that organizations can marshal prosocial actions and purpose to motivate employees to adopt new behaviors, sometimes in ways that transcend traditional financial incentives. The prediction based on this body of work, therefore, is that a prosocial motive will be relatively effective for prompting employee uptake of desired behaviors.

On the other hand, alternative theories suggest that a prosocial motive may not be effective. Though generally not considered in this area of inquiry, another characteristic of an organization’s stated motive may also influence individuals’ reactions to it; namely, how genuine it seems. By

genuine, we mean the degree to which a motive is perceived to be an honest and authentic reflection of the organization's true interests (see Forehand and Grier 2003; Radoynovska and King 2019; Ryan and Ryan 2019), or more pragmatically, if individuals think that what the organization *says* it cares about is what it actually cares about. The implication of this perspective is twofold. First, it suggests that employees in many settings will perceive an instrumental motive to be more genuine than a prosocial motive, in part, because as compared with prosocial motives, some degree of instrumental motives (e.g., reducing costs) are widely assumed in organizational contexts. Second, this perspective suggests that an instrumental motive may be more effective at encouraging employees to adopt new behaviors than a prosocial motive because individuals tend to comply with relatively straightforward, transparent requests to change behavior (Feiler et al. 2012; Forehand and Grier 2003; Osterhus 1997; Williams et al. 2004).

Amid this backdrop of contrasting theoretical possibilities, we examine the impact of an organization's expression of prosocial or instrumental motives for a new behavior on employees' adoption of it. We do so by deploying a field experiment in the context of a real-world problem a large university in the northeastern U.S. sought to address: the practice among the university's administrative employees of purchasing small items (e.g., paper, pens, cleaning products) in separate orders. The university wanted to encourage these employees to change their behavior—to put in the additional effort of planning and coordinating orders so that they reach a value of at least \$50, which we refer to as “bundling.” The university had determined that doing so would reduce environmental impacts and shipping costs in the long run. We programmed a message to present to administrative employees as they navigated to the university's supplier website to purchase goods. We manipulated only the content of the text that indicated *why* the university cared about bundling (where “X” designates the name of the university): “X cares about limiting pollution” (prosocial motive), “X

cares about limiting costs” (instrumental motive), and “X cares about limiting pollution and limiting costs” (mixed motive; order counterbalanced).

While a growing body of related research suggests that expressing prosocial motives will be particularly effective for stimulating behavioral change, we find that the instrumental motive is significantly more effective at getting employees to bundle. We then devise two follow-up vignette experiments with separate online samples to evaluate plausible theoretical processes. In the first, we find that the instrumental motive seems more genuine to individuals than the prosocial motive; in the second, we find that seeming genuine drives individuals’ intention to bundle (in both the instrumental and prosocial conditions).

This paper makes important contributions to theory and managerial practice. Our findings stand in stark contrast to predictions from literatures in psychology, economics, and management on prosocial appeals to employees. A clear prediction that emerges from these related streams of research is that highlighting the prosocial virtues of one’s work, products, and organization is highly effective for engaging employees. While prosocial appeals may have the capacity to inspire purpose and motivate, our results suggest that when an organization trumpets an exclusively prosocial motive for its interest in a new behavior that also advances instrumental goals, it may fall flat.

Furthermore, the findings from our vignette experiments advance theory by suggesting one reason why it may be more effective for organizations to express instrumental than prosocial motives. Namely, organizations may seem more genuine when they acknowledge instrumental motives, and consequently, employees will be more inclined to change their behavior in response to such requests. Our findings suggest that prosocial motives also have the capacity to stimulate the desired uptake of behaviors—as a great deal of previous research suggests—but only when employees perceive these motives to be a genuine. In this way, our research highlights an

underappreciated dimension of prosocial and instrumental appeals that may shape their effectiveness.

Our findings have direct, and potentially substantial, implications for public and private organizations seeking to promote a wide range of activities that can be justified on both prosocial and instrumental grounds (e.g., promoting diversity, Apfelbaum et al. 2016; supporting local economic development, Franks et al. 2014). Our study contributes an important piece of behavioral evidence that can help guide senior managers' and policy makers' efforts to motivate employees to undertake behaviors that advance these initiatives. While prosocial values may have the potential to inspire and motivate under certain conditions, in many organizational contexts, it may simply be more effective to acknowledge instrumental concerns. Our research suggests that leaders must choose their words carefully when communicating the importance of behaviors that contribute to both prosocial and instrumental goals. When they stray from employees' sense of the organization's true motives, words designed to pull at heartstrings may ring hollow.

## **2. Literature review**

Organizations regularly call on their employees to adopt new behaviors that can improve organizational functioning. While organizations have many ways of directing behavior, they often seek to do so by expressing their values (Mayer et al. 2019). That is, organizations communicate the underlying motive, rationale, and importance of a new behavior to promote its adoption. This lever for encouraging behavioral change is distinct from traditional incentives and less formalized understandings of reciprocity in which employees expect something in exchange for performing the behavior (Robinson and Rousseau 1994). Indeed, for the class of organizational situations in which the use of incentives is not desirable or viable (Gneezy et al. 2011), the motives organizations express represent a key lever for influencing employee behaviors. This may be particularly so when the behavior the organization promotes can be justified as advancing both prosocial objectives (e.g.,

pollution reduction) and instrumental objectives (e.g., cost reduction). In these contexts, leadership must determine how to communicate *why* they care—be it an expression of instrumental, prosocial, or mixed motives. While previous research has not directly examined the question of which of these motives most effectively spurs adoption of new behaviors, relevant and interrelated streams of research in psychology, economics, management, and marketing offer theoretical guidance as to how employees may respond to these motives. From these extant literatures, we derive contrasting theoretical perspectives, leading to distinct predictions about the motive that will be most effective for encouraging employees to change their behavior.

The first theoretical account, buoyed by a broad wave of recent research, suggests that prosocial motives may be particularly well suited to inspire and motivate employees. This relatively new stream of research emerged, in part, as a response to the longstanding monolithic view of incentives and instrumentality as the means to motivate workers. The psychological underpinning of this work is that individuals seek to maintain a favorable view of themselves (i.e., “I am a moral person”; Bolderdijk et al. 2013), and by extension, of the groups to which they belong (“I am a member of a virtuous organization”; Ellemers et al. 2011; Griskevicius et al. 2010). Employees therefore strive to affirm their view of themselves vis-à-vis their work on behalf of organizations that are committed to prosocial goals (Cassar and Meier 2018). Accordingly, this research posits that when employees understand an organization’s purpose as extending beyond purely instrumental goals, it can powerfully shape their behavior (Cassar and Meier 2018; Henderson and Van Den Steen 2015; Gartenberg et al. 2019).

A body of research spanning psychology and management demonstrates that employees (and individuals, more generally) are often driven by prosocial values. This research generally finds that it is more persuasive to appeal to individuals’ prosocial values (e.g., regarding the environment or public health) than to their instrumental values (Ryan and Deci 2000; Bolderdijk et al. 2013;

Evans et al. 2013). For example, recent work suggests that managers can more effectively persuade others in their organization by framing their appeals in moral as opposed to instrumental terms (Mayer et al. 2019). Related research has found that attaching a broader prosocial purpose and meaning to work can inspire engagement, a sense of belonging, commitment, and identification with one's organization (Bartel 2001; Bauman and Skitka 2012; Brammer et al. 2007; Grant 2008, 2012).

Further supporting this interpretation is evidence from a number of studies indicating that organizations can motivate workers by touting their prosocial actions. For example, a large number of studies provide evidence that prosocial incentives—such as linking a participant's work to charitable giving on behalf of the organization—increase participants' effort (Anik et al. 2013; Burbano 2019; Cassar 2018; Chandler and Kapelner 2013; Charness et al. 2016; Imas 2014; Tonin and Vlassopoulos 2015; but see Gosnell et al. 2019). Prosocially-imbued jobs may attract more prospective employees who ultimately become more productive and work longer hours (Hedblom et al. 2016). Evidence also suggests that individuals will work for lower compensation when told that their work has a prosocial mission (Burbano 2016; Frank and Smith 2016). Finally, indirect support also comes from a complementary area of research on firm-level strategy and CSR. This work provides evidence that when jobs incorporate prosocial initiatives, employees are less likely to turnover (Bode et al. 2015), and congruently, that firms use CSR strategically to boost employee engagement (Flammer and Luo 2017) and enhance their human capital (Burbano et al. 2018).

Overall, this web of indirect evidence from multiple fields and disciplines converges on the idea that an organization's prosocial actions can inspire and motivate employees in ways that instrumental objectives cannot. The prediction that flows from this work therefore is that to encourage employees to adopt a behavior that advances both prosocial and instrumental aims, it will be more effective for organizations to highlight their prosocial versus instrumental motives.

Given this past research, it would be reasonable for managers to expect that when they have the opportunity to frame a behavior as being prosocial, they should do so. Yet an alternative theoretical perspective provides reason to weigh a different characteristic of an organization's expressed motive: namely, whether it seems genuine, as in a sincere and authentic reflection of its true interests (Forehand and Grier 2003; Radoynovska and King 2019). Consider that when an organization expresses its motives (e.g. "here's why we care about behavior X"), employees become observers of how the organization has framed that behavior in accordance with a specific portrayal of its objectives. Employees are likely well aware of the contrasting prosocial and instrumental reasons the organization may have for advocating a given behavior, and given this ambiguity, will weigh the substance of their organization's stated motive carefully. Namely, they will consider whether the expressed motive seems to be a credible account of why the organization wants them to undertake the new behavior (i.e., "is what the organization says it cares about what it *really* cares about?").

According to this theoretical account, the perceived genuineness of an organization's motive can influence employees' willingness to change their behavior on its behalf. When an organization's stated motives seem genuine, employees should be more likely to perform the desired behavior (Osterhus 1997). Conversely, when employees sense that an expressed motive is disingenuous—for example, that it is an attempt to influence them or mask an ulterior motive—they should be less moved to change their behavior (Feiler et al. 2012; Williams et al. 2004).

Organizations that acknowledge an instrumental motive are likely to seem honest and transparent. Consider that unlike prosocial goals, some degree of instrumental goals (e.g., reducing costs) are integral to most organizations, and thus employees likely assume they play some role in management decisions. Moreover, highlighting an instrumental motive for a behavior that also furthers prosocial ends is unlikely to be viewed as a strategic form of impression management as, if



anything, this admission risks making an organization seem self-interested or callous (Berman et al. 2015; Newman and Cain 2014). Contrariwise, employees may question whether an organization's use of prosocial motives to promote the same behavior is genuine, unless they are privy to clear and compelling evidence supporting the credibility of this claim. And while we recognize that purely prosocial claims may seem genuine in select organizational contexts, these contexts are likely more the exception than the rule. In general, this theory suggests that employees will often question the genuineness of an organization that expresses purely prosocial motives for a behavior that also serves instrumental objectives (Barnett 2019; Bénabou and Tirole 2006; Hahl 2016; Makov and Newman 2016; Newman and Cain 2014). This argument is congruent with studies of consumer behavior, as Forehand and Grier (2003) note in discussing consumer skepticism to how marketers frame CSR: "...skepticism toward a firm is driven not simply by beliefs that the firm's motives are self-serving but rather by the perception that the firm is being deceptive about its true motives" (p. 350). This theoretical account thus leads to the prediction that employees will perceive an organization as more genuine when it expresses an instrumental (versus prosocial) motive for a win-win behavior. And because motives that seem genuine should compel action, it further suggests that employees will be more inclined to change their behavior in response to an instrumental versus prosocial motive.

This account is distinct from, and importantly extends, existing theories of why an organization's prosocial claims may be of limited benefit. For example, research suggests that prosocial claims are less effective when the workers targeted do not value the prosocial goal at stake (e.g., people who tend not to give money to charity are less apt to be motivated by an organization touting its philanthropy; Burbano 2019; Tonin and Vlassopoulos 2015). The perceived genuineness account predicts that even employees who value the prosocial end may remain unmoved by an

organization's expression of prosocial motives if they perceive it to be an inauthentic attempt to advance an instrumental goal.

Relatedly, recent scholarship suggests that prosocial incentives are ineffective at motivating workers when they are discernably self-serving for the organization (e.g., charitable donations are only made if workers help drive profits). The idea is that “employees use prosocial incentives as a signal of the firm’s type” (Cassar and Meier 2017, p.2) to infer whether it is benevolent and likely to treat them fairly (e.g., with respect to compensating them; Burbano 2016). Making a donation conditional on workers’ performance may thus be ineffective at motivating employees because it sends a signal that the organization will not treat them well in the future (Cassar and Meier 2017). Our theory broadens the context and processes by which prosocial actions can be ineffective. We suggest that, even in the absence of an explicit performance-conditionality, the simple expression of a prosocial motive for a behavior that also advances instrumental aims can fall flat. We propose that one reason this can be the case is that employees tend to change their behaviors when they believe their organization’s stated motives are true. Our interpretation is thus highly relevant to traditional workplace contexts in which employees’ everyday experiences provide myriad clues and opportunities to learn about their organization’s actual interests. In these contexts, employees can draw on a deep reservoir of aggregated experiences to inform their impression of their organization’s true goals and priorities. By contrast, impressions of an employer’s “type” may be most pertinent to contexts in which individuals have limited knowledge of an organization (e.g., prospective employees, gig workers).

Finally, though not the primary research focus, one may question the effects of combining prosocial and instrumental motives. If an organization’s use of a prosocial motive to explain its interest in a behavior inspires employees and calls to mind a broader purpose, then one may expect the addition of an instrumental motive to interfere with, and thus reduce the efficacy of, this

message. This prediction flows from the results of a number of experiments in psychology and economics that show instrumental motives can reduce (or “crowd out”) motivation on tasks that are personally important to individuals (Deci et al. 1999; Frey and Oberholzer-Gee 1997; Lepper et al. 1973; Wrzesniewski et al. 2014). However, if expressing why an organization cares about a behavior leads employees to consider the genuineness of these motives and these perceptions influence uptake of the behavior, then one may expect expressing mixed motives to be more effective than a prosocial motive alone. This prediction flows from the fact that the mixed motive acknowledges some instrumental interest on the part of the organization, which may seem more genuine, thereby increasing employees’ willingness to change their behavior.

### **3. Study 1: The Field Experiment**

**3.1. Context and Design.** We devised an experiment in the context of a real-world problem a large university in the northeastern U.S. sought to address: the common practice among the university’s administrative employees of purchasing small items (e.g., papers, pens, cleaning products) separately. The university determined that these orders were wasteful in terms of cost and environmental impact. The university wanted to encourage these employees to change their behavior: to increase the proportion of orders over \$50, or “bundle.”

The organization identified bundling as a behavior that would advance both a prosocial aim (i.e., limiting pollution) and an instrumental aim (i.e., limiting costs) in the long run. These contrasting prosocial and instrumental motives for the *same* behavior make it an ideal candidate for experimental manipulation and investigation. That is, rather than link the key behavior the organization sought to change to a separate, consequent prosocial incentive as in past research (e.g., a donation), we are able to manipulate the (prosocial or instrumental) motive for the target behavior itself. We designed the field experiment to evaluate which expressed motive for bundling—prosocial, instrumental, or both—would be most effective at getting employees to adopt this

behavior. The setting allowed for a “natural” field experiment (List 2009): it was timed and tailored to minimally interfere with an ongoing change effort the university had undertaken to increase the proportion of orders over \$50, and to do so using communication rather than incentives. We were able to embed messages communicating the university’s stated motive for bundling as employees navigated to the internal website where they purchase office supplies. We later obtained corresponding data from existing university administrative systems that track procurement. These features allowed us to manipulate the university’s stated motives and measure behavior without participants being aware that they were a part of a study.

The nature of bundling in this organizational context made it further advantageous for theoretical investigation. Bundling represented a “discretionary” or organizational citizenship behavior (Van Dyne et al. 1994)—it was voluntary, decoupled from incentives, and unmonitored by employees’ supervisors. Bundling did not factor into employees’ salary or performance evaluations (nor did the costs of ordering office supplies or sustainability goals, more generally).<sup>1</sup> Bundling was not part of employees’ formal job requirements, yet it required effort in the form of coordination and planning. When stock of an office supply runs low, or when a needed product is not in the supply cabinet, the simplest way for staff to satisfy demand for supplies is to purchase them as needed, regardless of whether orders meet the \$50 threshold. Bundling made this more complicated for staff because it required them to anticipate demand and proactively monitor stocks of supplies. Finally, it is noteworthy that shipping costs are part of a university-wide contract with the supplier (akin to electricity, or library subscriptions to academic journals), and thus they are not a relevant factor in office-level procurement decisions. Together, these context-specific features of bundling

---

<sup>1</sup> We obtained the formal criteria for performance evaluations for staff in one of the university’s largest departments. Bundling is not included in the criteria. More broadly, ordering office supplies (of which bundling is one aspect) played a minimal role in performance evaluations, and to the extent that it did play a role, it was based on whether needed supplies were immediately available to faculty and staff.

provide a unique opportunity to isolate the influence of an organization's stated motives for bundling on employees' uptake of this behavior.

Our research setting is also distinct in terms of our participants. Researchers have made important strides in this area of research using field experiments conducted with online labor markets (e.g., Mturk) and web-based work (Chandler and Kapelner 2013; Burbano 2016, 2019; Cassar and Meier 2017; Hedblom et al. 2019). However, gig workers often experience the organization in limited stints and as relative outsiders, thus limiting the depth of their knowledge regarding their organizations. Our participants are traditional employees working in their organization, on average for over seven years, with acquired insider knowledge of their organization's practices and priorities. Of course, employees of this type are common, but rarely have they been the focus of field experiments on prosocial claims and employee effort.<sup>2</sup>

**3.2 Participants.** Participants were 1,471 administrative employees at a large university in the northeastern U.S. who purchased office supplies at least once in the six-month baseline period. In accordance with the institutional review board and waiver of informed consent, we did not collect any identifying (or demographic) information about individual employees. The university's procurement office assigned a unique numerical identifier to each employee that was the basis for our randomization and analysis. In the aggregate, the sample draws from a population of administrative employees at this university that is 72% female, on average 40 years of age, employed by the university on average 7.6 years, and, in terms of educational attainment, 86% hold a bachelor's degree or greater.

**3.3 Procedure and Materials.** When employees purchase items through the university's online procurement system, they navigate to an externally-hosted office goods supplier website,

---

<sup>2</sup> We are only aware of one other study that shares this feature (Gosnell et al. 2019) and, unlike research using online labor markets, it found that prosocial incentives had no positive effect on inducing behavioral change. This may suggest that different processes and outcomes emerge in the context of online labor markets versus inside organizations.

where they select items and complete a purchase. We programmed a message to present (see Figure 1) as employees were directed to the supplier site. Employees viewed the same message (or no message, for those in the control condition) every time they clicked on the link to the supplier site during the six-month study period (mean = 40 views; median = 22 views), whether or not they ultimately placed an order. During the study, a very small number of users ( $N = 11$ , or  $<0.01\%$  of participants) complained about the messages.<sup>3</sup> As part of our agreement with the university, we removed them from the study. None of these individuals inquired further about these messages or indicated any knowledge that they were part of a research project.

All three treatment conditions began with the same phrase that presented the university's view of what bundling achieves: "Do you know that bundling items into orders over \$50 means fewer shipments and less wasted fuel?" We manipulated the content of the text that followed, which indicated *why* the university cared about bundling (where "X" designates the name of the university): "X cares about limiting pollution" (prosocial condition), "X cares about limiting costs" (instrumental condition), and "X cares about limiting pollution and limiting costs" (mixed condition; order counterbalanced).

---

<sup>3</sup> Of those participants that requested to be removed, 3 were assigned to the prosocial condition, 3 to the instrumental condition, and 5 to the mixed condition.

*Figure 1: Messages presented to employees  
in prosocial, instrumental, and mixed conditions*<sup>4</sup>

Prosocial Condition	Instrumental Condition	Mixed Condition
<p>Do you know that bundling items into orders over \$50 means fewer shipments and less wasted fuel?</p> <p><b>X cares about limiting pollution.</b></p> <p>Bundle today.</p>	<p>Do you know that bundling items into orders over \$50 means fewer shipments and less wasted fuel?</p> <p><b>X cares about limiting costs.</b></p> <p>Bundle today.</p>	<p>Do you know that bundling items into orders over \$50 means fewer shipments and less wasted fuel?</p> <p><b>X cares about limiting pollution and limiting costs.</b></p> <p>Bundle today.</p>

**3.4 Pilot Study.** Prior to administering the field experiment, we conducted a pilot study using a separate online sample of participants ( $N=166$ ; recruited via Amazon’s Mechanical Turk) to assess the effectiveness of our manipulation. We randomly assigned these participants to view either the prosocial, instrumental, or mixed message (Figure 1). We provided no information about bundling or the context of the organizational change effort in which they would be employed. We did this because the primary goal of the pilot was to evaluate the manipulation; namely, whether participants reliably recognized the organizational motive we sought to highlight. We also aimed to establish that the treatment messages did not systematically differ in other idiosyncratic respects (e.g., understandability, detail) that would threaten the internal validity of the ensuing experiment.

Participants indicated their agreement with two key items that assessed the effectiveness of our manipulation, “According to this message, Organization X cares about saving [the environment/money],” and five additional items that assessed differences on alternative dimensions that we expected to be orthogonal to the organization’s motives. Overall, results of the pilot study (Table 1) support the effectiveness of our manipulation. For example, participants understood the

<sup>4</sup> Participants in the control condition did not view a message. Within the mixed condition, we randomized the order in which the motives were presented (i.e., limiting pollution and limiting costs vs. limiting costs and limiting pollution).

organization to care more about saving the environment in the prosocial versus instrumental condition, but to care more about saving money in the instrumental versus prosocial condition. Further, we observed no differences in how our treatment messages were perceived on secondary dimensions that we did not seek to manipulate. See Supplemental Information (SI) for further details and pairwise comparisons of all combinations.

**Table 1: Pilot Study ANOVA**

	Condition	Mean	Std. Err.	<i>F</i>	<i>p</i>
According to this message, Organization X cares about saving the environment	Prosocial	6.15	0.15	13.89	<0.001
	Instrumental	4.36	0.31		
	Mixed	5.55	0.17		
According to this message, Organization X cares about saving money	Prosocial	4.83	0.29	7.51	0.001
	Instrumental	5.59	0.25		
	Mixed	5.95	0.13		
I understand this message	Prosocial	6.30	0.14	0.05	0.95
	Instrumental	6.23	0.14		
	Mixed	6.27	0.13		
This message is detailed	Prosocial	4.80	0.24	0.76	0.47
	Instrumental	4.73	0.21		
	Mixed	5.04	0.16		
This message is persuasive	Prosocial	5.33	0.23	0.54	0.58
	Instrumental	5.00	0.19		
	Mixed	5.15	0.17		
This message is forceful	Prosocial	4.40	0.27	0.72	0.49
	Instrumental	4.28	0.22		
	Mixed	4.05	0.18		
I like this message	Prosocial	5.40	0.20	0.55	0.58
	Instrumental	5.09	0.20		
	Mixed	5.23	0.16		

Note: P-value from a one-way ANOVA. 1 = *strongly disagree*, 7 = *strongly agree*.

**3.5 Design.** In the course of designing the field experiment, we learned that staff within offices were organized in a hierarchical structure. Interaction between offices was very infrequent; however, employees within the same office sometimes coordinated purchases of supplies, rendering their behavior interdependent, and thus creating the potential for treatment contamination and/or



spillover. Therefore, the appropriate unit of analysis is the office, not the individual. To identify the office to which employees belonged, we used employees' departmental affiliation and the physical location of their desk (i.e., the building and floor).

We randomly assigned each office cluster to an experimental condition such that all employees in a given cluster viewed the same message throughout the experimental period. When assigning the treatments, we included all 556 office clusters (comprised of 1,471 individuals) that made at least one purchase in the six months prior to the experiment. At the end of the six-month study period, we learned of an error that occurred when the university technology liaison programmed our randomization. From the outset, a group of clusters assigned to the control condition was swapped with those assigned to the mixed condition. Although assignment to condition remained random (i.e., the effect was taking two randomly assigned groups and interchanging them), the result was that the sample size for the control condition was smaller than planned (71 offices), and sample size for the mixed condition was larger than planned (214 offices). The sample sizes of the instrumental condition (142 offices) and prosocial condition (129 offices) were unaffected. This error does not reduce our ability to make comparisons *among* treatment conditions, which are focal to our research question, but it does limit our empirical precision in comparing the effects of the treatment conditions to the control. Notwithstanding this limitation, we include the control condition in all analyses. Table 2 presents descriptive statistics at baseline by experimental condition. We find no significant differences between conditions on these measures, suggesting that randomization was successful (see Table S2 in the Supplement for all pairwise comparisons).

**Table 2: Descriptive Statistics at Baseline**  
Mean (Standard Deviation)

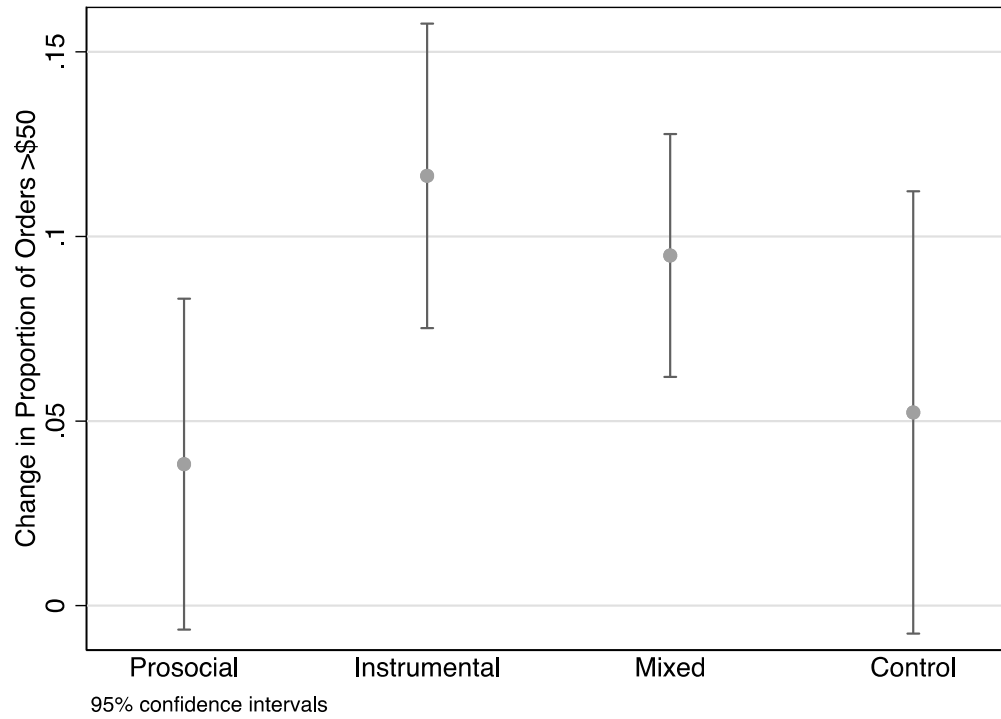
Variable	Prosocial	Instrumental	Mixed	Control	All	<i>P</i>
	(N = 129)	(N = 142)	(N = 214)	(N = 71)	(N = 556)	
Proportion of Orders > \$50	0.69 (0.28)	0.64 (0.29)	0.66 (0.28)	0.68 (0.26)	0.66 (0.28)	0.42
Count of Orders > \$50 (log)	0.86 (0.44)	0.86 (0.46)	0.90 (0.48)	0.93 (0.47)	0.88 (0.46)	0.68
Total Number of Orders (log)	1.02 (0.45)	1.05 (0.44)	1.07 (0.47)	1.10 (0.47)	1.06 (0.45)	0.66
Average Number of Items Per Order	4.00 (3.07)	3.50 (1.73)	3.65 (2.04)	3.94 (3.04)	3.73 (2.39)	0.28
Total Money Spent (log)	2.98 (0.66)	3.00 (0.67)	3.02 (0.70)	3.06 (0.69)	3.01 (0.68)	0.87
Number of Employees in Each Office	2.29 (2.36)	2.63 (3.50)	2.86 (4.05)	2.66 (3.04)	2.65 (3.45)	0.54

Note: P-values from one-way ANOVA. See Supplement for pairwise comparisons between all conditions for all variables at baseline. We log transform Count of Orders > \$50, Total Number of Orders, and Total Money Spent because they were highly positively skewed. We added a value of 1 before taking the log. We present untransformed distributions of these variables in the Supplement.

## 4. Results

**4.1 Primary analyses.** The primary goal of the field experiment was to compare the relative effectiveness of our treatment messages—statements indicating *why* the organization cares about bundling supplies—for getting employees to bundle more frequently. Our analyses measure uptake of bundling, from pre- to post-study periods, in terms of the proportion of orders over \$50: the central behavior that the university sought to change. For each office, we calculated the proportion of orders over \$50 in the in the six-month pre-study baseline period and the six-month study period. In the baseline period, the average proportion of orders over \$50 was 0.66. During the study period, the average was 0.75. This increase was apparent across all conditions (see Figure 2). Our primary analyses centered on whether the change in the proportion of orders over \$50 between the baseline and study period differed by experimental condition.

**Figure 2: Average Change in Proportion of Orders Over \$50 by Condition**



We estimate average treatment effects using difference-in-differences panel regressions in which the control condition is the omitted category. We report OLS models with office fixed effects (Table 3, Model 1). Then, to answer our research question directly—that is, which expressed motive is most effective—we conduct Wald tests (Table 4) that compare treatment effects to one another (instead of to the control). With regard to these focal comparisons between treatment conditions, we observe that the average treatment effect of the instrumental condition was significantly greater than that of the prosocial condition ( $F(1, 487)=6.45, p=0.011$ ). Compared with the prosocial condition, the instrumental condition led to a 7.8 percentage point increase in the proportion of orders over \$50 (95% CI [0.018, 0.138]). The average treatment effect of the mixed condition was also significantly larger than that of the prosocial condition ( $F(1, 487)=4.05, p=0.045$ ). Compared with the prosocial condition, the mixed condition led to a 5.7 percentage point increase in the proportion of orders over \$50 (95% CI [0.001, 0.112]). There was no difference in the effect on proportion of

orders over \$50 between the instrumental and mixed conditions ( $F(1, 487)=0.65, p=0.419, 95\% \text{ CI } [-0.074, 0.031]$ ). The estimates do not substantively change when we remove office fixed effects (Table 3, Model 2). These results thus demonstrate that the instrumental and mixed motive conditions were more effective than the prosocial motive condition at getting employees to bundle.

We also examined the effects of these motives in relation to the control condition (in which no message was presented) to enrich our interpretation of the primary results. The average treatment effect for the instrumental condition was positive, but did not reach conventional levels of statistical significance ( $b=0.064, 95\% \text{ CI } [-0.007, 0.135], p=0.079$ ). Neither the average treatment effect of the prosocial condition ( $b=-0.014, 95\% \text{ CI } [-0.087, 0.059], p=0.709$ ) nor the mixed condition ( $b=0.043, 95\% \text{ CI } [-0.025, 0.109], p=0.214$ ) was significantly different than the control condition. Despite the lack of precision in these estimates, the direction of the coefficients and the relative positioning of the CIs is more consistent with the interpretation that the instrumental condition increased (versus that the prosocial condition decreased) bundling as compared with the control.

**Table 3: Average Treatment Effects, Proportion of Orders >\$50**

	(1) Proportion of Orders >\$50	(2) Proportion of Orders >\$50
Prosocial*Study Period	-0.0140 (0.0374)	-0.0140 (0.0374)
Instrumental*Study Period	0.0641* (0.0364)	0.0641* (0.0364)
Mixed*Study Period	0.0425 (0.0341)	0.0425 (0.0342)
Prosocial		-0.0145 (0.0377)
Instrumental		-0.0483 (0.0363)
Mixed		-0.0449 (0.0337)
Study Period	0.0523* (0.0298)	0.0523* (0.0298)
Constant	0.668*** (0.00528)	0.701*** (0.0280)
Office Fixed Effects	Y	N
Observations	976	976
R-squared	0.125	0.030
Number of Offices	488	488

Note: OLS Models. Unit of analysis is office. Robust standard errors clustered by office in parentheses.  
 \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

**Table 4: Differences in Average Treatment Effect Between Conditions**

<i>Dependent Variable</i>	<i>Comparison</i>	<i>Difference</i>	<i>F</i>	<i>p</i>
Proportion of Orders > \$50	Instrumental vs Prosocial	0.078	6.45	0.011
	Mixed vs Prosocial	0.056	4.05	0.045
	Mixed vs Instrumental	-0.022	0.65	0.419

Note: Wald Tests.  $F(1, 487)$ . Calculated based on model with fixed effects. The results for models without fixed effects are substantively unchanged.

We operationalize bundling behavior as a proportion to isolate the behavior the organization prioritized. However, because we cannot compute a proportion with zero in the denominator, these analyses require us to drop 68 offices (12%) that placed at least one order in the pre-study baseline period (and thus were randomly assigned to conditions), but did not place at least one order in the study period. If our manipulation drove these offices to stop placing orders during the study period,

dropping these offices would lead to biased estimates of the average treatment effects. This interpretation seems unlikely because whereas staff can decide when, and in what combinations to buy supplies, it is not under their discretion to halt procurement altogether (for half a year). Nevertheless, we cannot rule out this possibility. Thus, we examined this issue empirically. Specifically, we constructed an alternative, but conceptually analogous model of bundling that analyzes the count of orders over \$50 conditional on the total number of orders placed by each office location. The count of orders over \$50 variable assigns zeros to those offices that placed no orders in the study period, thereby including all 556 offices in the analysis. We log transform this variable and then estimate difference-in-difference models.

We present these models in Table 5 (Model 3) and present the tests among conditions in Table 6. The results are largely consistent with, and to some extent sharpen, our previous analyses using the proportion of orders over \$50. The treatment effect of the instrumental condition was significantly greater than that of the prosocial condition, ( $F(1,555)=7.38, p=0.007$ ), and nearly surpassed that of the mixed condition ( $F(1,555)=3.78, p=0.052$ ). Unlike in the previous analyses, the treatment effect of the mixed condition did not differ from that of the prosocial condition ( $F(1,555)=1.34, p=0.248$ ). In addition, as compared with the control condition, the instrumental condition had a significant positive effect ( $b=0.074$ , 95% CI [0.008, 0.141],  $p=0.028$ ), indicating that the instrumental message increased bundling as compared with the control. The other treatments did not differ from the control.

**Table 5: Count of Orders >50**

	(3) Count of Orders >\$50 (log)	(4) Count of Orders >\$50 (log)	(5) Count of Orders >\$50 (log)	(6) Count of Orders >\$50 (log)
Prosocial*Study Period	0.00126 (0.0359)	0.00303 (0.0363)	0.0204 (0.0443)	0.0204 (0.0444)
Instrumental*Study Period	0.0747** (0.0338)	0.0750** (0.0339)	0.0773* (0.0404)	0.0773* (0.0404)
Mixed*Study Period	0.0310 (0.0327)	0.0314 (0.0329)	0.0362 (0.0394)	0.0362 (0.0395)
Prosocial		0.00920 (0.0424)		-0.0695 (0.0681)
Instrumental		-0.0118 (0.0424)		-0.0660 (0.0677)
Mixed		-0.0139 (0.0430)		-0.0312 (0.0646)
Study Period	-0.0432 (0.0287)	-0.0474 (0.0297)	-0.0888** (0.0357)	-0.0888** (0.0358)
Total Purchases Made	0.0177*** (0.00184)	0.0161*** (0.00191)		
Constant	0.560*** (0.0327)	0.596*** (0.0489)	0.883*** (0.00553)	0.928*** (0.0557)
Office Fixed Effects	Y	N	Y	N
Observations	1,112	1,112	1,112	1,112
R-squared	0.309	0.572	0.045	0.005
Number of Offices	556	556	556	556

Note: OLS Models. Unit of analysis is office. Robust standard errors clustered by office in parentheses. As a robustness check, we also ran Models 3 and 4 including squared and cubed terms for Total Purchases Made, thereby allowing more flexibility in the relationship between Total Purchases made Count of Orders > \$50. The difference between instrumental and prosocial remains statistically significant at  $p < 0.05$  levels in these models. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

**Table 6: Differences in Average Treatment Effect Between Conditions**

		<i>Difference</i>	<i>F</i>	<i>p</i>
Count of Orders >\$50 (log) (Conditional on number of orders)	Instrumental vs Prosocial	0.073	7.38	0.007
	Mixed vs Prosocial	0.029	1.34	0.248
	Mixed vs Instrumental	-0.044	3.78	0.052
Count of Orders >\$50 (log) (Not conditional on number of orders)	Instrumental vs Prosocial	0.057	3.10	0.079
	Mixed vs Prosocial	0.016	0.26	0.612
	Mixed vs Instrumental	-0.041	2.68	0.102

Note: Wald Tests.  $F(1,555)$ . Calculated based on models with fixed effects. The results for models without fixed effects are substantively unchanged.

We also report the same analyses using the count of orders over \$50 without controlling for the total number of orders (Table 5, Model 5). The pattern of results is largely unchanged; however, our estimates are less precise. Finally, we reach the same conclusions without the inclusion of office fixed effects (Table 5, Models 4 and 6), and when decomposing the six-month baseline and six-month study periods into two-month periods (SI Tables S3A and S3B), indicating that our results are not an artifact of our choice of sampling intervals when creating the panel. Overall, these results demonstrate that the instrumental condition increased bundling as compared with the prosocial and control conditions.

**4.2 Secondary Effects.** In addition to the key behavior that the organization sought to change—the proportion of orders over \$50—we also explored relevant secondary behaviors to provide additional insight into the scope and theoretical underpinnings of our main results.

#### **Total Money Spent**

We examined whether our treatments influenced the total amount of money that offices spent. This measure is particularly relevant because it bears on a potential explanation for our main findings. That is, one possibility is that the instrumental condition increased bundling behavior because it made cost reduction more salient to employees (and, perhaps, the feeling that keeping costs down may be professionally advantageous for them—even though, as described earlier with respect to the organizational context, this was not actually the case). To examine evidence for this possibility, we constructed a measure of total money spent during each period by each office (logged). As displayed in Table 7 (Model 7) and Table 8, none of the treatment conditions significantly differed from one another or in relation to the control condition. While the results are somewhat imprecise, they suggest that, if anything, the offices in the instrumental condition spent more than those in the prosocial condition: the average treatment effect of the instrumental condition was higher than the prosocial condition ( $F(1, 555)=3.38, p=0.067$ ), 95% CI [-0.012,



0.376]). This result is thus not congruent with the alternative explanation for our results regarding the salience of, or a strategic professional interest in, reducing costs. Those employees that received the instrumental message (i.e., “limiting costs”) did not reduce overall spending.

### **Average Number of Items Per Order**

We also examined whether the treatments influenced the average number of items placed in an order (Table 7, Model 8). The number of items per order variable is different, but conceptually related to the proportion of orders over \$50 in that it captures the underlying behavior of including more items in each order that is integral to bundling. We constructed a measure of the average number of items per order for each office. Due to the same constraint on our analyses when using the proportion of orders over \$50, we only include offices that made purchases in both periods. We observe that the instrumental condition significantly increased the number of items per order compared with the prosocial condition ( $F(1,487)=7.56, p=0.006$ ). Moreover, the mixed condition also significantly increased the average number of items compared with the prosocial condition ( $F(1,487)=4.82, p=0.029$ ). The prosocial and mixed conditions did not differ from one another, and none of the conditions was different from control.

### **Total Number of Orders**

Thus far, analyses show that, as compared with the prosocial condition, the instrumental condition led to a greater proportion of orders over \$50 and to more items in each order. Though the treatment messages did not specifically reference the total number of orders, the question remains whether the treatment messages also influenced this outcome during the six months of the study. The results in Table 7 indicate that this was not the case. Using the dependent variable of the number of orders (logged), we observe that the treatment messages did not differ from one another or the control (Wald tests in Table 8). This indicates that offices in the instrumental condition increased bundling by purchasing relatively more items during the study period, not by making fewer

purchases. One possibility is that the instrumental condition prompted employees to stockpile needed office supplies during the study period. These stocks may eventually reduce their total purchases in the months following; however, we cannot observe orders beyond the six-month study period. A second possibility is that the offices in the instrumental condition purchased unneeded office supplies. If this was the case, the core metric chosen by the university (i.e., to increase the proportion of orders over \$50) had an unintended side effect of increasing overall purchasing. Regardless of the longer-term effect on number of orders, these analyses provide additional evidence that the instrumental message was more effective than the prosocial message at changing the target behavior.

**Table 7: Secondary Effects**

	(7) Total Money Spent (log)	(8) Average Number of Items Per Order	(9) Total Number of Orders (log)
Prosocial*Study Period	-0.00752 (0.136)	-0.166 (0.526)	0.0280 (0.0462)
Instrumental*Study Period	0.174 (0.119)	0.371 (0.521)	0.0528 (0.0423)
Mixed*Study Period	0.0518 (0.121)	0.266 (0.521)	0.0246 (0.0413)
Study Period	-0.362*** (0.107)	0.125 (0.505)	-0.140*** (0.0378)
Constant	3.010*** (0.0175)	3.801*** (0.0468)	1.058*** (0.00560)
Office Fixed Effects	Yes	Yes	Yes
Observations	1,112	976	1,112
R-squared	0.123	0.029	0.153
Number of Offices	556	488	556

Note: OLS Models. Unit of analysis is office. Robust standard errors clustered by office in parentheses. Effects not substantively changed without fixed effects. For Total Money Spent and Total Number of Orders, a value of one was added before taking the log.

\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

**Table 8: Secondary Effects, Contrasts Between Conditions**

		<i>Difference</i>	<i>F</i>	<i>p</i>
Total Money Spent (log)	Instrumental vs Prosocial	0.182	3.38	0.067
	Mixed vs Prosocial	0.059	0.34	0.56
	Mixed vs Instrumental	-0.123	2.51	0.11
Average Items Per Order	Instrumental vs Prosocial	0.534	7.59	0.006
	Mixed vs Prosocial	0.431	4.83	0.03
	Mixed vs Instrumental	-0.105	0.34	0.56
Total Number of Orders (log)	Instrumental vs Prosocial	0.025	0.57	0.45
	Mixed vs Prosocial	-0.003	0.01	0.91
	Mixed vs Instrumental	-0.028	1.25	0.26

Note: Wald Tests. For Total Money Spent and Total Number of Orders:  $F(1, 555)$ . For Average Items Per Order:  $F(1, 487)$ .

### **Do Concerns about Environmental Sustainability Matter?**

One salient theoretical prediction, as described earlier, is that employees will respond to prosocial motives to the degree that they are concerned with the prosocial objective at stake. Indeed, studies suggest that an organization's expression of a prosocial motive is most influential for changing behavior among employees who already believe in the importance of the objective and their capacity to meaningfully impact it (Burbano 2019; Fehrler and Kosfeld 2014; Grant and Hofmann 2011; Tonin and Vlassopoulos 2015, but see Cassar 2018). If this mechanism is at play in our context, we would expect the effect of the prosocial treatment to be conditional on environmental concern. In other words, we may expect the prosocial treatment to be more effective for employees particularly concerned with environmental sustainability than those who are not.

To investigate this possibility, we constructed an office-level measure of concern for sustainability. This measure exploits the fact that staff have discretion over whether or not to buy a “sustainable” version of the products they procure. For instance, while staff may have to purchase office paper, they can decide whether to purchase paper that is recycled or not. We used data on these choices as a proxy for employees' concern about the prosocial goal in question. Specifically, we

flagged products as sustainable based on keywords in the product description.<sup>5</sup> We then constructed a variable that reflected the proportion of orders with a sustainable product in the baseline period (18% of all orders). By using data from the baseline period, we are able to estimate employees' concern for environmental sustainability *before* our manipulation was introduced.

Our preliminary analyses demonstrated a significant positive correlation between orders containing sustainable products (our measure of environmental concern) and the proportion of orders over \$50 during the study period ( $r=0.18, p<0.001$ ) as well as the count of orders over \$50 ( $r=0.14, p<0.001$ ). This relationship provides internally-consistent evidence for our conceptualization of environmental concern: all else equal, we would expect those employees concerned with sustainability to be more apt to engage in a discretionary behavior like bundling.

We then shifted to the key empirical question of whether the average treatment effect of the prosocial versus instrumental condition is greater for offices concerned with environmental sustainability than those that are not. The test for such a relationship is a three-way, environmental concern  $\times$  treatment condition  $\times$  study period, interaction. We mean-centered our measure of environmental concern before estimating these models in Table 9. The key three-way interaction between environmental concern, prosocial condition, and the study period was not significant. We observe that the three-way environmental concern  $\times$  mixed condition  $\times$  study period interaction on the proportion of orders over \$50 approaches significance ( $p = 0.06$ ), suggesting that the mixed motive message may have been more effective than the control for those who were less (versus more) concerned with sustainability. We report Wald tests among the interaction terms in the Supplement (Table S4A). These analyses show no significant difference between the instrumental and prosocial conditions. We also run models taking an alternative approach to modeling these

---

<sup>5</sup> We read all product descriptions to generate a list of keywords. The list included: recycled, sustainable, earth, certified, renewable, compostable, Seventh Gen, Seventh Generation, Method, Reman. Note that Seventh Generation, Method and Reman are brands that sell sustainable products.

three-way interactions wherein we “bin” levels of environmental concern (see Supplement). We do not observe any significant three-way interactions in these models either. In sum, these analyses suggest that environmental concern is related to bundling in general, but that employees demonstrating higher (versus lower) levels of environmental concern do not bundle more frequently in response to the prosocial motive.

**Table 9: Treatments Conditional on Environmental Concern**

	(10)	(11)
	Proportion Over \$50	Count of Orders Over \$50 (log)
Prosocial	-0.0198 (0.0391)	0.00307 (0.0433)
Instrumental	-0.0508 (0.0370)	-0.0125 (0.0433)
Mixed	-0.0358 (0.0335)	0.000828 (0.0434)
Study Period	0.0557* (0.0287)	-0.0424 (0.0294)
Prosocial*Study Period	-0.0133 (0.0369)	0.000180 (0.0359)
Instrumental*Study Period	0.0578 (0.0358)	0.0682** (0.0337)
Mixed Condition*Study Period	0.0366 (0.0329)	0.0255 (0.0327)
Sustainable Purchasing	-0.0133 (0.0369)	0.000180 (0.0359)
Prosocial*Sustainable Purchasing	0.313 (0.266)	0.276 (0.246)
Instrumental*Sustainable Purchasing	0.253 (0.276)	0.247 (0.258)
Mixed*Sustainable Purchasing	0.646** (0.263)	0.674*** (0.238)
Sustainable Purchasing*Study Period	0.130 (0.158)	0.185 (0.146)
Prosocial*Sustainable Purchasing*Study Period	-0.257 (0.210)	-0.248 (0.184)
Instrumental*Sustainable Purchasing*Study Period	-0.0191 (0.189)	-0.0700 (0.164)
Mixed*Sustainable Purchasing*Study Period	-0.335* (0.178)	-0.212 (0.171)
Total Purchases Made		0.0159*** (0.00189)
Constant	0.699*** (0.0289)	0.597*** (0.0496)
Observations	976	1,112
R-squared	0.083	0.593

Note: OLS Models. Unit of analysis is office. Robust standard errors clustered by office in parentheses. Sustainable Purchasing is mean centered. As a robustness check, we ran Model 11 including squared and cubed terms for Total Purchases Made; the results are substantively the same.  
\*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

### **Follow-up experiments to clarify theoretical process**

Secondary analyses of Study 1 refine our understanding of the theoretical process insofar as they suggest that two plausible explanations for our results are unlikely. Specifically, these analyses show that the instrumental condition did not uniquely push employees to engage in cost-saving behaviors (as would be expected if they were concerned about being monitored and/or being rewarded). They also show no evidence consistent with the idea that the prosocial (versus instrumental) motive was more effective for employees who showed concern for the prosocial goal at stake.

To further probe theoretical processes, we devised a sequence of vignette experiments with separate online samples (see Spencer, Zanna, and Fong 2005). The goal of the first of these experiments (presented in full in the Supplement) was to explore the plausibility of multiple possible mechanisms (e.g., perceived genuineness, concern for prosocial/instrumental objectives, concern with being monitored). To do so, we presented participants the same treatment messages employed in Study 1, with the inclusion of a matched control condition, and asked them to complete measures relevant to these mechanisms. For each set of measures, we examined evidence consistent with possible mechanisms. As reported in full in Supplement, we only observed significant support for the theorized account regarding perceived genuineness. Given this evidence, in Study 2, we administer a preregistered test focused on the prediction that an organization's expression of an instrumental motive seems more genuine than its expression of a prosocial motive. We then administer a subsequent experiment, Study 3, in which we directly manipulate whether an organization's motives seem genuine to assess the impact on individuals' intention to bundle, thereby testing the full scope of the theorized account in the context of vignette experiments.

## 5. Study 2

**6.1 Participants and procedure.** We preregistered all measures, predictions, and analyses (OSF: <https://osf.io/er365/>). We recruited 786 U.S. participants (447 female;  $M_{\text{age}}=37.14$  years,  $SD=11.27$ ) from Amazon's Mechanical Turk to complete an online study on "organizations" in exchange for payment. In terms of educational attainment, 70.0% held a bachelor's degree or greater, 21.5% completed some college, and 8.6% held a high school degree or only completed some high school. The majority of participants were employed ( $M=87.8\%$ ); of those employed, participants reported working in private (64.8%), non-profit (8.4%), governmental (10.9%), other organizational types (2.6%), or were self-employed (13.3%).

To place participants in the mindset of an employee, we told them that we are exploring the messages that one organization is using to change how its administrative staff orders office supplies (see Supplement for full transcript of materials). We described the difference between bundling (i.e., creating orders greater than \$50) and not bundling (i.e., placing orders less than \$50). Further, we informed participants that this organization wants their administrative staff to bundle and that bundling was voluntary. We then randomly assigned participants to one of four conditions: prosocial, instrumental, mixed (order counterbalanced), and control. Participants in the prosocial, instrumental, and mixed conditions viewed the same message employed in the field experiment. We asked all participants to consider the perspective of an employee at this organization.

**6.2 Process measures.** We asked participants to respond to two complementary sets of genuineness measures, which took different approaches to assessing whether individuals think what the organization stated it cares about is truly what it cares about. In one attribute-based measure, we asked participants to evaluate their impression of the organization on a series of qualities emblematic of genuineness. In a second match-based measure, we examined whether participants think what the organization stated it cares about matches what it actually cares about and seeks to achieve. Finally,



we included secondary measures associated with the alternative theoretical accounts as in the preceding exploratory study. We describe these measures in brief below, and in full in the Supplement.

### **Measures of genuineness.**

**Genuineness index.** To capture impressions of how genuine the organization seemed, we asked participants to complete the following phrase: “Organization X seems \_\_\_\_.” They did so by indicating their agreement (1 = *strongly disagree*, 7 = *strongly agree*) with the following seven attributes: genuine, authentic, honest, transparent, straightforward, sincere, and truthful. As preregistered, we averaged these seven items to form an index of genuineness ( $M=5.19$ ,  $SD=1.24$ ;  $\alpha=.95$ ).

**Believing that stated motives match true motives and goals.** We also assessed how genuine the organization’s motives seemed based on whether participants thought that what the organization stated it cared about matched or mismatched what participants thought it actually cared about and sought to achieve. Following our preregistration, we computed this by comparing participants’ response to the force-choice item—“Organization X tells its employees that it cares about \_\_\_\_.” (limiting costs, limiting pollution, limiting costs and limiting pollution)—with their response to the following two forced-choice items: “I think Organization X asks its employees to bundle because it wants \_\_\_\_.” (primarily to limit costs, primarily to limit pollution, equally to limit costs and limit pollution); “What do you think that bundling helps Organization X do?” (It primarily helps limit costs, It primarily helps limit pollution, It helps equally to limit costs and limit pollution). We constructed two variables based on whether participants’ response to the first item (i.e., the stated motive) matched their response to the second item (i.e., the actual motive), and whether participants’ response to the first item (i.e., the stated motive) matched their response to the third item (i.e., the actual goal). These variables were coded 1 (“genuine”) when the two corresponding items matched, and 0 (“not genuine”) when they did not match. For example, if a participant

indicated that the organization's stated motive is limiting pollution and believed that the organization wants primarily to limit pollution, the variable was coded 1. If a participant indicated that the organization's stated motive is limiting pollution but instead believed that it wants primarily to limit costs (or equally to limit costs and limit pollution), the variable was coded 0. In this way, these match-based variables capture the essence of our theorized process: whether individuals think the organization's stated motives are its true motives.

### **Additional Measures.**

As noted in the preregistration, we included secondary measures to probe plausible alternative processes (i.e., concern with the objectives at stake) or to further contextualize our findings in the literature (i.e., by measuring impressions of kindness and trustworthiness as in past work). Most notably, while past theorizing suggests the expression of a prosocial motive can inspire and motivate, we examined whether it could also be that the instrumental motive does so. However, as with the preceding exploratory study, we find no evidence suggesting that the instrumental motive uniquely amplified participants' interest in the goal, belief in its importance, or the sense that they personally have the capacity to impact it. We include all secondary measures and analyses in the Supplement.

***Table 10: Study 2 Summary Statistics***

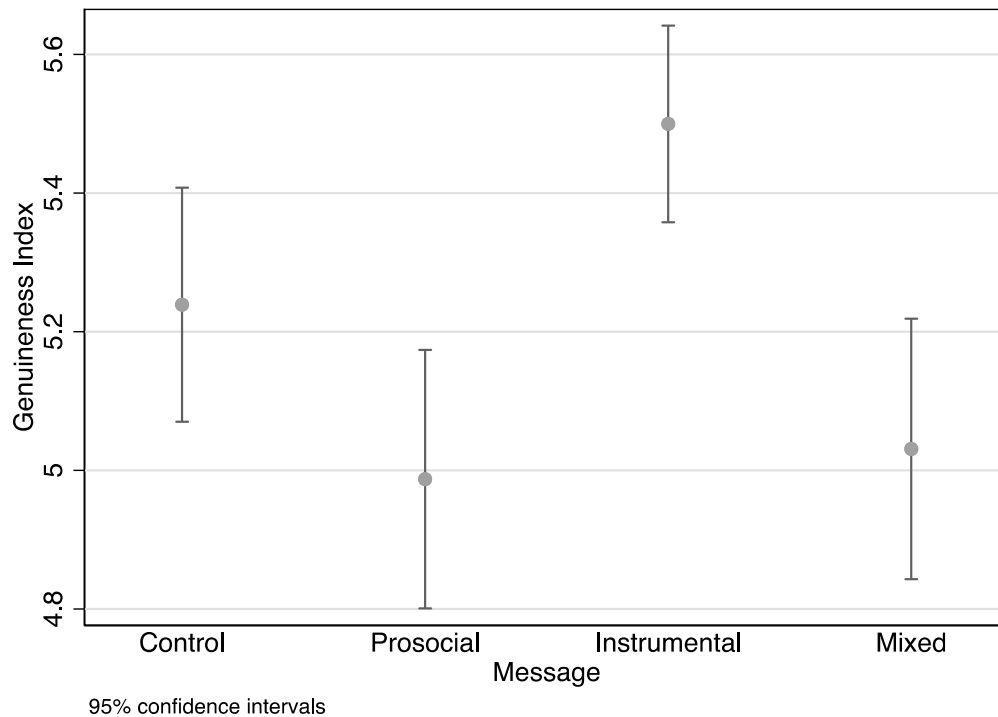
Mean (standard deviation)

<b>Dependent Variable</b>	<b>Prosocial (N = 201)</b>	<b>Instrumental (N = 193)</b>	<b>Mixed (N = 200)</b>	<b>Control (N = 192)</b>	<b>Total</b>
Genuine index	4.99 (1.34)	5.50 (1.00)	5.03 (1.35)	5.24 (1.19)	5.19 (1.24)
Believing stated motive matches true motive (proportion)	0.43 (0.50)	0.69 (0.46)	0.55 (0.50)	0.62 (0.49)	0.57 (0.50)
Believing stated motive matches true goal (proportion)	0.39 (0.49)	0.61 (0.49)	0.56 (0.50)	0.59 (0.49)	0.54 (0.50)

### 6.3 Results.

**Genuineness index.** We submitted participants' scores on the genuineness index to a one-way ANOVA. The ANOVA was significant ( $F(3,785)=7.12, p<0.001$ ). As Figure 3 displays, participants perceived the organization to be more genuine in the instrumental condition ( $M=5.50, SD=1.00$ ) as compared with prosocial condition ( $M=4.99, SD=1.34; t=4.14, p<0.001$ ), supporting our main prediction. They also perceived the organization to be more genuine in the instrumental condition as compared with the control condition ( $M=5.24, SD=1.35; t=2.08, p=0.038$ ) and mixed condition ( $M=5.03, SD=1.19; t=3.78, p<0.001$ ). Finally, we observed that participants in the prosocial condition perceived the organization to be *less* genuine than those in the control condition ( $t=2.03, p=0.043$ ). There were no other significant differences (see Table S12 in the Supplement). These results support our theoretical predictions regarding perceived genuineness.

**Figure 3: Differences in Genuineness Index Across Conditions**

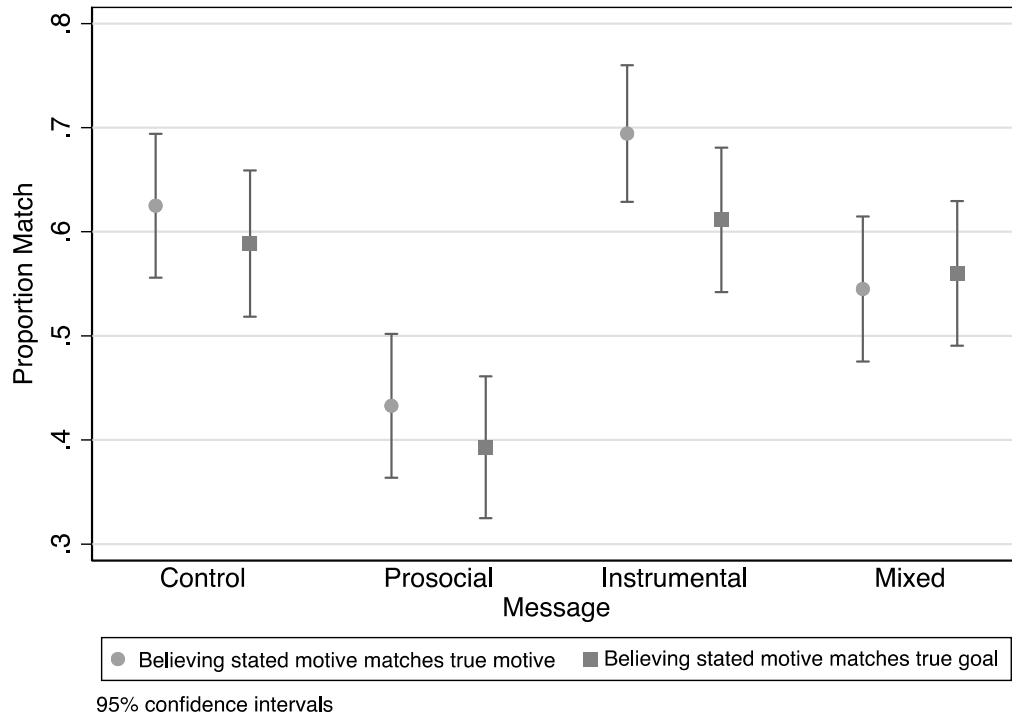


**Believing that stated motives match true motives and goals.** We then turned to our second genuineness measure, reflecting whether participants thought the organization's stated motive matched its true motives and true goals. We first observed that, overall, the two binary variables derived from this operationalization of genuineness (i.e., matching true motives and matching true goals) both were significantly positively correlated with the attributed-based genuineness index described above ( $r=0.37, p<0.01$  and  $r=0.31, p<0.01$ , respectively), consistent with the notion that these measures tap a similar underlying construct.

We then examined the frequency of matches across conditions, with higher values reflecting greater perceived genuineness (Figure 4). Because our outcome measures are binary, we used linear probability models to examine the effects of our treatment conditions on perceived genuineness (Table S10A). In these models, we enter three dummy-coded predictor variables for each of our treatment conditions (with the control condition as the omitted variable). We report tests of all pairwise comparisons in the Supplement (Table S10B). As compared with the prosocial condition, participants in the instrumental condition were 26 percentage points more likely to believe the organization's stated motive matched its true motive ( $F(1,782)=29.30, p<0.001$ ) and were 22 percentage points more likely to believe that the organization's stated motive matched its true goal ( $F(1,782)=19.62, p<0.001$ ). These results support our main prediction.

With respect to the control, we also observe that the prosocial condition led participants to view the organization as significantly *less* genuine both in terms of its true motives ( $t=3.88, p<0.001$ ), and true goals ( $t=3.84, p<0.001$ ). The instrumental and mixed conditions did not significantly differ from the control condition in perceived genuineness in terms of motives.

**Figure 4: Believing Stated Motive Matches True Motive and Goal**



Overall, we observe consistent support for the theorized account regarding perceived genuineness, regardless of how we operationalize the measure. These data from Study 2 cannot definitively explicate the results of Study 1; nevertheless, they provide empirical support for one theorized process by which an instrumental motive may be more effective at promoting behavioral change.

### Study 3

Building on evidence from Study 2, we administer an experiment that examines the subsequent link between perceived genuineness and the intention to bundle, thereby testing the full scope of the theorized mechanism in the context of a laboratory experiment. To do so, we directly manipulated the genuineness (i.e., high versus low) of the same prosocial and instrumental messages used in Studies 1 and 2. We did so by turning our matched-based measure of genuineness in Study 2 into an experimental manipulation. Specifically, participants learned that the prosocial or instrumental message presented was either consistent or inconsistent with the organization's track

record. Based on the theory we develop, we expected individuals to show greater intentions to bundle when the organization's stated motive for bundling seemed more versus less genuine. Further, given evidence that people are more skeptical of prosocial motives, we expected the impact of (high versus low) genuineness to be a particularly influential in the prosocial condition.

**6.1 Participants and procedure.** We recruited 752 U.S. participants (39% female;  $M_{\text{age}}=35.33$  years,  $SD=10.9$ ) from Amazon's Mechanical Turk to complete an online study on "opinions about organizations" in exchange for payment. In terms of educational attainment, 63.2% held a bachelor's degree or greater, 26.9% completed some college, and 9.5% held a high school degree or only completed some high school. The majority of participants were employed ( $M=86.0\%$ ); of those employed, participants reported working in private (68.0%), non-profit (7.9%), governmental (9.7%), other organizational types (3.9%), or were self-employed (10.5%).

We introduced participants to the study and the notion of bundling using the identical procedure as in Study 2 (see Supplement for full transcript of materials). We then randomly assigned participants to one of four conditions. They viewed the same prosocial or instrumental message used in previous studies, followed by a message that manipulated its genuineness (high versus low) in the vein of our match-based genuine measure in Study 2. Participants who viewed the prosocial message subsequently viewed one of the following statements:

**High Genuineness / Prosocial**

"Behind the scenes, numerous employees knowledgeable about Organization X confirm that Organization X genuinely cares about limiting pollution. Organization X's message is consistent with their track record of prioritizing environmentally-friendly operations."

**Low Genuineness / Prosocial**

"Behind the scenes, numerous employees knowledgeable about Organization X doubt whether Organization X only cares about limiting pollution. Organization X's message is not consistent with their track record of prioritizing low-cost operations."

Participants who viewed the instrumental message subsequently viewed one of the following statements:

### **High Genuineness / Instrumental**

“Behind the scenes, numerous employees knowledgeable about Organization X confirm that Organization X genuinely cares about limiting costs. Organization X’s message is consistent with their track record of prioritizing low-cost operations.”

### **Low Genuineness / Instrumental**

“Behind the scenes, numerous employees knowledgeable about Organization X doubt whether Organization X only cares about costs. Organization X’s message is not consistent with their track record of prioritizing environmentally-friendly operations.”

We then asked participants to complete a series of questions that measured the likelihood that they would bundle in situations in which doing so was discretionary or unmonitored, as in the context of our field experiment. Last, participants provided demographic information.

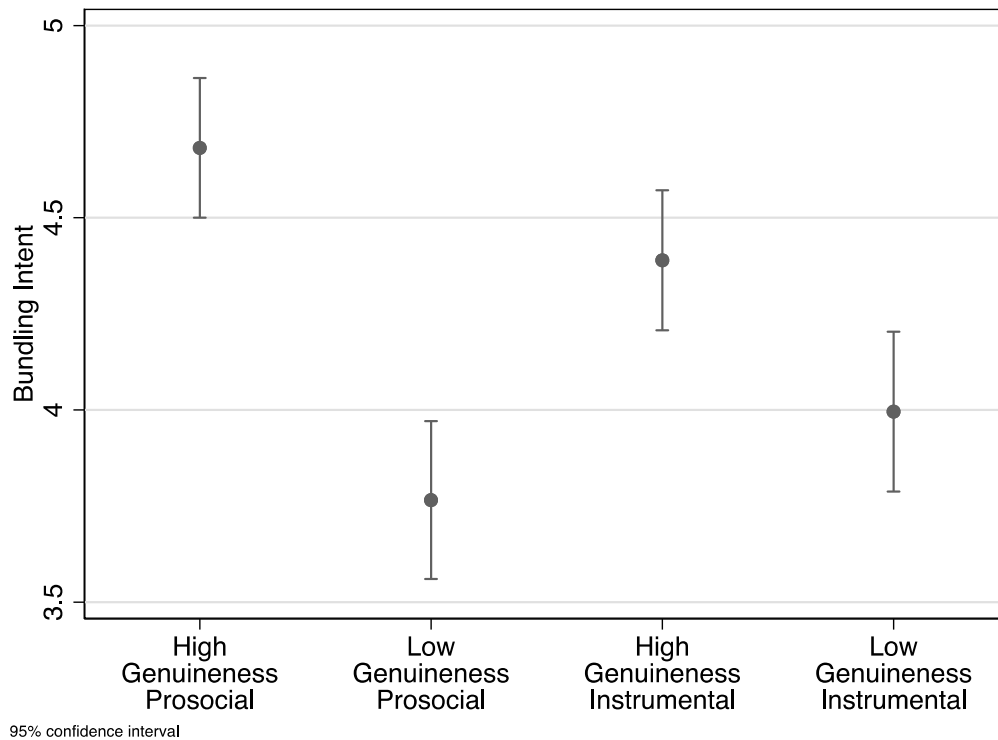
**6.2 Bundling intentions.** To assess participants’ intention to bundle, we asked: “As an employee at Organization X, how likely would you be to bundle if \_\_\_\_.” Participants then completed the question by responding to six statements (1 = *extremely unlikely*, 7 = *extremely likely*). The statements were: “\_\_\_\_ you were not paid to do so?”; “\_\_\_\_ it meant staying late at work without extra pay?”; “\_\_\_\_ you were already busy with your core job responsibilities?”; “\_\_\_\_ your supervisors did not monitor whether or not you bundled?”; “\_\_\_\_ it meant explaining to managers that they might have to wait for the office supplies they want?”; “\_\_\_\_ it was an optional task on top of your actual job requirements?” We averaged responses to these six items to form an index of bundling intention ( $M=4.21$ ,  $SD=1.40$ ;  $\alpha=.86$ ).

**6.3 Results.** We submitted participants’ scores on the bundling intention index to a 2 (message type: prosocial vs. instrumental)  $\times$  2 (genuineness: high vs. low) between-subjects ANOVA. We observed a significant effect of genuineness on bundling intention ( $F(1,749)=41.63$ ,  $p<.001$ ). Specifically, participants were more likely to bundle in response to messages in which

genuineness was high ( $M=4.52$ ,  $SD=1.31$ ) than low ( $M=3.88$ ,  $SD=1.43$ ), supporting our theorized prediction. The effect of message type was not significant ( $F(1,749)=0.04$ ,  $p=0.85$ ).

Further, these main effects were qualified by a significant interaction ( $F(1,749)=6.25$ ,  $p=0.013$ ). As Figure 5 displays, though participants in the instrumental condition showed greater intention to bundle when genuineness was high as compared with low ( $M=4.39$ ,  $SD=1.30$  vs.  $M=4.00$ ,  $SD=1.44$ ;  $t=2.84$ ,  $p<0.01$ ), for participants in the prosocial condition, the impact of high versus low genuineness was particularly pronounced ( $M=4.66$ ,  $SD=1.32$  vs.  $M=3.77$ ,  $SD=1.40$ ;  $t=6.49$ ,  $p<0.001$ ). Notably, the data show that an instrumental message high in genuineness is significantly more effective than an prosocial message low in genuineness ( $t=4.45$ ,  $p<0.001$ )—a practically-informative comparison for organizations choosing to convey prosocial motives for actions that employees believe are intended for instrumental ends.

**Figure 5: Bundling Intention**





Overall, Study 3 provides evidence that perceived genuineness drives individuals' intention to bundle. Together with evidence from Study 2, which indicates that the instrumental motive seems more genuine than the prosocial motive, these follow-up experiments provide empirical support for one theorized process by which an instrumental motive may be more effective than a prosocial motive. Furthermore, they also point to the very conditions under which the reverse can be true: the prosocial condition demonstrated the capacity to be the *most* effective when genuineness was high. This may represent a key theoretical bridge to previous research demonstrating the capacity for prosocial activities to inspire and motivate individuals. In sum, these results suggest that to predict how prosocial versus instrumental motives will affect behavioral change may require an understanding of how genuine they are perceived to be.

## **Conclusions and Future Directions**

Organizations of all types seek to motivate their employees to change routines and adopt new behaviors by communicating the rationale for, and importance of, the behavior. When organizations want their employees to adopt new behaviors that advance both instrumental and prosocial outcomes, organizations face a critical choice regarding how to communicate why they care about them. We investigate the implications of this choice in the context of a natural field experiment that varies the reason a university provides for caring about changing how staff purchase office supplies. We observe robust evidence that expressing an instrumental motive (i.e., limiting costs) is more effective at getting employees to undertake this behavior than expressing a prosocial motive (i.e., limiting pollution). Evidence from subsequent studies that probe reasons why an instrumental motive may be most effective suggests that an organization seems more genuine when it expresses an instrumental (versus prosocial) motive, and in turn, individuals are more inclined to change their behavior on its behalf. These findings and theoretical process sharply deviate from the

preponderance of related research in this domain, which underscores the benefits of promoting prosocial activities for inspiring and motivating employees.

One important implication of our research is that the same contexts in which organizations have much to gain (e.g., financially) from prosocial behaviors may be the very contexts in which they should be most wary about extolling their prosocial motives. This conclusion dovetails with recent work that finds that prosocial incentives are less effective at motivating workers when prosocial actions are clearly self-serving (Cassar and Meier 2017). It may be that only in certain contexts, such as an organization that has consistently acted in ways that align with prosocial values (perhaps, even when costly), that a purely prosocial motive seems genuine, and in turn, proves effective for changing behavior.

Another noteworthy contribution of this research is the (lack of) evidence for “crowding out” or “over-justification” effects—the notion that combining prosocial and instrumental motives leads to negative outcomes, as previous research in psychology and economics has documented. While it may be the case that introducing instrumental incentives to tasks that are personally valuable to individuals can undermine their motivation, learning, and performance (Deci et al. 1999; Frey and Oberholzer-Gee 1997; Lepper et al. 1973) and that holding competing prosocial and instrumental motives is associated with poorer career outcomes (Wrzesniewski et al. 2014), we do not observe evidence of any such maladaptive effects on behavioral change in our context. Our findings for the mixed condition are less clear than for the other two treatment conditions. We observe some evidence that, as compared with the prosocial motive, the mixed motive was more effective for stimulating behavioral change (Study 1); however, the findings are not definitive.

Finally, this research is also noteworthy in terms of its design and measures. Field experiments enhance the credibility of findings by modeling actual behavior at scale, particularly when (as in the present context) they minimally intervene in the natural process through which

individuals engage in a target behavior (List 2009). This research analyzes the behavior of employees who hold conventional positions in a typical organizational setting, and tracks changes in behavior over a twelve-month period. These features, in conjunction with a sequence of two vignette experiments devised to test theorized processes, strengthen the internal and external validity of our findings.

As with all studies, ours has limitations and unanswered questions that suggest fruitful avenues for future research. While the results of the vignette experiments suggest that perceptions of genuineness influence efforts to promote behavioral change, we cannot rule out the possibility that other processes contributed to the results of our field experiment. It will be important for future work examining the impact of prosocial and instrumental motives to directly measure or manipulate perceived genuineness in the field. We note also that the degree to which our results generalize to other forms of prosocial actions (e.g., those focused more concretely on helping people versus the environment), and to other organizational contexts, remains untested. Future scholarship evaluating the generalizability of our findings will be important not only for situating these findings in a larger theoretical framework, but also, prescriptively, as a tool guiding leaders' efforts to enact behavioral change on a range of activities that can be justified on both prosocial and instrumental grounds.

## REFERENCES

- Anik L, Aknin LB, Norton MI, Dunn EW, Quoidbach J (2013) Prosocial bonuses increase employee satisfaction and team performance. *PloS one* 8(9): e75509.
- Apfelbaum EP, Stephens NM, Reagans RE (2016) Beyond one-size-fits-all: Tailoring diversity approaches to the representation of social groups. *Journal of Personality and Social Psychology* 111(114): 547-566.
- Barnett, ML (2019) The business case for corporate social responsibility: A critique and an indirect path forward. *Business & Society* 58(1): 167-190.
- Bartel CA (2001) Social comparisons in boundary-spanning work: Effects of community outreach on members' organizational identity and identification. *Administrative Science Quarterly* 46(3): 379-413.
- Bauman CW, Skitka LJ (2012) Corporate social responsibility as a source of employee satisfaction. *Research in Organizational Behavior*, 32: 63-86.
- Bénabou R, Tirole J (2006) Incentives and prosocial behavior. *American Economic Review*, 96(5): 1652-1678.
- Berman, JZ, Levine EE, Barasch A, Small DA (2015) The Braggart's dilemma: On the social rewards and penalties of advertising prosocial behavior. *Journal of Marketing Research*, 52(1): 90-104.
- Bode, C., Singh, J., & Rogan, M. (2015) Corporate social initiatives and employee retention. *Organization Science*, 26(6): 1702-1720.
- Bolderdijk JW, Steg L, Geller ES, Lehman PK, Postmes, T (2013) Comparing the effectiveness of monetary versus moral motives in environmental campaigning. *Nature Climate Change* 3(4): 413-416.
- Burbano VC (2016) Social responsibility messages and worker wage requirements: Field experimental evidence from online labor marketplaces. *Organization Science* 27(4): 1010-1028.
- Burbano, VC (2019) Getting gig workers to do more by doing good: Field experimental evidence from online platform labor marketplaces. *Organization & Environment*: 1-26.
- Burbano, VC, Mamer J, Snyder J (2018) Pro bono as a human capital learning and screening mechanism: Evidence from law firms. *Strategic Management Journal* 39(11): 2899-2920.
- Brammer S, Millington A, Rayton B (2007) The contribution of corporate social responsibility to organizational commitment. *The International Journal of Human Resource Management* 18(10): 1701-1719.
- Cassar L (2018) Job Mission as a Substitute for Monetary Incentives: Benefits and Limits. *Management Science* 65(2): 896-912.

- Cassar L, Meier S (2018) Nonmonetary Incentives and the Implication of Work as a Source of Meaning. *Journal of Economic Perspectives* 32(3), 215-238.
- Cassar, L, Meier S (2017) Intentions for Doing Good Matter for Doing Well: The (Negative) Signaling Value of Prosocial Incentives. National Bureau of Economic Research. No. w24109.
- Chandler D, Kapelner A (2013) Breaking monotony with meaning: Motivation in crowdsourcing markets. *Journal of Economic Behavior & Organization* 90: 123-133.
- Charness G, Cobo-Reyes R, Sanchez A (2016) The effect of charitable giving on workers' performance: Experimental evidence. *Journal of Economic Behavior & Organization*, 131: 61-74.
- Deci EL, Koestner R, Ryan RM (1999) A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin* 125(6): 627-28.
- Ellemers N, Kingma, L, van den Burgt, J, Barreto, M (2011) Corporate social responsibility as a source of organizational morality, employee commitment and satisfaction. *Journal of Organisational Moral Psychology*, 1: 97-124.
- Evans L, Maio GR, Corner A, Hodgetts CJ, Ahmed S, Hahn U (2013) Self-interest and pro-environmental behaviour. *Nature Climate Change* 3(2): 122.
- Fehrler S, Kosfeld M (2014) Pro-social missions and worker motivation: An experimental study. *Journal of Economic Behavior & Organization* 100: 99-110.
- Feiler DC, Tost, LP, Grant AM (2012) Mixed reasons, missed givings: The costs of blending egoistic and altruistic reasons in donation requests. *Journal of Experimental Social Psychology* 48(6):1322-1328.
- Flammer C, Luo J (2017) Corporate social responsibility as an employee governance tool: Evidence from a quasi-experiment. *Strategic Management Journal* 38(2): 163-183.
- Forehand, MR, Grier, S. (2003) When is honesty the best policy? The effect of stated company intent on consumer skepticism. *Journal of consumer psychology* 13(3): 349-356.
- Frank DH, Smith C (2016) Will Employees Pay to Work for a More Socially Responsible Organization?. *Academy of Management Annual Meeting Proceedings* 2016(1): 868-873.
- Franks DM, Davis R, Bebbington AJ, Ali SH, Kemp D, Scurrah M (2014) Conflict translates environmental and social risk into business costs. *Proc. Natl. Acad. Sci.* 111(21): 7576-7581.
- Frey BS, Oberholzer-Gee F (1997) The cost of price incentives: An empirical analysis of motivation crowding-out. *American Economic Review* 87(4): 746-755.
- Gartenberg C, Prat A, Serafeim G (2019) Corporate purpose and financial performance. *Organization Science* 30(1): 1-18.

- Gneezy U, Meier S, Rey-Biel P (2011) When and why incentives (don't) work to modify behavior. *Journal of Economic Perspectives* 25(4): 191-210.
- Gosnell, GK, List JA, Metcalfe RD (2019) The Impact of Management Practices on Employee Productivity: A Field Experiment with Airline Captains. (No. w25620) *National Bureau of Economic Research*.
- Griskevicius V, Tybur JM, Van den Bergh B (2010) Going green to be seen: status, reputation, and conspicuous conservation. *Journal of personality and social psychology* 98(3): 392.
- Grant AM (2008) Does intrinsic motivation fuel the prosocial fire? Motivational synergy in predicting persistence, performance, and productivity. *Journal of Applied Psychology* 93(1): 48.
- Grant AM, Hofmann DA (2011) Outsourcing inspiration: The performance effects of ideological messages from leaders and beneficiaries. *Organizational Behavior and Human Decision Processes* 116: 173–187.
- Grant AM (2012) Leading with meaning: Beneficiary contact, prosocial impact, and the performance effects of transformational leadership. *Academy of Management Journal* 55(2): 458-476.
- Hahl O (2016) Turning back the clock in baseball: The increased prominence of extrinsic rewards and demand for authenticity. *Organization Science* 27(4): 929-953.
- Hedblom D, Hickman BR, List JA (2019) Toward an understanding of corporate social responsibility: Theory and field experimental evidence (No. w26222). *National Bureau of Economic Research*.
- Henderson R, Van den Steen E (2015) Why do firms have “purpose”? The firm’s role as a carrier of identity and reputation. *American Economic Review*, 105(5): 326-30.
- Imas A (2014) Working for the “warm glow”: On the benefits and limits of prosocial incentives. *Journal of Public Economics* 114: 14-18.
- List JA (2009) An introduction to field experiments in economics. *Journal of Economic Behavior and Organization*, 3(70): 439-442.
- Lepper MR, Greene D, Nisbett RE (1973) Undermining children’s intrinsic interest with extrinsic reward: A test of the “overjustification” hypothesis. *Journal of Personality and Social Psychology* 28(1): 129.
- Makov T, Newman GE (2016) Economic gains stimulate negative evaluations of corporate sustainability initiatives. *Nature Climate Change* 6(9): 844.
- Mayer DM, Ong M, Sonenshein S, Ashford SJ (2019) The money or the morals? When moral language is more effective for selling social issues. *Journal of Applied Psychology* 104(8): 1058-1076.

- Newman, GE, Cain, DM (2014) Tainted altruism: When doing some good is evaluated as worse than doing no good at all. *Psychological science* 25(3): 648-655.
- Osterhus, TL (1997) Pro-social consumer influence strategies: When and how do they work? *Journal of marketing* 61(4): 16-29.
- Radoynovska, N., & King, B. G. (2019). To Whom Are You True? Audience Perceptions of Authenticity in Nascent Crowdfunding Ventures. *Organization Science*, 30(4), 781-802.
- Robinson SL, Rousseau DM (1994) Violating the psychological contract: Not the exception but the norm. *Journal of organizational behavior* 15(3): 245-259.
- Ryan RM, Deci EL (2000) Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American psychologist* 55(1): 68-78.
- Ryan, W. S., & Ryan, R. M. (2019). Toward a social psychology of authenticity: exploring within-person variation in autonomy, congruence, and genuineness using self-determination theory. *Review of General Psychology*, 23(1), 99-112.
- Spencer SJ, Zanna MP, Fong GT (2005) Establishing a causal chain: why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology* 89(6): 845-851.
- Tonin, M, Vlassopoulos M (2014). Corporate philanthropy and productivity: Evidence from an online real effort experiment. *Management Science* 61(8): 1795-1811.
- Van Dyne L, Graham JW, Dienesch RM (1994) Organizational citizenship behavior: Construct redefinition, measurement, and validation. *Academy of management Journal*, 37(4): 765-802.
- Williams P, Fitzsimons GJ, Block LG (2004) When consumers do not recognize “benign” intention questions as persuasion attempts. *Journal of Consumer Research* 21: 540–550.
- Wrzesniewski A, Schwartz B, Cong X, Kane M, Omar A, Kolditz, T (2014) Multiple types of motives don't multiply the motivation of West Point cadets. *Proceedings of the National Academy of Sciences* 111(30): 10990-10995.

## **SUPPORTING INFORMATION**

True Motives: Prosocial and Instrumental Justifications for Behavioral Change in Organizations

Accepted at *Management Science*, May 7, 2020

Matthew Amengual  
University of Oxford

Evan P. Apfelbaum  
Boston University

1. Pilot study
2. Field Experiment Balance, Randomization, and Untransformed Distributions of Logged Dependent Measures
3. Field Experiment Robustness Checks and Additional Analyses
4. Materials for Exploratory Study and Study 2
5. Exploratory Study
6. Study 2 Additional Analyses
7. Pre-registration



## 1. Pilot Study

We ran the pilot study before the field experiment. In the main text, we describe the participants and objective of the study. To recap, the introduction to the study informed participants: “In this study, we are interested in how people view messages that organizations communicate to their employees. Specifically, we are exploring messages that organizations may use to help their administrative staff purchase office supplies.” In the main text, we report the mean responses. In the table below, we report pairwise comparisons between conditions for each variable.

***Table S1: Pilot Study Pairwise Comparisons***

	<b>Comparison</b>	<b>Mean Diff.</b>	<b>Std. Err.</b>	<b><i>p</i></b>
Organization X cares about saving the environment	Prosocial vs Instrumental	1.79	0.35	0.000
	Prosocial vs Mixed	0.60	0.31	0.13
	Instrumental vs. Mixed	-1.19	0.30	0.000
Organization X cares about saving money	Prosocial vs Instrumental	-0.77	0.33	0.05
	Prosocial vs Mixed	-1.13	0.29	0.000
	Instrumental vs. Mixed	-0.36	0.28	0.41
I understand this message	Prosocial vs Instrumental	0.07	0.23	0.95
	Prosocial vs Mixed	0.03	0.20	0.99
	Instrumental vs. Mixed	-0.04	0.20	0.98
This message is detailed	Prosocial vs Instrumental	0.07	0.32	0.97
	Prosocial vs Mixed	-0.24	0.28	0.68
	Instrumental vs. Mixed	-0.31	0.27	0.49
This message is persuasive	Prosocial vs Instrumental	0.33	0.31	0.55
	Prosocial vs Mixed	0.18	0.28	0.80
	Instrumental vs. Mixed	-0.15	0.27	0.85
This message is forceful	Prosocial vs Instrumental	0.13	0.35	0.93
	Prosocial vs Mixed	0.35	0.31	0.49
	Instrumental vs. Mixed	0.22	0.30	0.74
I like this message	Prosocial vs Instrumental	0.31	0.30	0.55
	Prosocial vs Mixed	0.17	0.26	0.80
	Instrumental vs. Mixed	-0.14	0.25	0.84

Note: P-value from pairwise comparisons using Tukey’s HSD

## 2. Field Experiment Balance and Randomization

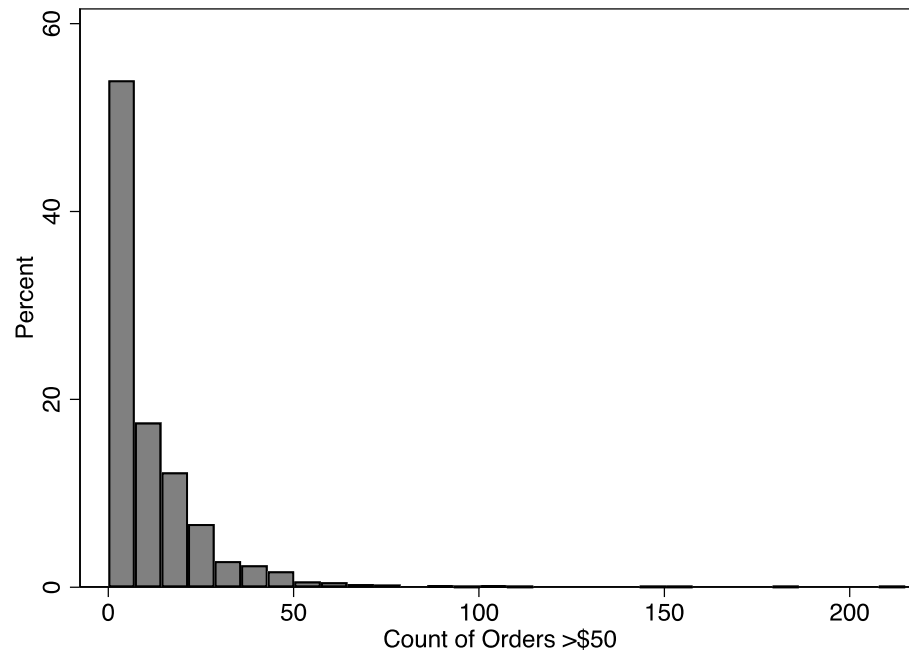
*Table S2: Baseline Balance Checks*

	Comparison	Contrast	Std. Err.	<i>t</i>	<i>p</i>
Proportion of Orders > \$50	Instrumental vs Prosocial	-0.052	0.034	-1.52	0.130
	Mixed vs Prosocial	-0.035	0.031	-1.12	0.262
	Control vs Prosocial	-0.009	0.041	-0.22	0.829
	Mixed vs Instrumental	0.017	0.030	0.55	0.584
	Control vs Instrumental	0.043	0.041	1.05	0.295
	Control vs Mixed	0.026	0.038	0.68	0.497
Count of Orders > \$50 (log)	Instrumental vs Prosocial	0.004	0.057	0.06	0.950
	Mixed vs Prosocial	0.038	0.052	0.74	0.459
	Control vs Prosocial	0.070	0.069	1.01	0.312
	Mixed vs Instrumental	0.035	0.050	0.69	0.489
	Control vs Instrumental	0.066	0.068	0.98	0.329
	Control vs Mixed	0.031	0.064	0.49	0.625
Total Number of Orders (log)	Instrumental vs Prosocial	0.033	0.055	0.59	0.554
	Mixed vs Prosocial	0.052	0.051	1.02	0.310
	Control vs Prosocial	0.076	0.067	1.12	0.261
	Mixed vs Instrumental	0.019	0.049	0.38	0.703
	Control vs Instrumental	0.043	0.066	0.65	0.518
	Control vs Mixed	0.024	0.062	0.39	0.700
Average Number of Items Per Order	Instrumental vs Prosocial	-0.504	0.291	-1.73	0.084
	Mixed vs Prosocial	-0.353	0.267	-1.32	0.186
	Control vs Prosocial	-0.057	0.354	-0.16	0.871
	Mixed vs Instrumental	0.151	0.259	0.58	0.559
	Control vs Instrumental	0.447	0.348	1.29	0.199
	Control vs Mixed	0.296	0.328	0.90	0.367
Total Money Spent (log)	Instrumental vs Prosocial	0.014	0.083	0.17	0.863
	Mixed vs Prosocial	0.038	0.076	0.50	0.616
	Control vs Prosocial	0.080	0.101	0.79	0.430
	Mixed vs Instrumental	0.024	0.074	0.32	0.748
	Control vs Instrumental	0.065	0.099	0.66	0.511
	Control vs Mixed	0.041	0.093	0.44	0.657
Number of Employees in Each Office	Instrumental vs Prosocial	0.339	0.420	0.81	0.419
	Mixed vs Prosocial	0.565	0.385	1.47	0.142
	Control vs Prosocial	0.367	0.510	0.72	0.472
	Mixed vs Instrumental	0.226	0.374	0.60	0.545
	Control vs Instrumental	0.028	0.502	0.06	0.955
	Control vs Mixed	-0.198	0.473	-0.42	0.676

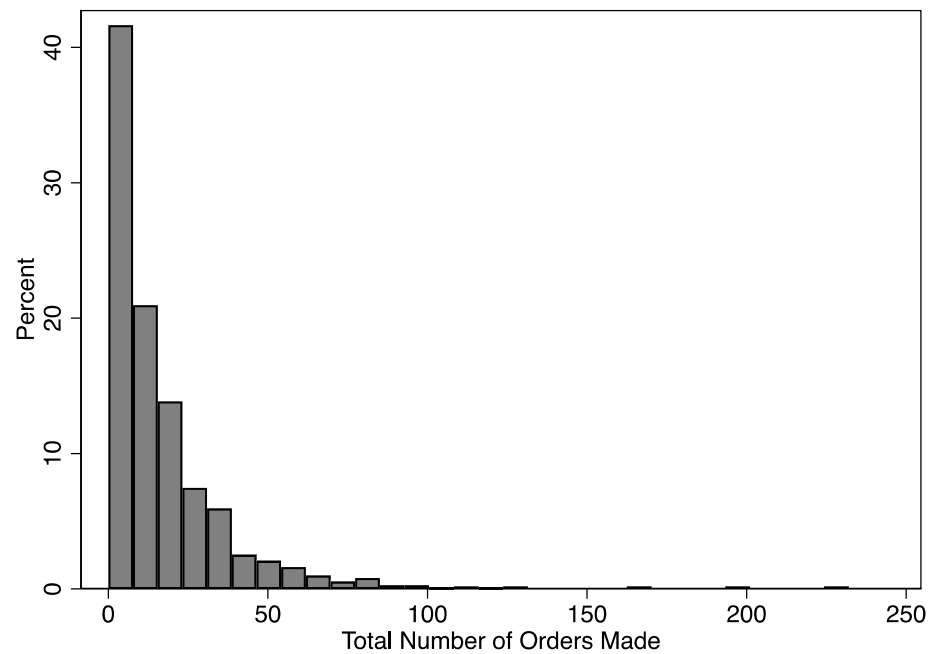
Note: Means appear in Table 2.

Untransformed distributions of logged dependent measures.

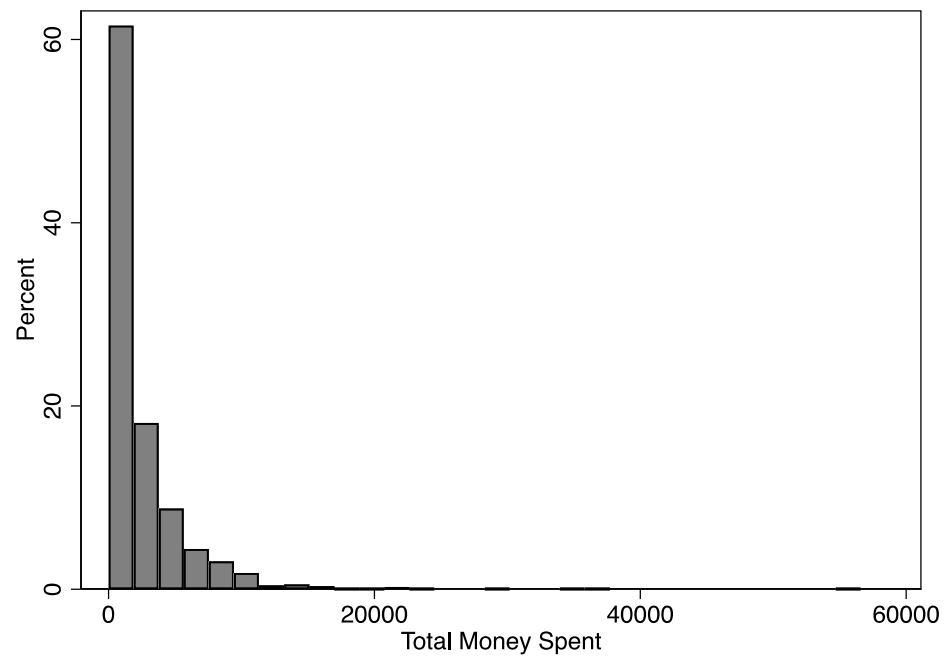
**Figure S1A: Untransformed Distribution of Count of Orders > \$50**



**Figure S1B: Untransformed Distribution of Total Number of Orders Made**



**Figure S1C: Untransformed Distribution of Total Money Spent**



### 3. Field Experiment Robustness Checks and Additional Analyses

In our main analyses, we aggregate our data into six-month time periods (the full 6-month baseline period and the full 6-month study period). In the analyses below, we aggregate the data at two-month intervals. We include fixed effects for each time period to absorb any period-specific changes across all conditions. In Model 1 we do not include office fixed effects. In Model 2 we do. In both models, the instrumental treatment has a positive and statistically significant effect compared with the control. Wald tests (Table S3B) show that our main finding—that instrumental outperforms prosocial—is robust.

**Table S3A Two-Month Panels**

	(1) Count of Orders > \$50 (log)	(2) Count of Orders > \$50 (log)
Prosocial*Study Period	0.0175 (0.0487)	0.0149 (0.0484)
Instrumental*Study Period	0.103** (0.0480)	0.102** (0.0481)
Mixed Condition*Study Period	0.0568 (0.0449)	0.0562 (0.0448)
Prosocial	0.0216 (0.0680)	
Instrumental	-0.0176 (0.0670)	
Mixed	-0.0182 (0.0692)	
Study Period	-0.0318 (0.0433)	0.0980*** (0.00655)
Number of Purchases in Month	0.0907*** (0.00947)	0.592*** (0.0429)
Constant	0.645*** (0.0807)	0.0980*** (0.00655)
Two-Month Fixed Effects	Yes	Yes
Office Fixed Effects	No	Yes
Observations	3,336	3,336
R-squared	0.625	0.389
Number of Offices	556	556

Note: OLS Models. Robust standard errors clustered by office in parentheses.

**Table S3B: Differences in Average Treatment Effect Across Conditions  
(Two-Month Models)**

		<i>Difference</i>	<i>F</i>	<i>p</i>
Count of Orders > \$50 (log) No Office Fixed Effects	Instrumental vs Prosocial	0.085	5.06	0.02
	Mixed vs Prosocial	0.039	1.36	0.24
	Mixed vs Instrumental	-0.046	1.96	0.16
Count of Orders > \$50 (log) Office Fixed Effects	Instrumental vs Prosocial	0.087	5.40	0.02
	Mixed vs Prosocial	0.041	1.53	0.21
	Mixed vs Instrumental	-0.046	1.96	0.16

Note: Wald Tests.  $F(1,555)$

### Conditional Effects

We test for a three-way, environmental concern  $\times$  treatment condition  $\times$  study period, interaction to probe whether environmental concern conditioned responses to the prosocial motive. In Table S4, we report Wald tests among the three-way interactions. We observe a significant difference between mixed and instrumental conditions on the proportion of orders > \$50, suggesting that the instrumental condition is more effective than the mixed condition at high (versus low) levels of environmental concern.

To complement the analysis of the three-way interactions, we undertake two additional analyses. First, recent research suggests caution in using linear interactions and recommends ensuring robustness by binning.<sup>6</sup> Following this advice, we create a dummy variable with high / low levels of environmental concern (Table S5A). Fitting models with control condition as the omitted condition, we find no evidence of significant three-way interactions. We also run similar models with prosocial as the omitted condition, we do not find any three-way interactions in these models either (full models available upon request).

Second, we estimate separate regressions splitting at the median (Table S5B). Once again, we test our central hypotheses with Wald tests (Table S5C). For the offices below the median (i.e., low concern for environmental sustainability), the patterns are the same as in the analyses of the full sample, but they are imprecise and are not statistically significant. For those offices above the median (i.e., high concern for environmental sustainability), we observe statistically significant differences in the treatment effects; however, these results are not consistent with the prediction that concern for environmental sustainability conditions a positive response to the prosocial motive. Rather, they mirror our findings from the main analyses. Specifically, even for those employees particularly concerned about environmental sustainability, the instrumental message was significantly more effective than the prosocial message, both in terms of the proportion of orders over \$50 ( $F(1,241)=6.62, p=0.01$ ) and the count of orders over \$50 ( $F(1,277)=6.32, p=0.01$ ).

<sup>6</sup> Hainmueller, Jens, Jonathan Mummolo, and Yiqing Xu. "How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice." *Political Analysis* 27.2 (2019): 163-192.

***Table S4: Wald Tests Among Interaction Terms***

		<i>F</i>	<i>p</i>
Proportion of Orders > \$50	Instrumental vs Prosocial	1.89	0.17
	Mixed vs Prosocial	0.24	0.63
	Mixed vs Instrumental	5.71	0.02
Count of Orders > \$50 (log)	Instrumental vs Prosocial	1.76	0.19
	Mixed vs Prosocial	0.06	0.80
	Mixed vs Instrumental	1.51	0.23

Note: Wald Tests based on models 10 and 11 in Table 9.  $F(1,487)$  for Proportion Over \$50.  
 $F(1,555)$  for Count of Orders > \$50 (log)

**Table S5A: Interactions with Binned Measure of Environmental Concern**

	(5) Proportion Over \$50	(6) Count of Orders Over \$50 (log)
Prosocial	-0.0987 (0.0616)	-0.134** (0.0603)
Instrumental	-0.160*** (0.0545)	-0.110* (0.0593)
Mixed	-0.187*** (0.0492)	-0.127** (0.0561)
Study	0.0278 (0.0449)	-0.0782* (0.0439)
Sustainable Purchasing (dummy)	-0.121** (0.0543)	-0.0368 (0.0701)
Prosocial*Study Period	0.00298 (0.0606)	0.0363 (0.0539)
Instrumental*Study Period	0.0659 (0.0564)	0.0941* (0.0512)
Mixed Condition*Study Period	0.0821 (0.0519)	0.0642 (0.0483)
Sustainable Purchasing (dummy)*Study Period	0.0461 (0.0600)	0.0569 (0.0591)
Prosocial*Sustainable Purchasing (dummy)	0.154** (0.0770)	0.252*** (0.0835)
Instrumental*Sustainable Purchasing (dummy)	0.211*** (0.0714)	0.195** (0.0835)
Mixed*Sustainable Purchasing (dummy)	0.284*** (0.0646)	0.242*** (0.0831)
Prosocial*Sustainable Purchasing (dummy)*Study Period	-0.0334 (0.0771)	-0.0625 (0.0729)
Instrumental*Sustainable Purchasing (dummy)*Study Period	-0.00295 (0.0734)	-0.0353 (0.0679)
Mixed*Sustainable Purchasing (dummy)*Study Period	-0.0776 (0.0685)	-0.0634 (0.0667)
Total Purchases Made		0.0156*** (0.00186)
Constant	0.766*** (0.0398)	0.624*** (0.0567)
Observations	976	1,112
R-squared	0.080	0.606

Note: OLS Models. Unit of analysis is office. Robust standard errors clustered by office in parentheses. Sustainable Purchasing Dummy = 1 if above median, 0 if below. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$



**Table S5B: Split Samples**

	(7) Proportion of Orders > \$50 (Low Environmental Concern)	(8) Proportion of Orders > \$50 (High Environmental Concern)	(9) Count of Orders > \$50 (log) (Low Environmental Concern)	(10) Count of Orders > \$50 (log) (High Environmental Concern)
Prosocial*Study Period	0.00713 (0.0563)	-0.0376 (0.0508)	0.0366 (0.0539)	-0.0361 (0.0473)
Instrumental*Study Period	0.0672 (0.0525)	0.0590 (0.0502)	0.0947* (0.0515)	0.0533 (0.0437)
Mixed*Study Period	0.0858* (0.0482)	-0.00647 (0.0478)	0.0648 (0.0487)	-0.00588 (0.0442)
Study Period	0.0277 (0.0411)	0.0786* (0.0429)	-0.0777* (0.0433)	-0.00573 (0.0369)
Total Purchases Made			0.0161*** (0.00242)	0.0195*** (0.00265)
Constant	0.625*** (0.00814)	0.712*** (0.00670)	0.508*** (0.0372)	0.597*** (0.0534)
Office Fixed Effects	Yes	Yes	Yes	Yes
Observations	492	484	558	554
R-squared	0.111	0.160	0.271	0.356
Number of Office Clusters	246	242	279	277

Note: OLS Models. Unit of analysis is office. “Low Environmental Concern” refers to those offices less than the median proportion of orders with sustainable products; “High Environmental Concern” refers to those offices greater than or equal to the median. Robust standard errors clustered by office in parentheses. All models include office-level fixed effects. The results do not substantively change without the inclusion of fixed effects. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

**Table S5C: Wald Tests for Split Samples**

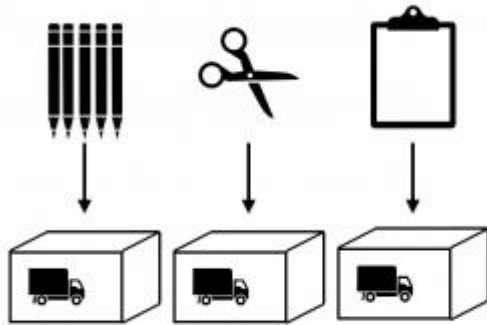
		<b>Difference</b>	<b>F</b>	<b>p</b>
Proportion of Orders > \$50 (Low Environmental Concern) <i>F</i> (1, 245)	Instrumental vs Prosocial	0.060	1.42	0.23
	Mixed vs Prosocial	0.079	2.95	0.09
	Mixed vs Instrumental	-0.018	0.20	0.65
Proportion of Orders > \$50 (High Environmental Concern) <i>F</i> (1, 241)	Instrumental vs Prosocial	0.097	6.62	0.01
	Mixed vs Prosocial	0.031	0.83	0.36
	Mixed vs Instrumental	-0.065	3.87	0.05
Count of Orders > \$50 (log) (Low Environmental Concern) <i>F</i> (1, 278)	Instrumental vs Prosocial	0.058	2.02	0.16
	Mixed vs Prosocial	0.028	0.58	0.45
	Mixed vs Instrumental	-0.029	0.82	0.36
Count of Orders > \$50 (log) (High Environmental Concern) <i>F</i> (1, 276)	Instrumental vs Prosocial	0.089	6.32	0.01
	Mixed vs Prosocial	0.030	0.72	0.39
	Mixed vs Instrumental	-0.059	3.61	0.06

#### 4. Materials for Exploratory Study and Study 2

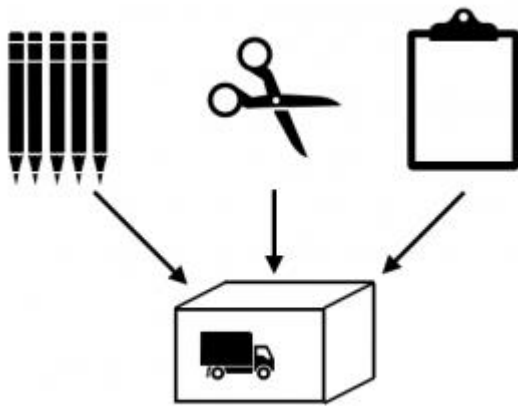
Below, we present the full transcript of the materials employed in the Exploratory Study and Study 2, in the order that participants viewed them. First, all participants saw:

Thank you for participating in today's study. We are interested in the way organizations communicate to their employees. Specifically, in this study, we are exploring the messages that one organization is using to change how its administrative staff orders office supplies.

Staff can place several small orders that are each less than \$50,



or staff can combine items to create orders greater than \$50.



**This organization wants its administrative staff to combine items, or "bundle," to create orders greater than \$50.**

---Page Break---

We then administered our experimental manipulation.

Participants in the moral, instrumental, and mixed conditions saw:

On the following page, you will see the message this organization plans to send its employees each time they start an order. The message communicates WHAT bundling achieves and WHY the organization says it cares about bundling.

Did you know that bundling items into orders over \$50 means fewer shipments and less wasted fuel?

**Organization X cares about [limiting pollution / limiting costs / limiting pollution and limiting costs]. Bundle today.**

**Participants in the control condition saw:**

On the following page, you will see the message this organization plans to send its employees each time they start an order. The message communicates WHAT bundling achieves.

Did you know that bundling items into orders over \$50 means fewer shipments and less wasted fuel?

**Bundle today.**

**---Page Break---**

**Participants then completed the dependent measures reported in the main text / supplement.**

**END.**

## 5. Exploratory Study

Prior to administering Study 2, which we preregistered, we conducted an exploratory version of a similar study that aimed to examine plausible processes that may have been at play in Study 1. The exploratory study presented the identical cover story and materials as in Study 2; it differed only with respect to some of the dependent measures collected. Specifically, it did not include all seven items that comprise the genuineness index and it did not include the Likert-based measures of perceived concern for the organization's objectives (i.e., in terms of interest, importance, and impact). The exploratory study also included some questions that were not part of Study 2. Specifically, the exploratory study included match-based measures of perceived concern for the organization's objectives (i.e., in terms of interest, importance, and impact), which we describe below, as well as a set of measures assessing the perceived professional benefits of bundling. We report the full method and results from this exploratory study below.

**Participants and procedure.** We recruited 757 U.S. participants (347 female;  $M_{\text{age}}=35.33$  years,  $SD=11.21$ ) from Amazon's Mechanical Turk to complete an online study on "organizations" in exchange for payment. In terms of educational attainment, 62.2% held a bachelor's degree or greater, 25.9% completed some college, and 11.9% held a high school degree or only completed some high school. The majority of participants were employed ( $M=86.3\%$ ); of those employed, participants reported working in private (64.5%), non-profit (10.8%), governmental (8.3%), other organizational types (3.1%), or were self-employed (13.2%).

**Process measures.** We asked participants to respond to separate blocks of measures that probed mechanisms through which instrumental and prosocial motives may influence the adoption of bundling. Specifically, we examined perceived genuineness, perceived concern for the (prosocial/instrumental) motives expressed, and participants' subjective impressions of the organization (e.g., how kind it seemed). As a secondary measure, we assessed participants' perception that engaging in bundling would be professionally advantageous for them.

**Perceived genuineness of motives.** We assessed perceived genuineness using the same match-based approach as in Study 2. We also coded perceived genuineness using a second match-based approach. In this approach, we coded whether the treatment condition to which participants were randomly assigned (versus participants' *perception* of the organization's stated motive) matched their belief about the organization's true motive and true goal. This alternative match-based approach offered another theoretically congruent way to operationalize perceived genuineness. However, coding genuineness in terms of whether the treatment message to which participants were assigned matched their reported beliefs about the true motive and goal precluded us from incorporating the control condition (which, deliberately, did not specify a motive).

**Perceived concern for objectives.** While past theorizing suggests the expression of a prosocial motive can inspire and motivate, it could also be that the instrumental motive does so. For instance, the instrumental motive may uniquely amplify participants' interest in helping the organization reduce costs, their belief in the importance of doing so, or the sense that they personally have the capacity to impact such objectives. This notion that the instrumental motive triggers greater concern with the instrumental objectives thus represents a different potential underlying process.

To probe this possibility, we enlisted an analogous match-based approach to construct variables that assessed whether our treatments enhanced perceived concern for the prosocial and instrumental objectives in question. Participants completed the following three forced-choice items regarding their concern for the objectives: “As a staff member, I would feel most interested in helping Organization X \_\_\_\_” (reduce costs, reduce pollution, reduce costs and pollution); “As a staff member, I would feel that it is most important to help Organization X \_\_\_\_” (reduce costs, reduce pollution, reduce costs and reduce pollution); “As a staff member, I would feel that my actions could have a direct impact on \_\_\_\_” (reducing costs, reducing pollution, reducing costs and reducing pollution). As above, we first constructed three variables based on whether participants’ perception of what the organization tells its employees matched their responses to these three items. In addition, we also examined whether the treatment message to which participants were randomly assigned matched their responses to these three items. For both types of match-based measures, when responses matched, the variable was coded 1 (“genuine”); when they did not match, the variable was coded 0 (“not genuine”).

**Impressions of the organization.** We also explored participants’ subjective impressions of the organization in terms of how genuine, kind, and trustworthy it seemed. To assess the effects of our treatments on these impressions of the organization, we asked: “To what extent does Organization X seem \_\_\_\_.” Participants then indicated their agreement (1 = *strongly disagree*, 7 = *strongly agree*) with the following five attributes: honest, transparent, straightforward, kind, and trustworthy.

## Results.

The primary goal of our analyses was to examine whether our treatment messages influenced participants’ responses to the process measures described above. In particular, we sought to examine whether any observed effects of the treatment messages on these process measures could speak to a plausible account as to why an instrumental motive would be more effective at changing employees’ behavior than the prosocial motive. We report conditional means and standard deviations for all measures in Table S6.

**Perceived genuineness of motives.** We begin by assessing how genuine the organization’s motives seemed based on whether participants’ thought that what the organization stated it cared about matched or mismatched what participants thought it actually cared about and sought to achieve. In other words, as in Study 2, we construct match-based variables based on whether participants’ perceptions of the stated motive matched their beliefs about the true motive and goals. In general, participants perceived the instrumental motive to be highly genuine, both when defined in terms of its true motive ( $M = 0.75$ ) and true goal ( $M = 0.65$ ). Because our outcome measures are binary, we used linear probability models to examine the effects of our treatment conditions on perceived genuineness (Table S7A). In these models, we enter three dummy-coded predictor variables for each of our treatment conditions. We report tests between conditions in Table S7B. Compared with the prosocial condition, we find that participants in the instrumental condition are 31 percentage points more likely to believe the organization’s expressed motive matched its actual motive ( $F(1,753) = 43.1, p < 0.001$ ) and 22 percentage points more likely to believe the organization’s expressed motive matched its true goal ( $F(1,753) = 18.9, p < 0.001$ ). With respect to the control, we observe that the instrumental condition significantly increased perceived genuineness in terms of the organization’s motive ( $t = 3.08, p < 0.01$ ), but not in terms of its true goal ( $t = 1.23, p = 0.22$ ). We further observe that, in relation to the control, the prosocial condition significantly decreased perceived

genuineness both in terms of what the organization's motive ( $t=3.22, p<0.01$ ) and goal ( $t=3.00, p=0.003$ ). These findings point to a divergent effect of our treatment messages on perceived genuineness: perceived genuineness increased in response to the instrumental message, but also *decreased* in response to the prosocial message.

We observe a similar pattern of results between the prosocial and instrumental messages when we analyze the match variables based upon the condition to which participants were randomly assigned (rather than the participants perception of what the organization is telling its employees). Overall, these results support the theorized account regarding genuineness.

**Perceived concern for objectives.** Next, we probe the plausibility of the perceived concern mechanism. Based on this mechanism, we would expect the instrumental motive to trigger greater levels of concern for the (instrumental) objective described, thereby motivating individuals to adopt behaviors in service of this objective. We report means, standard deviations for these measures in Table S6 and report linear probability models to test for differences across conditions in Table S8. In short, while there are some significant effects, we do not observe patterns of evidence compatible with this theory, regardless of whether concern is operationalized as perceived interest, importance, or potential to personally impact the objective in question. These data do not support the idea that expressing an instrumental motive uniquely amplifies perceived concern for this objective.

**Impressions of the organization.** Turning to impressions of the organization, we observe additional evidence consistent with the perceived genuineness theory (see Table S6 for means and Table S9 for all pairwise comparisons). We use one-way ANOVA to probe the differences across treatments. These analyses reveal that participants in the instrumental condition perceived the organization to be more straightforward ( $t=4.82, p<0.01$ ), honest ( $t=4.03, p<0.01$ ), and transparent ( $t=4.14, p<0.01$ ) than participants in the prosocial condition. We also find that, compared with the control condition, participants in the instrumental condition perceived the organization to be more straightforward ( $t=4.15, p<0.01$ ), honest ( $t=2.99, p<0.01$ ), and transparent ( $t=3.34, p<0.01$ ). The prosocial condition was not different from the control for any of these measures.

We find a different pattern for perceptions of the organization as kind. Participants perceived the organization as more kind in the prosocial condition than both the instrumental ( $t=3.60, p<0.01$ ) and control conditions ( $t=2.56, p=0.01$ ). Finally, for trustworthy, we did not observe a statistically significant difference between instrumental and prosocial conditions ( $t=1.09, p=0.28$ ), or between these conditions and the control.

**Professional benefits of bundling.** To assess the degree to which the participants expected the organization to monitor and reward their bundling, we asked participants: "As an employee at Organization X, I think that \_\_\_\_." Participants then indicated their agreement (1 = *strongly disagree*, 7 = *strongly agree*) with four statements: "Organization X would monitor how frequently I bundle"; "how frequently I bundle would influence whether I am promoted"; "how often I bundle would influence how much I am paid"; "how often I bundle would influence my performance reviews." We averaged responses to these four items to form an index (Career Index) of perceived benefits of bundling ( $\alpha = .85$ ). The means across conditions were: Prosocial ( $M=4.63, SD=1.40$ ); Instrumental ( $M=4.61, SD=1.26$ ); Mixed ( $M=4.60, SD=1.34$ ); Control ( $M=4.62, SD=1.30$ ).

We submitted participants' scores on the perceived benefit of bundling index to a one-way ANOVA. The effect of treatment condition on the perception that bundling would be

professionally beneficial was not significant ( $F(3, 753) = 0.03, p = .99$ ), nor was the pairwise relationship between the instrumental and prosocial condition (see Table S9). These results are not consistent with the notion that the instrumental condition triggers a perception that bundling is professionally advantageous.

**Table S6: Summary Statistics (Exploratory Study)**

Mean (standard deviation)

Dependent Variable	Prosocial (N = 191)	Instrumental (N = 197)	Mixed (N = 187)	Control (N = 182)	Total
Perceived Genuineness (based on perception/DV response match)					
Believing stated motive matches true motive (proportion)	0.44 (0.50)	0.75 (0.43)	0.55 (0.50)	0.60 (0.49)	0.59 (0.49)
Believe stated goal is true goal (proportion)	0.44 (0.50)	0.65 (0.48)	0.64 (0.48)	0.59 (0.49)	0.58 (0.49)
Perceived Genuineness (based on condition/DV response match)					
Believing stated motive matches true motive (proportion)	0.33 (0.47)	0.76 (0.43)	0.48 (0.50)		0.53 (0.50)
Believe stated goal is true goal (proportion)	0.30 (0.46)	0.63 (0.48)	0.59 (0.49)		0.51 (0.50)
Concern for Objectives (based on perception/DV response match)					
Most interested in helping...	0.60 (0.49)	0.49 (0.50)	0.67 (0.47)	0.60 (0.49)	0.59 (0.49)
Most important to help...	0.51 (0.50)	0.55 (0.50)	0.67 (0.47)	0.55 (0.50)	0.57 (0.50)
Could have a direct impact on...	0.49 (0.50)	0.54 (0.50)	0.68 (0.47)	0.67 (0.47)	0.59 (0.49)
Concern for Objectives (based on condition/DV response match)					
Most interested in helping...	0.43 (0.50)	0.31 (0.47)	0.61 (0.49)		0.45 (0.50)
Most important to help...	0.39 (0.49)	0.39 (0.49)	0.61 (0.49)		0.46 (0.50)
Could have a direct impact on...	0.30 (0.46)	0.37 (0.48)	0.65 (0.48)		0.44 (0.50)
Impressions of the Organization					
Straightforward	5.18 (1.48)	5.81 (1.00)	5.75 (1.23)	5.26 (1.40)	5.51 (1.31)
Honest	5.11 (1.40)	5.62 (1.04)	5.58 (1.32)	5.24 (1.20)	5.39 (1.26)
Transparent	4.98 (1.49)	5.54 (1.06)	5.47 (1.34)	5.08 (1.39)	5.27 (1.35)
Kind	5.34 (1.20)	4.87 (1.21)	5.36 (1.38)	5.00 (1.26)	5.14 (1.28)
Trustworthy	5.13 (1.40)	5.26 (1.11)	5.58 (1.26)	5.17 (1.23)	5.28 (1.26)

**Table S7A: Perceived Genuineness of Motives (Exploratory Study)**

	(1) <i>Based on Perception of Message</i>	(2) <i>Based on Perception of Message</i>	(3) <i>Based on Condition</i>	(4) <i>Based on Condition</i>
	Believing stated motive matches true motive	Believing stated motive matches true goal	Believing stated motive matches true motive	Believing stated motive matches true goal
Prosocial	-0.165*** (0.0512)	-0.154*** (0.0513)		
Instrumental	0.147*** (0.0477)	0.0614 (0.0499)	0.432*** (0.0457)	0.336*** (0.0478)
Mixed	-0.0589 (0.0515)	0.0483 (0.0507)	0.146*** (0.0500)	0.295*** (0.0490)
Constant	0.604*** (0.0363)	0.593*** (0.0365)	0.330*** (0.0341)	0.298*** (0.0332)
Observations	757	757	575	575
R-squared	0.054	0.030	0.131	0.090

Note: OLS Models with robust standard errors in parentheses. In Models 1 and 2, omitted condition is control. In Models 3 and 4, omitted condition is the prosocial. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

**Table S7B: Comparisons in Perceived Genuineness of Motives (Exploratory Study)**

		<i>F</i>	<i>p</i>
Believing stated motive matches true motive	Instrumental vs Prosocial	43.1	<0.001
	Mixed vs Prosocial	4.25	0.04
	Mixed vs Instrumental	18.5	<0.001
Believing stated motive matches true goal	Instrumental vs Prosocial	18.9	<0.001
	Mixed vs Prosocial	16.1	0.001
	Mixed vs Instrumental	0.07	0.79

Note: Wald tests for regressions presented in Table S7A, Models 1 and 2.  $F(1,753)$ .



**Table S8: Perceived Concern for Objectives (Exploratory Study)**

	(5)	(6)	(7)	(8)	(9)	(10)
	<i>Based on Perception of Message</i>			<i>Based on Conditions</i>		
	Most interested in helping...	Most important to help...	Could have a direct impact on...	Most interested in helping...	Most important to help...	Could have a direct impact on...
Prosocial	-0.00204 (0.0509)	-0.0364 (0.0518)	-0.183*** (0.0504)			
Instrumental	-0.107** (0.0510)	0.00385 (0.0513)	-0.127** (0.0499)	-0.120** (0.0489)	0.00343 (0.0496)	0.0618 (0.0479)
Mixed	0.0749 (0.0501)	0.124** (0.0505)	0.0142 (0.0488)	0.175*** (0.0507)	0.228*** (0.0502)	0.349*** (0.0483)
Constant	0.599*** (0.0364)	0.549*** (0.0370)	0.670*** (0.0349)	0.435*** (0.0360)	0.387*** (0.0353)	0.304*** (0.0334)
Observations	757	757	757	575	575	575
R-squared	0.018	0.015	0.029	0.059	0.045	0.092

Note: OLS Models with robust standard errors in parentheses. In Models 5, 6 and 7, omitted condition is control. In Models 8, 9, and 10, omitted condition is the prosocial. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

*Table S9: Contrasts (Exploratory Study)*

	Comparison	Contrast	Std. Err.	<i>t</i>	<i>p</i>
Straightforward	Control vs Prosocial	-0.080	0.133	0.60	0.55
	Control vs Instrumental	0.548	0.132	4.15	0.00
	Control vs Mixed	0.490	0.134	3.66	0.00
	Instrumental vs Prosocial	0.629	0.131	4.82	0.00
	Mixed vs Prosocial	0.571	0.132	4.32	0.00
	Mixed vs Instrumental	-0.058	0.131	0.44	0.66
Honest	Control vs Prosocial	-0.126	0.129	0.98	0.33
	Control vs Instrumental	0.383	0.128	2.99	0.00
	Control vs Mixed	0.347	0.130	2.68	0.01
	Instrumental vs Prosocial	0.509	0.126	4.03	0.00
	Mixed vs Prosocial	0.473	0.128	3.70	0.00
	Mixed vs Instrumental	-0.036	0.127	0.29	0.78
Transparent	Control vs Prosocial	-0.10	0.14	0.75	0.45
	Control vs Instrumental	0.46	0.14	3.34	0.00
	Control vs Mixed	0.38	0.14	2.77	0.01
	Instrumental vs Prosocial	0.56	0.13	4.14	0.00
	Mixed vs Prosocial	0.49	0.14	3.56	0.00
	Mixed vs Instrumental	-0.07	0.14	0.54	0.59
Kind	Control vs Prosocial	0.335	0.131	2.56	0.01
	Control vs Instrumental	-0.127	0.130	0.98	0.33
	Control vs Mixed	0.358	0.132	2.72	0.01
	Instrumental vs Prosocial	-0.462	0.128	3.60	0.00
	Mixed vs Prosocial	0.023	0.130	0.18	0.86
	Mixed vs Instrumental	0.485	0.129	3.76	0.00
Trustworthy	Control vs Prosocial	-0.045	0.130	0.34	0.73
	Control vs Instrumental	0.094	0.129	0.73	0.47
	Control vs Mixed	0.407	0.130	3.12	0.00
	Instrumental vs Prosocial	0.138	0.127	1.09	0.28
	Mixed vs Prosocial	0.452	0.129	3.51	0.00
	Mixed vs Instrumental	0.314	0.128	2.45	0.01
Career (Index)	Control vs Prosocial	0.011	0.137	0.08	0.94
	Control vs Instrumental	-0.014	0.136	0.11	0.92
	Control vs Mixed	-0.026	0.138	0.19	0.85
	Instrumental vs Prosocial	-0.026	0.135	0.19	0.85
	Mixed vs Prosocial	-0.037	0.137	0.27	0.79
	Mixed vs Instrumental	-0.012	0.135	0.09	0.93

## 6. Study 2 Additional Analyses

After completing the exploratory study described above, we conducted a preregistered experiment (Study 2) focused on testing the effects of our treatment messages on perceived genuineness. We report the principal results in the main text. In this section, we present all tables referred to (but not presented) in the main text, supplemental analyses of the main dependent measures, and analyses of all secondary measures.

**Table S10A: Matched-Based Measure of Genuineness**

	(1) Believing stated motive matches true motive	(2) Believing stated motive matches true goal
Prosocial	-0.192*** (0.0495)	-0.196*** (0.0496)
Instrumental	0.0693 (0.0483)	0.0229 (0.0501)
Mixed	-0.0800 (0.0497)	-0.0285 (0.0501)
Constant	0.625*** (0.0350)	0.589*** (0.0356)
Observations	786	786
R-squared	0.039	0.030

Note: OLS Models with robust standard errors in parentheses.

Omitted condition is the control. \*\*\*  $p < 0.01$ , \*\*  $p < 0.05$ , \*  $p < 0.1$

**Table S10B: Matched-Based Measure of Genuineness, Comparisons of Conditions**

		<i>F</i>	<i>p</i>
Believing stated motive matches true motive	Instrumental vs Prosocial	29.30	<0.001
	Mixed vs Prosocial	5.09	<0.024
	Mixed vs Instrumental	9.78	0.002
Believing stated motive matches true goal	Instrumental vs Prosocial	19.62	<0.001
	Mixed vs Prosocial	11.47	<0.001
	Mixed vs Instrumental	1.07	0.302

Note: Wald tests for regressions presented in Table S10A.  $F(1,782)$ .

In addition to the analysis of genuineness, we sought to improve measurement of a key alternative theory, namely that the instrumental motive is seen as more interesting, important, or impactful than the prosocial motive. We asked participants to respond to the following three statements using a 7-point scale (1 = *strongly disagree*, 7 = *strongly agree*): “I think that Organization X’s goal is interesting,” “I think that Organization X’s goal is important,” and “I think that I could make a direct impact on Organization X’s goal.” We submitted these measures to a one-way ANOVA. There were no statistically significant effects or pairwise comparisons (Table S12).

We also maintained two measures from the exploratory study: kind and trustworthy. Unlike the exploratory study, we find no differences in terms of kindness across condition. By contrast, for trustworthy, we find significant differences, with subjects in the prosocial condition perceiving the organization to be more trustworthy than those in the instrumental condition ( $t=3.21$ ,  $p=0.01$ ). This latter result is broadly consistent with our theorizing that organizations that express an instrumental motive seem more genuine.

***Table S11: Descriptive Statistics for Additional Dependent Measures in Study 2***

<b>Dependent Variable</b>	<b>Prosocial (N =201)</b>	<b>Instrumental (N =193)</b>	<b>Mixed (N =200)</b>	<b>Control (N = 192)</b>	<b>Total</b>	<b><i>p</i></b>
Interest in objective	5.11 (1.74)	4.87 (1.47)	4.95 (1.69)	5.04 (1.59)	4.99 (1.63)	0.578
Importance of objective	5.52 (1.52)	5.28 (1.53)	5.39 (1.52)	5.55 (1.42)	5.44 (1.50)	0.286
Ability to have an impact on objective	5.00 (1.70)	5.02 (1.65)	5.00 (1.58)	5.28 (1.43)	5.07 (1.60)	0.244
Kind	5.05 (1.41)	5.03 (1.24)	5.03 (1.39)	5.12 (1.25)	5.06 (1.32)	0.856
Trustworthy	5.00 (1.42)	5.42 (1.13)	5.05 (1.44)	5.28 (1.23)	5.19 (1.32)	0.004

Note: P-values from one-way ANOVA. See Table S12 for all pairwise comparisons. For the interest, importance, and impact measures, a design error in the survey allowed some subjects to leave responses blank. For this reason, the number of valid responses for the importance measure is 777 and for the impact measure is 785.

**Table S12: Study 2 Contrasts**

	<b>Comparison</b>	<b>Contrast</b>	<b>Std. Err.</b>	<b><i>t</i></b>	<b><i>p</i></b>
Genuineness Index	Control vs Prosocial	-0.252	0.124	2.03	0.043
	Control vs Instrumental	0.261	0.125	2.08	0.038
	Control vs Mixed	-0.208	0.124	1.68	0.094
	Instrumental vs Prosocial	0.512	0.124	4.14	<0.001
	Mixed vs Prosocial	0.044	0.123	0.35	0.723
	Mixed vs Instrumental	-0.469	0.124	3.78	<0.001
Important	Control vs Prosocial	-0.029	0.151	-0.19	0.846
	Control vs Instrumental	-0.262	0.153	1.71	0.087
	Control vs Mixed	-0.157	0.152	1.04	0.301
	Instrumental vs Prosocial	-0.232	0.151	1.54	0.124
	Mixed vs Prosocial	-0.127	0.150	0.85	0.395
	Mixed vs Instrumental	0.105	0.151	0.69	0.488
Interesting	Control vs Prosocial	0.069	0.165	0.42	0.676
	Control vs Instrumental	-0.163	0.167	0.98	0.329
	Control vs Mixed	-0.082	0.165	0.50	0.619
	Instrumental vs Prosocial	-0.232	0.165	1.41	0.160
	Mixed vs Prosocial	-0.151	0.163	0.93	0.355
	Mixed vs Instrumental	0.081	0.165	0.49	0.624
Impact	Control vs Prosocial	-0.273	0.161	1.69	0.091
	Control vs Instrumental	-0.262	0.163	1.61	0.108
	Control vs Mixed	-0.277	0.161	1.72	0.086
	Instrumental vs Prosocial	0.011	0.161	0.07	0.948
	Mixed vs Prosocial	-0.005	0.159	0.03	0.975
	Mixed vs Instrumental	-0.016	0.161	0.10	0.923
Kind	Control vs Prosocial	-0.075	0.134	0.56	0.574
	Control vs Instrumental	-0.099	0.135	0.73	0.464
	Control vs Mixed	-0.100	0.134	0.75	0.456
	Instrumental vs Prosocial	-0.024	0.134	0.18	0.858
	Mixed vs Prosocial	-0.025	0.132	0.19	0.852
	Mixed vs Instrumental	-0.001	0.134	0.01	0.995
Trustworthy	Control vs Prosocial	-0.281	0.133	2.12	0.034
	Control vs Instrumental	0.144	0.134	1.07	0.284
	Control vs Mixed	-0.226	0.133	1.71	0.089
	Instrumental vs Prosocial	0.425	0.132	3.21	0.001
	Mixed vs Prosocial	0.055	0.131	0.42	0.675
	Mixed vs Instrumental	-0.370	0.132	2.79	0.005

## **7. Study 2 Preregistration.**

Posted on OSF: <https://osf.io/er365/>