The Good-on-Paper Effect: How the Decision Context Influences Virtuous Behavior

Maferima Touré-Tillery
Lili Wang

Maferima Touré-Tillery (m-touretillery@kellogg.northwestern.edu) is associate professor of marketing at Northwestern University, Kellogg School of Management, Evanston, IL 60208, USA. Lili Wang (lw122@zju.edu.cn) is associate professor of marketing at Zhejiang University, School of Management, Hangzhou, China 310058.

The authors are grateful to the university teams who provided support at various stages in the development of this work. The authors thank Echo (Menglu) Dong, Ege Gokturk, Emily Kelleher, Jessie Jia, Katherine Feng, Senere Ong, Xiao Su, Xuan Zhang, Yuki (Hanyu) Chen, and Yuting Pang for their indispensable research assistance. This research was supported in part by funding from the Richard M. Clewett Endowed Professorship (Kellogg School of Management) awarded to the first author, and by grants from the National Natural Science Foundation of China (Grant No. 71572179) awarded to the second author. Both authors contributed equally to this research.

Correspondence concerning this article may be addressed to Lili Wang (lw122@zju.edu.cn) or to Maferima Touré-Tillery (m-touretillery@kellogg.northwestern.edu).
The Good-on-Paper Effect: How the Decision Context Influences Virtuous Behavior

Abstract:
We find that people are more likely to make virtuous decisions on paper than on a digital device because they perceive choices on paper as more real (i.e., tangible, actual, and belonging to the physical rather than the virtual world), and hence as more self-diagnostic (i.e., representative of who they are). We first show people express more interest in donating and volunteering (Study 1a and 1b), are more likely to donate (Study 2), and put more effort into helping a charitable cause (Study 3) when these choices occur on paper (vs. tablet)—a pattern of decision-making we label the “good-on-paper effect.” Study 4 extends these findings to book choices (highbrow vs. lowbrow) and to a device interaction that closely mimics writing on paper (i.e., tablet with digital pen). In the context of volunteering decisions, we then provide evidence for the sequential mediating roles of perceptions of realness and self-diagnosticity in the good-on-paper effect (Study 5 and Studies 6a and 6b). Finally, we show that chronic (Study 7) and situational (Study 8) perceptions of self-diagnosticity moderate this effect in the contexts of environmental protection and food choices (healthy vs. indulgent), respectively. We discuss the theoretical and practical implications of these findings.

Keywords:
Good-on-paper effect, digital device, paper, realness, self-diagnosticity, virtuous behavior
1. Introduction

Digital devices such as computers, tablets, and smart phones are increasingly pervasive in today’s societies. According to recent reports by Statista (2020a, 2000b), the number of tablet users worldwide grew from about 840 million in 2014 to over 1.28 billion in 2018, whereas the number of smartphone users rose from 2.5 billion in 2016 to around 3.5 billion in 2020—a 40% increase! Many decisions that used to occur solely on paper (e.g., choosing what to order from a restaurant menu, deciding whether and what to donate to charity) now also take place on these digital devices, with different types of devices having distinct influences on emotions, cognitions, and choices (Brasel and Gips 2014; Ghose et al. 2013; Shen et al. 2016). However, the use of paper is still prevalent in everyday contexts, ranging from restaurant menus to charitable pledge forms. Recognizing the importance of the medium through which an action is performed, prior research has compared the effects of using digital devices to using paper on reading, learning, and test-taking performances (e.g., Clariana and Wallace 2002; DeAngelis 2000; Mangen, Walgermo, and Brønnick 2013; Mazzeo et al. 1991; Watson 2001). The present research broadens these lines of inquiry to the domain of virtuous behavior by asking the question: How does using a digital device instead of paper influence people’s likelihood to engage in virtuous behavior?

A virtuous behavior is an act or a choice that provides delayed personal benefits (e.g., good health, financial stability, academic success) or indirect benefits (e.g., helping others, protecting the environment) and adheres to standards such as morals, values, and ideals (see Baumeister, Vohs, and Tice 2007; Read, Loewenstein, and Kalyanaraman 1999). Such behaviors are typically costly, requiring resources such as time, money, effort, or self-control. For example, a consumer making a virtuous choice might sign-up to donate her hard-earned money to charity.
(instead of spending it on herself) or to volunteer during her free time (instead of resting or partying). Other examples of virtuous behaviors include forgoing an indulgent cheeseburger in favor of a healthier salmon dish or choosing a highbrow/intellectual book that would contribute to one’s growth and learning (instead of a lowbrow/fun book that would only entertain in the moment). Virtuous behaviors have self-signaling potential, that is, they can send positive signals to the self (and others) about a person’s traits and values, and hence boost the self-concept (Bem 1972; Dhar and Wertenbroch 2012; Diener and Srull 1979; Prelec and Bodner 2003).

We propose that people will be more likely to engage in virtuous behavior on paper than on a digital device—a pattern of decision-making we label the “good-on-paper effect”—because they will perceive their choices on paper as more real (i.e., more tangible, actual, and belonging to the physical rather than the virtual world; Deleuze 1990; Girvan 2018; Lehdonvirta 2010; Shields 2003), and thus as more representative of who they are (i.e., self-diagnostic; Bodner and Prelec 1996; Touré-Tillery and Light 2018). Indeed, a study using functional magnetic resonance imaging to examine the brain activities of participants viewing ads as physical mail on paper cards or as digital messages on a computer screen found that materials viewed on paper (vs. digital) media produced more activity in the left and right parietal areas of the brain. Because these areas are associated with the integration of visual and spatial information, the authors concluded that participants perceived paper (vs. digital) materials as more real, tangible, and belonging in space (see Millward Brown Study 2009).

The fact that digital contexts are inextricably linked to virtuality supports this notion (Yoh 2001). Indeed, the virtual is conceptualized as not actual (Deleuze 1990), not real (Lehdonvirta 2010), and not of the physical, natural, or material world (Shields 2003). Furthermore, virtual is often used to describe simulated experiences, that is, experiences that are
almost—but not quite—real (i.e., fictitious) and that lack physical properties beyond the screen (see also, Flavián, Ibáñez-Sánchez, and Orús 2021; Girvan 2018; Zavoleas 2006). In sum, we posit that people will perceive a decision context (i.e., the setting in which a decision is made) that is digital/virtual as less real than one that is on paper/physical.

Moreover, the knowledge or perception that an event is real (vs. unreal, virtual) influences judgment and decision making (Bostyn et al. 2018; FeldmanHall et al. 2012; Francis et al. 2016; Imas 2016; Imas and Loewenstein 2018; Patil et al. 2014). For example, Raghubir and Srivastava (2008) have advanced the greater perceived realness and vividness of using cash (physical payment) relative to a credit card (digital payment) as a contributing factor to differences in pain of paying, spending (Prelec and Simester 2001; Thomas, Desai, and Seenivasan 2010) and post-purchase satisfaction (Shah et al. 2016) between these payment modalities. At a basic level, real choices (e.g., making a $50 donation to charity) are indeed more consequential than unreal or fictitious ones (e.g., imagining making the $50 donation or pretending to make the $50 donation). We propose decisions that appear more (vs. less) real will seem more diagnostic of a person’s traits and characteristics, such that a person making a real $50 donation will be deemed—and will deem herself—more generous than a person making an unreal, fictitious $50 donation.

Thus, drawing from the proposition that decision contexts on paper seem more real than those on digital devices, we advance that people will view decisions on paper (vs. digital devices) as more self-diagnostic—a notion consistent with recent work showing that physical goods (e.g., printed books) have a greater capacity to garner an association with the self than their digital counterpart (e.g., kindle book; Atasoy and Morewedge 2018). In turn, these differential perceptions of self-diagnosticity will have consequences for decision-making on
paper (vs. digital devices). Research shows when people consider their actions self-diagnostic, they are more likely to behave in line with standards such as ideals, values, morals, or social expectations—to maintain a positive self-concept (i.e., self-signaling; Bryan et al. 2011; Gneezy et al. 2012; Prelec and Bodner 2003; Savary et al. 2015). In sum, we propose the “good-on-paper effect,” whereby people will exhibit more virtuous (i.e., “good”) decision-making on paper than on a digital device (H1) and posit this effect will occur because decisions made on paper will seem more real, and hence more self-diagnostic than those made on a digital device (sequential mediation; H2).

Research shows people vary in the extent to which they see their actions as self-diagnostic, which in turn determines their tendency to engage in self-signaling virtuous behavior (Bodner and Prelec 1996; Touré-Tillery and Light 2018). At higher levels of self-diagnosticity, people tend to see most of their actions as representative of who they are and hence behave mostly virtuously, whereas the opposite is true for lower levels of self-diagnosticity. In contrast, perceptions of self-diagnosticity should be more malleable at moderate levels, such that people should be more responsive to situational cues of self-diagnosticity, such as the decision context. Thus, given the central role of perceptions self-diagnosticity in our theorizing, we propose that chronic levels of self-diagnosticity will moderate the good-on-paper effect, such that it will occur at moderate—but not at high or low—levels of self-diagnosticity (H3).

Another important moderator emerges from this last proposition: whether the target of a choice is the self or another person. Indeed, choices made for another person tend to depend on—often inaccurate—inferences about the recipient’s preferences or about most people’s preferences (Kray 2000; Kray and Gonzalez 1999). For example, Laran (2010) showed that when making food choices for others, choosers tend to infer recipients will most likely indulge
rather than exercise self-control, and hence choose accordingly for others (see also Lu, Liu, and Fang 2016). In sum, choices made for others are reflective of the chooser’s beliefs about others (rather than of the chooser’s preferences), and thus such choices should generally be seen as less representative of the chooser’s personal traits and characteristics than choices made for the self (See also Touré-Tillery and Kouchaki 2021). Therefore, we expect the good-on-paper effect to occur only when consumers are making personal choices that can reflect their own preferences and characteristics, but not when consumers are making choices for others—i.e., non-self-diagnostic choices (H4).

Finally, we note that a construal-level-theory (CLT; Trope and Liberman 2003) perspective might offer an alternative set of predictions. Indeed, according to CLT, realness is one of four types of psychological distances, with the other three being temporal, spatial, and social distance (Trope and Liberman 2010). CLT advances that people represent objects that are psychologically close (e.g., real events) at a lower, more concrete level of construal, whereas they conceive of objects that are psychologically distant (e.g., events that are not real) at a higher, more abstract level. Furthermore, psychologically distant (vs. proximal) situations are more likely to activate values and moral principles (Eyal, Liberman, and Trope 2008), which are in turn more likely to guide people’s judgments (Agerström and Björklund 2009). Within this perspective, people should behave more virtuously for seemingly less real (i.e., higher-level-construal) decisions on a digital device than for seemingly more real (i.e., lower-level-construal) decisions on paper. Contrary to this alternative prediction, we expect people to make more virtuous choices on paper (vs. digital devices) due to differential perceptions of realness and self-diagnosticity. In addition to enriching the literature on digital technologies by providing new insights into the consequences of using paper compared to digital devices, the present research
extends knowledge on self-signaling and virtuous behavior by identifying the perception of realness as an antecedent to both phenomena.

2. Overview of Studies

We test our hypotheses in ten studies\(^1\) using various types of participant populations (American and Chinese adults and students), digital devices, and virtuous behaviors. The first set of studies explored the good-on-paper effect (H1) in various prosocial contexts: expressing an interest in giving (Study 1a) and volunteering (Study 1b), making monetary and in-kind donations (Study 2; preregistered), and earning to give to a charitable cause (Study 3; preregistered). Study 4 (preregistered) extended these findings beyond prosocial decisions to the context of book choices and to a device interaction that more closely mimics writing on paper (i.e., tablet with pen). In the context of volunteering decisions, the next two studies tested the sequential mediating roles of perceptions of realness and self-diagnosticity on the good-on-paper effect (H2) using classic mediation (Study 5; preregistered) and causal-chain mediation (Studies 6a, preregistered and Study 6b) approaches. The last two studies tested the moderating role of perceptions of self-diagnosticity (H3), operationalized through (i) an individual difference measure (Study 7, preregistered) and (ii) the target of the choice: choosing for oneself as opposed to choosing for another person (i.e., an act low in self-diagnosticity; H4; Study 8). Whereas Study 7 explored a different type of prosocial behavior (environmental protection), Study 8 extended our investigation to the domain of food choices.

For all experiments, we estimated a minimum required sample size of 60 participants per experimental condition to achieve a power of .80 at an alpha level of \(p = .05\). This sample size

\(^1\) All studies except Studies 2, 4, and 5 were conducted before the emergence of the COVID-19 pandemic.
calculation was based on the results of Study 1b. To maximize power, we collected between 70 and 100 responses per experimental condition for each study. We note that our final sample sizes were often slightly higher or lower than the pre-registered numbers due to the difficulty inherent in keeping an exact count in the type of multimedia field experiments we conducted. All experimental stimuli are available in the Supplemental Materials. De-identified data for all experiments are posted on the Open Science Framework (OSF) and can be accessed from the following link: https://tinyurl.com/good-on-paper.

3. Study 1: Expressing Interest in Helping

This study tested the good-on-paper hypothesis (H1) in the contexts of signing up to donate (Study 1a) and to volunteer (Study 1b). Participants read a charitable appeal presented on paper or on a digital device (tablet) and indicated their willingness to help. We predicted participants making the decision on paper (vs. tablet) would be more willing to help.

3.1. Study 1a: Signing up to Donate

3.1.1. Methods. We recruited 200 adults (gender: 116 female, 84 male; age: \( M = 37.87, SD = 16.10 \)) in the downtown area of a large American city and gave them candy bars as tokens of appreciation. We did not exclude any participant.

The study employed a 2-level (decision context: paper vs. digital device) between-subjects design. Research assistants (RAs) approached passersby about completing a short study. Those who agreed to participate received either a pen and a paper survey titled “A Short Paper and Pen Survey” or a tablet (iPad) displaying an online survey titled “A Short Tablet Survey.” We designed both surveys using similar formatting and typeface (see Appendix B-I). In the paper condition, participants used a pen to check off small circles or checkboxes corresponding to their choices, or to handwrite/inscribe their answers, whereas in the tablet condition,
participants used their fingers to tap on radio buttons or checkboxes, or to type their answers using the tablet’s digital keyboard.

Unlike participants completing the survey on a tablet, those completing the survey on paper did not have the option to click a button to exit out of the survey (although they could stop at any time in both conditions). This inability to conceal their answers when handing the survey back to the RA might lead them to perceive their responses as less private or less anonymous, which might in turn elicit socially desirable responding. Thus, after providing the survey materials to participants, the RA stepped aside to let participants complete the study with some degree of privacy.

To ensure participants experienced the decision context (paper or tablet) before the critical dependent measure, the survey started with demographic questions (gender, age, whether they reside in the US\(^2\), and whether English is their native language) and the following (size-of-characters) question\(^3\): “When you write on paper [on a tablet], how small or large do your characters tend to be?” (1 = very small, 7 = very large). Next, participants completed a filler “preference survey,” designed to appear as the main part of the study. They made five separate selections: coffee vs. tea; football vs. basketball; gelato vs. frozen yogurt; minty gum vs. fruity gum; and winter vs. summer. The options were intentionally neutral—such that one was not more virtuous than the other—so as not to influence the subsequent measure of virtuous behavior (e.g., through licensing or consistency; see Mullen and Monin 2016). Participants then saw a

---

\(^2\) These data were collected in a section of the city often visited by tourists. We asked participants to indicate whether they resided in the US. All participants residing outside of the US (N = 22) declined to provide their email address. We kept them in the analyses and note that excluding them did not change the pattern or the significance of the results.

\(^3\) Decision context did not influence participants’ responses to the size-of-character question \((M_{\text{paper}} = 3.94, SD = 1.11; M_{\text{digital}} = 3.83, SD = 1.02; t(198) = .78; p = .438; d = .10)\), and answers to this question had no effect on sign-up rate \((b = -.27 (.20), z = -1.36, p = .172; \text{odds ratio} = .76)\).
charitable appeal soliciting donations for No Kid Hungry, a not-for-profit organization that provides meals to disadvantaged children in America (see Appendix B-II).

The message ended with “If you are interested in donating to this organization, please provide your email address below for a follow-up*” Below this call to action was a text box in which participants could provide their email address. To further ensure a sense of anonymity and minimize self-presentation concerns, the asterisk next to “follow-up” referred participants in the paper [tablet] condition to the following information displayed below the text box: “*For your privacy, please fold this sheet before handing it back [please click the arrow to exit the survey before handing back the tablet]. Thank you!” Our measure of virtuous behavior was whether participants expressed interest in donating by providing their email address4 (i.e., sign-up rate).

2.1.2. Results. A logistic regression5 of sign-up rate (0 = did not provide email address, 1 = provided email address) on decision context (0 = paper, 1 = digital device) revealed a greater proportion of participants expressed interest in donating on paper (20.00%) than on tablet (7.27%; b = -1.16(.45), z = -2.56, p = .010; odds ratio = .31), a pattern consistent with the good-on-paper hypothesis. Additionally, in this study and in subsequent studies, the pattern and significance of the good-on-paper effect was unchanged when we included all available demographic variables (e.g., in Study 1a: age, gender, whether they reside in the US, and whether English is their native language) as covariates in the analysis (see Appendix C-I).

We note that, for a variety of reasons, older participants might perceive digital devices differently from younger ones (i.e., digital natives), which might elicit different responses on

---

4 One participant (paper condition) provided a clearly invalid email address (“$1.ABNOS. SKM”) and we coded this response as: 0 = did not provide email address.

5 We note Gomila (2021) recently argued that logistic regression is potentially problematic when estimating causal effects and generally suggests supplementing or replacing such analyses with linear regressions. Thus, we re-ran our key analyses using linear regressions instead if logistic regressions and found the pattern and significance of our key results were unchanged.
paper (vs. tablet). Thus, we tested whether age moderated the good-on-paper effect. We found no such moderation in this study or in subsequent studies (see Appendix C-I), suggesting that the tendency to make more virtuous decisions on paper (vs. tablet) is not a function of age.

3.2. Study 1b: Signing up to Volunteer

3.2.1. Methods. We recruited 194 students (gender: 86 female, 108 male; age: $M = 20.88$, $SD = 2.69$) in the study area of a large university in China and gave them candy bars in appreciation for their time. We excluded two participants from the analysis because they indicated they would be graduating within the next two weeks and leaving the city, which would preclude them from volunteering locally (our dependent variable). One hundred ninety-two participants (gender: 85 females, 107 males; $M_{\text{age}} = 20.87$, $SD_{\text{age}} = 2.70$) remained for the subsequent analyses.

The study employed a 2-level (decision context: paper vs. digital device) between-subjects design used the same recruiting procedure and survey titles as Study 1a. To ensure a sense of anonymity and minimize potential self-presentation concerns, after providing the survey materials to participants, the research assistant (RA) asked them to return their completed surveys to a box on a separate desk in the study area. Then, the RA stepped away to let each participant complete the study with some degree of privacy. As in Study 1a, after completing some brief demographic questions (gender and age) and the size-of-characters question, participants read a charitable appeal soliciting volunteering time for a well-known local orphanage (see Appendix B-III). To measure virtuous behavior, we asked participants (a) to

---

6 Retaining these participants in the analysis did not change the pattern or the significance of the results.
7 Decision context did not influence participants’ responses to the size-of-character filler question ($M_{\text{paper}} = 3.90$, $SD = 1.17$; $M_{\text{digital}} = 3.72$, $SD = 1.04$; $t(190) = 1.11; p = .271$; $d = .16$), and answers to this question had no effect on sign-up rate ($b = -.10(.15)$, $z = -.70$, $p = .486$; odds ratio = .90) or time commitment ($b = .33(.43)$, $t(186) = .76$, $p = .449$; $d = .11$).
indicate their interest in volunteering for the organization by leaving their cellphone numbers so the organization could contact them directly, and (b) to specify how many hours they could commit to volunteer in one month. Thus, our measures of virtuous behavior were whether participants provided their cellphone numbers (i.e., sign-up rate) and the number of hours they committed (i.e., time commitment).

Finally, we note that people (especially digital natives) might be more comfortable with tablets than with paper (i.e., digital fluency), which might in turn interact with the decision context to influence judgments and choices. To test this possibility, we included three questions assessing digital fluency (α = .85): (a) “Do you write more often on paper or on an electronic device (e.g., phone, tablet, and computer)?” (1 = definitely more often on paper, 7 = definitely more often on electronic device); (b) “In general, is it easier for you to write on paper or on an electronic device” (1 = definitely easier on paper, 7 = definitely easier on electronic device); and (c) “In general, is it more comfortable for you to write on paper or on an electronic device?” (1 = definitely more comfortable on paper, 7 = definitely more comfortable on electronic device).

3.2.2. Results. Consistent with the good-on-paper hypothesis, a logistic regression of sign-up rate (0 = did not provide their cellphone number, 1 = provided their cellphone number) on decision context (0 = paper, 1 = digital device) showed a greater proportion of participants signed-up to volunteer on paper (34.38%) than on the tablet (20.83%; b = -.69(.33), z = -2.08; p = .037; odds ratio = .50). Furthermore, a t-test of time commitment by decision context showed a similar pattern: participants committed more volunteering hours on paper (M = 4.37, SD = 7.87) than on the tablet (M = 2.24, SD = 4.99; t(186) = 2.21, p = .028; d = .32).

---

8 Four participants (three in the paper condition and one in the tablet condition) left their cellphone numbers but did not indicate the number of hours they would be willing to volunteer.
Next, to explore the possible role of digital fluency in the good-on-paper effect, we ran a logistic regression of sign-up rate on decision context, digital fluency (mean-centered), and their interaction. The analysis revealed a marginal effect of decision context on sign-up rate ($b = -.61(.34), z = -1.80, p = .072$; odds ratio = .54). Neither the main effect of digital fluency ($b = -.03(.14), z = -.23, p = .820$; odds ratio = .97) nor the interaction between decision context and digital fluency ($b = -.02(.22), z = -.10, p = .921$; odds ratio = .98) approached significance. Similarly, a linear regression of time commitment on decision context, digital fluency (mean-centered), and their interaction revealed a marginal effect of decision context ($b = -1.75(.89), t(182) = -1.95, p = .052$; $d = .29$) but no effect of digital fluency ($b = .097(.41), t(182) = .24, p = .813$; $d = .036$) and no interaction between decision context and digital fluency ($b = -.52(.57), t(182) = -.91, p = .363$; $d = .13$). Importantly, these (null) results suggested the good-on-paper effect is unlikely to stem from or be influenced by digital fluency.

3.3. Discussion

Using participants from different countries (i.e., United States and China), Studies 1a and 1b provided initial support for the good-on-paper effect (H1). We found people were more likely to sign-up to donate their money (Study 1a) and to volunteer their time (Study 1b), and committed to volunteer more time to charity (Study 1b) when these prosocial decisions were made on paper (vs. a digital device). Additional analyses (Study 1b) indicated participants’ level of digital fluency—that is, their frequency and ease of using digital devices relative to paper—could not explain this effect. Furthermore, in both studies, we took steps to ensure that participants felt they were completing their survey with a degree of privacy, such that they would not perceive their responses as more or less anonymous depending on the experimental condition (paper vs. tablet)—which might elicit a social signaling motive. Notably, the presence of others
was held constant in all experimental conditions, such that we would expect any social signaling effect to manifest as a main effect in both conditions. In Supplemental Study 2 (see Appendix A-II), we also measured perceived anonymity and found no effect of the experimental condition on this perception, indicating the good-on-paper effect is unlikely to stem from differential perceptions of anonymity. Thus, we believe our studies capture the effect of decision context (paper vs. digital device) on self-signaling above and beyond any incidental social-signaling effects which might naturally arise from the mere presence of others in the experimental context.

In Studies 1a and 1b (as well as Studies 5 and 6b), participants were told the charity intended to collect their personal data to contact them later, a common practice for non-profit organizations. This measure captured a consequential decision because personal data has become extremely valuable to corporations, and consumers are increasingly aware of this fact, and concerned about their privacy (Feijóo et al. 2014; Schwartz 2004). Thus, agreeing to provide their email address to a charity in this way is not a trivial decision. We take it as an expression of their desire to help—potentially, at some cost to their privacy. However, one might argue that this measure of prosocial/virtuous behavior is somewhat confounded with trust or privacy concerns and might indicate that consumers are less trusting or more worried about their privacy when providing their email addresses electronically (vs. on paper). In the next three studies, we tested the robustness of the good-on-paper effect by moving to more consequential and unequivocal measures of virtuous behavior.

In Studies 1a and 1b, the decision to help was both prompted and expressed in the same context (i.e., on the same medium), such that participants who read the charitable appeal on paper (on the tablet) expressed their decision to donate on paper (on the tablet). This raises the question of whether the good-on-paper effect is a function of the medium prompting the decision.
or of medium through which the decision is implemented. We note that our theorizing refers to the realness of the decision context—that is, the setting in which the decision is made—which may or may not be the setting in which the decision is expressed but is likely to corresponds to the setting in which the decision is prompted. In particular, the types of decisions we explore in this article (donations, food choices) are typically made while reviewing or recalling prompts, such that a person will decide whether to donate or what foods to buy while reading a charitable appeal or a menu or after recalling these prompts (unless new information is presented in the intervening time). After deciding what to do, the person will then take steps to express or implement her decision either in the same context (e.g., donate cash after reading a paper brochure at a charity event; order food online after reviewing the menu on a restaurant’s website) or in a different context (donate online after reading a paper brochure at a charity event; call to order food after reviewing the menu on a restaurant’s website). Indeed, research on deliberation and implementation shows that once people have decided what to do (in the deliberative phase), they tend to become narrow-mindedly focused on executing their decision (in the implementation phase; see Gollwitzer, Heckhausen, and Steller 1990).

In sum, we expect the good-on-paper effect to typically reflect the context in which the decision was prompted and made rather than the context in which the decision was implemented or expressed. For example, we would expect a person who read a charitable appeal on paper (vs. digital device) to decide whether she will donate or not in that initial context, such that she would be more likely to donate, whether she implements her decision online, verbally (e.g., on the phone), or on paper (e.g., by writing a check). In the next study, we address this question by testing the good-on-paper effect in a paradigm in which participants first make the decision to
donate on paper or on a tablet, and then make a monetary donation online (using their mobile phones) or an in-kind donation offline (at a different physical location).

4. Study 2: Granting a Wish

This study tested the good-on-paper hypothesis (H1) in the more consequential context of donating monetary or in-kind gifts. As part of a student-led charity event, potential donors read a charitable appeal presented on paper or on a digital device (tablet) and decided whether to help or not. We predicted participants making the decision on paper (vs. tablet) would be more likely to help. We preregistered the design and analysis plan for this study at:


4.1. Methods

4.1.1. Participants. We recruited 431 adults (gender: 254 female, 177 male; age: $M = 21.32$, $SD = 3.69$; student status: 424 students, 6 non-students, 1 undisclosed) on the campus of a large university in China as part of a three-day student-led charity event. We did not exclude any participant.

4.1.2. Design and Procedure. The study employed a 2-level (decision context: paper vs. digital device) between-subjects design. To test the good-on-paper hypothesis (H1) in a naturalistic context, we collaborated with a student-led charitable organization at a large university in China, called the “Love Club.” The organization launched a year-end “Big Hands Holding Small Hands” charity event with the mission to fulfill the wishes of disadvantaged children in various regions of China, by soliciting in-kind donations of the items on the children’s wish lists. Research assistants and members of the charitable organization approached passersby about completing a short survey, which followed the same template as Study 1a and 1b (paper vs. tablet/iPad), but without the size-of-character question. After completing a brief demographic questionnaire (gender, age, student status and monthly allowance or income) and
filler preference questions similar to the ones used in Study 1a (e.g., coffee vs. tea), potential donors read a charitable appeal to help a child in need by purchasing items from his or her wish list. The appeal was followed by a list of 36 wish lists corresponding to 36 different children. Wish lists included one or more items, and most wish lists featured school supplies (see Appendix B-IV).

Potential donors were asked to indicate which wish list(s) they would like to fulfill (if any) and learned that they could fulfill a wish list (a) by purchasing the item(s) and sending them to the student association (at the address provided) within three days, or (b) by making a monetary donation on the spot to allow the organization to fulfill the wish list. In this latter case, donors used their own mobile phones to donate, and the amount of the monetary donation was up to each donor. For ease of tracking, the organization set up two different mobile/online payment accounts (through Alipay), one for each experimental condition. Our measure of virtuous behavior was whether participants chose to help by selecting at least one wish list to fulfill (i.e., sign-up rate). Furthermore, working with the charitable organization, we were able to track the number of participants who followed through on their commitment by donating the corresponding monetary amount on the spot or by bringing the selected wish-list items to the organization within three days (i.e., donation rate).

4.2. Results

4.2.1. Sign-up rate. A logistic regression of sign-up rate (0 = did not select any wish list, 1 = selected at least one wish list) on decision context (0 = paper, 1 = digital device) revealed a greater proportion of participants signed up to fulfill a child’s wish list when they made this decision on paper (55.30%) compared to on tablet (43.93%; \( b = -.46(\cdot19), z = -2.36, p = .018; \) odds ratio = .63), a pattern consistent with the good-on-paper hypothesis.
4.2.2. Additional analyses. After this pre-registered analysis, we turned to actual donations to conduct two exploratory analyses. First, we found that, of the 214 participants who signed up to help by selecting a wish list, 16.36% followed through (i.e., donated money or items from the selected wish list(s)). Notably, the follow-through rate was higher in the paper condition (20.83%) than in the tablet condition (10.64%; $\chi^2(1) = 4.01, p = .045$; odds ratio = .45), which lent further support to the good-on-paper hypothesis.

Second, a chi-square analysis of actual donation rate ($0 = did not donate; 1 = donated money or items from selected wish list(s)) by decision context showed that a greater proportion of participants donated in the paper condition (11.52%) than in the digital device condition (4.67%, $\chi^2(1) = 6.77, p = .009$; odds ratio = .38).

4.3. Discussion

This field study provided further support for the good-on-paper hypothesis (H1), showing that people are not only more likely to make prosocial commitments on paper (vs. digital device), but are also more likely to follow-through on such commitments made on paper (vs. digital device). Using a large dataset of alumni donations to their alma mater, Supplemental Study 1 (Appendix A-I) provided another test of the good-on-paper hypothesis in a naturalistic context. Analyses of this dataset suggested that alums tended to give more money in response to mail solicitations than to email solicitations, controlling for the demographic and classification information that were available in the dataset. The next study explored a different type of prosocial behavior: earning to give (see MacAskill 2015).

5. Study 3: Earning to Give by Solving Anagrams

This study provided another test of the good-on-paper hypothesis (H1) in the context of volunteering, or more specifically, earning-to-give to a charitable organization. Participants read a charitable appeal presented on paper or on a digital device (tablet) and had the opportunity to
help by earning money—through solving anagrams—to give to the charity. We predicted participants making the decision on paper (vs. tablet) would be more willing to help. We preregistered the design and analysis plan for this study at:


5.1. Methods

5.1.1. Participants. We recruited 195 participants over the age of 18 (gender: 126 female, 66 male, 3 undisclosed; age: Median = 18-24, Mode: 18-24) on the campus of a large American university and gave them various snacks as tokens of appreciation. We did not exclude any participant.

5.1.2. Design and Procedure. The study employed a 2-level (decision context: paper vs. digital device) between-subjects design. Research assistants approached passersby and those who agreed to participate received a survey on a tablet (iPad) or on paper titled “Words for Charity.” As in previous studies, both surveys used similar formatting and typeface. However, unlike in Studies 1a and 1b—and to minimize potential demand effects—the survey title did not include the words “paper” or “tablet,” and we omitted the size-of-character question. Participants indicated some basic demographics (age, student status, and gender) before reading a charitable appeal about LitWorld, an organization working to eliminate illiteracy worldwide.

As a measure of virtuous behavior, participants were asked to volunteer to solve anagrams to raise money for LitWorld (i.e., an earning-to-give model of prosocial behavior). Indeed, effort is a common measure of prosocial behavior, both from a theoretical and practical perspective (e.g., volunteering; earning-to-give; see Brown, Meer, and Williams 2019; Imas 2014; MacAskill 2015). Specifically, for each anagram participants solved, 5 cents would be donated to LitWorld on their behalf. They could solve up to 20 anagrams—for a total donation of
$1—and they could stop at any point (see Appendix B-V). This paradigm allowed us to measure virtuous behavior both (a) as whether participants chose to help by solving at least one anagram (i.e., volunteering rate), and (b) as the number of anagrams they solved (volunteering effort, which ranged from 0 to 20). Proceeds from this study went to LitWorld.

5.2. Results

A logistic regression of volunteering rate ($0 = solved no anagrams, 1 = solved at least one anagram) on decision context ($0 = paper, 1 = digital device) showed participants were more likely to help (92.39%) on paper than on tablet (79.61%; $b = -1.14(.46), z = -2.45, p = .014; odds ratio = .32). Furthermore, a $t$-test of volunteering effort (number of anagrams solved correctly) by decision context showed participants solved more anagrams on paper ($M = 8.45, SD = 5.59$) than on tablet ($M = 4.69, SD = 3.85; t(193) = 5.52, p < .001; d = .79$).

5.3. Discussion

This study provided yet more support for the good-on-paper hypothesis ($H1$). We argue this effect occurs due to greater perceptions of realness and hence self-diagnosticity for decisions on paper (vs. digital device). However, other mechanisms may be at play. In the next studies, we begin to rule out alternative mechanisms and rule in our proposed underlying mechanisms. Furthermore, the next study extended our findings beyond prosocial decisions to a different type of virtuous behavior: choosing highbrow (instead of lowbrow) books.

6. Study 4: Choosing Highbrow Books

This study moved to the context of book choices to test the good-on-paper hypothesis ($H1$). Furthermore, we aimed to rule out an important alternative explanation for this effect: psychological ownership. Indeed, in our studies so far, we have compared responses using paper and pen to responses using a tablet without a pen. In the paper conditions, participants used a pen
to check off small circles or checkboxes corresponding to their choices, or to handwrite/inscribe their answers. By contrast, in the tablet conditions, participants used their fingers to tap on radio buttons or checkboxes, or to type their answers using the tablet’s digital keyboard.

Research shows that physical (vs. digital) objects (e.g., printed book vs. kindle book; Atasoy and Morewedge 2018) and digital devices with touch (vs. no touch) interface (e.g., tablet vs. computer; Brasel and Gips 2014) elicit a greater sense of psychological ownership—i.e., the sense that something belongs/relates to oneself. Notably, these perceptions of psychological ownership occur because manipulating and touching an object increases people’s perception of control over the object. Thus, our paper and tablet conditions should elicit similar degrees of psychological ownership, because they both allow for manipulating and touching. However, using one’s handwriting/inscribing (in our paper conditions) might elicit a greater sense of control and might feel more personal than clicking or typing with a digital keyboard (in our tablet conditions), such that handwriting one’s contact information, or inscribing a checkmark on paper (vs. typing on a digital keyboard or clicking a button on the tablet) might increase one’s psychological ownership over that response or decision. Within this perspective, the good-on-paper effect might occur through greater psychological ownership in the paper (vs. tablet) condition.

Study 4 tested this alternative account using a digital device that mimics the unique features of the paper condition more closely: a tablet with a digital pen, which allowed participants to use their own handwriting and inscribe their answers (rather than typing on the digital keyboard). If the good-on-paper effect stems from greater psychological ownership elicited by handwriting/inscribing, then this effect should attenuate when comparing the paper condition to a tablet-with-pen condition. If, however, the good-on-paper effect occurs when
comparing the paper condition to a tablet-with-pen condition, we can then conclude this effect occurs through other processes than psychological ownership. Accordingly, participants made book choices from a list featuring highbrow and lowbrow books either (a) on paper, or (b) on a tablet without pen (these two conditions were similar to our previous studies), or (c) on a tablet with a digital pen. We predicted participants making the decision on paper (vs. tablet with or without pen) would choose more highbrow books. We preregistered the design and analysis plan for this study at: https://aspredicted.org/blind.php?x=ih6up4.

6.1. Methods

6.1.1. Participants. We recruited 310 participants (gender: 188 female, 122 male; age: $M = 20.55, SD = 1.94$) on the campus of a large Chinese university and paid them for their time. As preregistered, we excluded nine participants who did not follow the survey instructions: one who did not make a choice (paper condition) and eight who made fewer choices than instructed (one in the paper condition, three in the tablet-without-pen condition, and four in the tablet-with-pen condition). Our final sample consisted of 301 responses (gender: 184 female, 117 male; age: $M = 20.55, SD = 1.95$).

6.1.2 Pretest of Highbrow and Lowbrow Books. Before the main study, we conducted a multi-step pretest to identify the highbrow and lowbrow books to use in the study and ensure they would not be too familiar to participants in our pool. First, two research assistants selected 20 popular books, which two other research assistants then rated as highbrow or lowbrow based on Read, Loewenstein and Kalyanarama’s (1999) definition of these concepts (see also and Khan and Dhar 2007). Specifically, highbrow books were defined as educationally or culturally enriching and offering long-term benefits, whereas lowbrow books were defined as providing

---

9 Retaining these participants in the analysis did not change the pattern or significance of the results.
short-term benefits and little educational or cultural benefits. Fifteen books were selected based on these initial ratings.

We then recruited 150 students (gender: 75 female, 75 male; age: $M = 21.87$, $SD = 1.62$) from the same population as the main study to further rate the virtuousness of each of these 15 books using the following three items ($\alpha = .83$): (a) To what extent do you believe this book is educational? (1 = not at all, 7 = very much); (b) To what extent do you believe this book is culturally enriching? (1 = not at all; 7 = very much); (c) What type of benefits do you believe this book can provide? (1 = short-term benefits, 7 = long-term benefits). Next, we sought to ensure the books selected for the main experiment would not be too familiar to our participants, which might influence their final choice (e.g., they might not choose a book because they already read it or heard a lot about it). Thus, we also measured the extent to which participants were familiar with each book (1 = very unfamiliar, 7 = very familiar).

The results of a one-sample t-test showed that, out of the 15 books, nine were rated higher than the midpoint (4) of virtuousness (all $M_s > 4.30$ and $ps < .010$), so we categorized these books as highbrow. Four books were rated lower than the midpoint of virtuousness (all $M_s < 3.70$ and $ps < .024$), so we categorized these books as lowbrow. The virtuousness scores of the remaining two books were not significantly different from the midpoint ($M_s = 3.94$ and $4.02$, $ps > .640$), so we could not categorize these books as highbrow or lowbrow. Finally, five of the nine highbrow books were very familiar to participants with familiarity score significantly higher than the midpoint (4) of familiarity (all $M_s > 4.35$ and $ps < .032$). Based on these analyses, we select four highbrow and four lowbrow books for the main experiment (See Appendix B-VI)

### 6.1.3. Design and Procedure

The main study employed a 3-level (decision context: paper and pen vs. tablet without pen vs. tablet with pen) between-subjects design. The study
followed a similar procedure as Study 1a: Research assistants approached passersby on various campus locations and those who agreed to participate received a survey (a) on paper with a pen or (b) on a tablet without a pen (these two conditions were similar to our previous studies), or (c) on a tablet with a pen (iPad pen). This last condition was new and mimicked the paper condition more closely than the tablet-without-pen condition, such that participants could handwrite/inscribe their responses. In the paper [tablets] condition, the survey was titled: “A Short Paper and Pen [Tablet] Survey.” As in previous studies, both surveys used similar formatting and typeface but unlike in Studies 1a and 1b, there was no size-of-character question. Participants completed some basic demographic and filler preference questions similar to the ones used in Study 2 (e.g., table tennis vs. badminton).

Next, participants read about a promotional event from a book seller, during which they could get a good deal by purchasing four books. A list of the eight pretested books was displayed below these instructions featuring four highbrow books and four lowbrow books in two separate columns. Each book was presented with a number, a picture of the book cover, and a brief description (see Appendix B-VI; see also Khan and Dhar 2007). Following Fishbach and Zhang’s (2008) logic to highlight the difference between vices and virtues through separation, the four highbrow books were displayed in the left-hand side (and numbered 1, 3, 5, and 7), whereas the four lowbrow books were displayed in the right-hand side (and numbered 2, 4, 6, and 8). We asked participants to select exactly four books from this list, by marking the box next to the corresponding books, and then writing the book numbers in the space provided below the book list. Twelve participants (five in the paper condition, three in the tablet-without-pen condition, and four in the tablet-with-pen condition) did not mark the boxes next to the book (or marked less than four books) but wrote their four selections in the space provided. We retained
these participants in the analysis and note that the results are unchanged when these participants are excluded. Our measure of virtuous behavior was the number of highbrow books participants selected from the list (ranging from 0 to 4).

6.2. Results

An analysis of variance (ANOVA) with the number of highbrow books participants selected as the dependent variable and the decision context (0 = paper, 1 = tablet without pen, 2 = tablet with pen) as the independent variable revealed a significant effect ($F(2, 298) = 3.57, p = .029; d = .31$). Specifically, in the paper condition, participants chose more highbrow books ($M = 3.05, SD = .76$) than in the tablet-without-pen condition ($M = 2.81, SD = .91, t(298) = 2.11, p = .036; d = .29$), which replicated our previous findings. Furthermore, participants in the paper condition chose more highbrow books than those in the tablet-with-pen condition ($M = 2.77, SD = .77; t(298) = 2.46, p = .014; d = .37$). There were no differences in the number of highbrow books selected when comparing the two tablet conditions ($|t| < 1$; see Figure 1).

**Figure 1.** Number of Highbrow Books Selected as a Function of the Decision Context (Study 4)

![Graph showing the number of highbrow books selected](image)

*Note:* Error bars represent standard error

6.3. Discussion

This study provided additional support for the good-on-paper hypothesis (H1) and fulfilled two important objectives. First, the study extended our findings to a different type of
device interaction, showing that the good-on-paper effect occurred when comparing paper to both a tablet without pen (i.e., touch) and a tablet with pen—the latter of which closely mimicked handwriting on paper. These results suggested psychological ownership—possibly elicited by the act of handwriting—is unlikely to account for the good-on-paper effect. Second, the study extended our findings to the domain of book choices, showing that the good-on-paper effect goes beyond prosocial behavior. We note that, in a similar vein, Supplemental Study 2 (Appendix A-II) further extended our finding to the context of academic performance and using a computer as the digital device: Participants working on a task described as diagnostic of skill and intelligence performed better when completing this task on paper (vs. computer).

Taken together, these results show the good-on-paper effect occurs when comparing paper to various types of digital devices and highlight that this effect does not depend on one’s ability to interact with the decision context through handwriting or touch-interfaces. Instead, we propose this effect occurs due to greater perceptions of realness and hence self-diagnosticity for decisions on paper (vs. digital device) and we test these mechanisms in the next two studies. Furthermore, the next study sought to rule out some additional alternative explanations stemming from the many differences between a sheet of paper and a digital device such as a tablet.

**7. Study 5: Serial Mediation**

In this study, we explored our proposed underlying mechanism by testing the sequential mediating roles of perceptions of realness and self-diagnosticity on the good-on-paper effect (H2). After making a series of choices on paper or on a digital device (tablet), participants indicated their perceptions of the realness and self-diagnosticity of choices in this context. They then read a charitable appeal and indicated their willingness to help. We predicted participants making the decision on paper (vs. tablet) would be more willing to help and that this effect
would be sequentially mediated by perceptions of realness and self-diagnosticity. To explore some additional alternative processes, we also measured several points of difference between a sheet of paper and a tablet (e.g., technology, weight, access to games). We preregistered the design and analysis plan for this study at: https://aspredicted.org/blind.php?x=ce7fk9.

7.1. Methods

7.1.1. Participants. We recruited 249 students (gender: 138 female, 110 male, 1 undisclosed; age: $M = 21.52$, $SD = 2.72$) on the campus of a large Chinese university to complete this survey and receive candy bars as tokens of appreciation. We did not exclude any participant.

7.1.2. Design and Procedure. The study employed a 2-level (decision context: paper vs. digital device) between-subjects design and followed a procedure similar to that of Study 1a, with research assistants randomly approaching participants and asking them to complete “A Short Paper and Pen [Tablet] Survey.” Participants provided some basic demographic information and completed filler preference questions similar to the ones used in Study 2 (e.g., coffee vs. tea). We then assessed their perceptions of the self-diagnosticity of choices in the survey context, using four items ($\alpha = .84$; e.g., “Right now, what I do in this survey says a lot about who I am;” see Appendix B-VII) adapted from Touré-Tillery and Light (2018). Next, we measured participants’ perceptions of the realness of these choices (two items, $r = .80$, $p < .001$; e.g., “To what extent do your choices in this survey seem real right now?” (see Appendix B-VII).

On the next page/screen, participants read a brief charitable appeal soliciting volunteering time for a student-led charitable organization supporting children with special needs (see Appendix B-VIII). Participants could indicate their interest in volunteering for the organization by leaving their WeChat account numbers so the organization could contact them directly and
put them on the schedule for a specific volunteering task. Our measure of virtuous behavior (DV) was whether participants provided their WeChat account numbers (i.e., sign-up rate).

Because there are many natural differences between a sheet of paper and a tablet that could account—at least partially—for the results we have documented so far, we included a comprehensive set of questions to measure these differences. Specifically, in the paper [tablet] condition, participant indicated on a series of bipolar scales whether they would describe the sheet of paper [tablet] they were using as: (a) 1 = very low tech, 7 = very high tech; (b) 1 = very light, 7 = very heavy; (c) 1 = not at all sensitive to touch, 7 = very sensitive to touch; (d) 1 = very fragile, 7 = very sturdy; and (e) 1 = easy to put in an envelope, 7 = hard to put in an envelope. They also indicated the extent to which they associated the sheet of paper [tablet] they were using with (f) playing games (1 = not at all; 7 = very much), and (g) social media networking (1 = not at all; 7 = very much).

7.2. Results

7.2.1. Sign-up Rate. A logistic regression of sign-up rate$^{10}$ (0 = did not provide WeChat account info, 1 = provided WeChat account info) on decision context (0 = paper, 1 = digital device) showed a greater proportion of participants signed-up to volunteer when this decision was made on paper (29.75%) than on the tablet (5.83%; b = -1.92(.44), z = - 4.40, p < .001; odds ratio = .15).

7.2.2. Perceived Realness and Self-Diagnosticity. In addition, analyses showed participants perceived their choices as more real on paper ($M = 6.05$, $SD = .98$) than on a tablet ($M = 5.73$, $SD = .96$; $t(247) = 2.55$, $p = .011$; $d = .33$), and as more self-diagnostic on paper ($M = 5.21$, $SD = 1.36$) than on a tablet ($M = 4.84$, $SD = 1.22$; $t(247) = 2.20$, $p = .029$; $d = .29$).

---

$^{10}$ Eight participants (all in the paper condition) did not complete the main dependent variable (sign-up rate), which appeared on the second page/screen.
7.2.3. Mediation Analysis. To test whether the effect of decision context on virtuous behavior (i.e., sign-up rate) occurred due to differential perceptions of realness ($M = 5.90, SD = .98$) and hence self-diagnosticity ($M = 5.03, SD = 1.31$), we conducted a serial mediation analysis using the bootstrap test of the indirect effect (Process Model 6; Hayes 2017). We found a significant mean indirect effect ($a_1 \times d \times b_2 = -.08(.05)$, 95% CI [-.202, -.004]), indicating perceived realness and self-diagnosticity sequentially mediated the effect of decision context on virtuous behavior (see Figure 2).

**Figure 2.** Serial Mediation by Perceived Realness and Self-Diagnosticity (Study 5)

Note. ***$p < .001$, *$p < .05$

7.2.4. Alternative Mechanisms. Compared to the tablet, participants perceived the sheet of paper as heavier ($M_{\text{paper}} = 3.08, SD_{\text{paper}} = 1.96; M_{\text{tablet}} = 4.22, SD_{\text{tablet}} = 1.58; t(236)^{11} = -4.92, p < .001; d = -.64$), less sensitive to touch ($M_{\text{paper}} = 3.71, SD_{\text{paper}} = 1.64; M_{\text{tablet}} = 5.23, SD_{\text{tablet}} = 1.49; t(236)^{11} = -7.44, p < .001; d = -.97$), less sturdy ($M_{\text{paper}} = 4.33, SD_{\text{paper}} = 1.65; M_{\text{tablet}} = 5.41, SD_{\text{tablet}} = 1.43; t(237)^{12} = -5.42, p < .001; d = -.70$), easier to store (in an envelope; $M_{\text{paper}} = 3.42, SD_{\text{paper}} = 2.13; M_{\text{tablet}} = 4.83; SD_{\text{tablet}} = 1.77; t(237)^{12} = -5.59, p < .001; d = -.72$), and less

---

11 Eleven participants (all in the paper condition) did not respond to this question.  
12 Ten participants (paper condition) did not respond to this question.
associated with games ($M_{paper} = 2.56, SD_{paper} = 1.49; M_{tablet} = 3.90, SD_{tablet} = 2.10\); $t(238^{13}) = -5.70, p < .001; d = -0.74$) and social networking ($M_{paper} = 3.64, SD_{paper} = 1.76; M_{tablet} = 4.66, SD_{tablet} = 1.70; t(239^{14}) = -4.56, p < .001; d = -.59$). However, there was no difference between the sheet of paper and tablet in terms of high-tech perceptions ($M_{paper} = 3.95, SD_{paper} = 1.43; M_{tablet} = 4.08, SD_{tablet} = 1.24; t(235^{15}) = -.77, p = .440; d = -.10$). Notably, bootstrap tests of the indirect effect (Process Model 4; Hayes 2017) showed that none of the differences observed here mediated the good-on-paper effect (all 95% CI included zero). Thus, while some of the alternative mediators were correlated to the proposed mediators (realness and self-diagnosticity; see Appendix C-II), they do not explain the causal relationship between decision context and virtuous behavior.

7.3. Discussion

This study supported our hypothesis about the sequential mediating roles of perceptions of realness and self-diagnosticity in the good-on-paper effect (H2). We found that participants completing a survey on paper (vs. tablet) perceived their choices as more real and hence as more self-diagnostic, which in turn increased their willingness to help.

In the next study, we conducted another test of these underlying processes using an experimental-causal-chain design, in which we manipulated the independent variable (paper vs. digital device) and one of the mediating variables (perceived realness) to draw causal inferences about the chain of events. This design allowed us to minimize interferences between measures (e.g., whereby measures are correlated simply because they are in the same survey). Indeed, purely statistical mediation analyses based on regression models have often been criticized as

---

13 Nine participants (paper condition) did not respond to this question.
14 Eight participants (paper condition) did not respond to this question.
15 Twelve participants (paper condition) did not respond to this question.
relying heavily on such correlations (Judd and Kenny 1981). Therefore, a series of experiments that demonstrates the proposed causal chain (i.e., an experimental-causal-chain design) is generally considered a more conservative approach (see, Spencer, Zanna, and Fong 2005).

8. **Study 6: Causal-Chain Mediation**

Using an experimental-causal-chain design, we examined the sequential mediating roles of perceptions of realness and self-diagnosticity on the good-on-paper effect (H2), testing these proposed psychological processes first as effects of the decision context (Study 6a) and then as predictors of virtuous behavior (Study 6b). Specifically, in Study 6a, participants made a series of choices as part of a survey on paper or on a digital device (tablet) and we measured their perceptions of the realness and self-diagnosticity of these choices. We predicted participants in the paper (vs. tablet) condition would perceive their choices as more self-diagnostic and that perceptions of realness would mediate this effect. We preregistered the design and analysis plan for this study at: [http://aspredicted.org/blind.php?x=rs9ed4](http://aspredicted.org/blind.php?x=rs9ed4). In study 6b, participants completed an experimental manipulation of realness, made a series of choices, and then indicated their perceptions of the self-diagnosticity of these choices and their willingness to help a charitable organization. We predicted a greater incidence of virtuous/prosocial behavior in the high (vs. low) perceived realness condition and a mediation of this effect by perceptions self-diagnosticity.

8.1. **Study 6a – Perceived Realness Mediates the Effect of the Decision context on Perceived Self-Diagnosticity**

8.1.1. **Method.** We recruited 172 students (gender: 108 female, 64 male; age: $M = 20.93, SD = 2.31$) on the campus of a large Chinese university and gave them candy bars in appreciation for their time. We were concerned that the length of the survey would lead to a larger incidence of low-quality responses, which might need to be excluded from the analysis. Thus, we included
an attention check for data-quality control, and we collected more participants than our target sample size to allow for these exclusions. As preregistered, we excluded\(^\text{16}\) 23 participants who did not follow our instructions and either failed (n = 21) or skipped (n = 2) this attention-check question. These exclusions left us with a final sample of 149 responses for subsequent analyses (gender: 93 females, 56 males; age: \(M_{\text{age}} = 20.95, SD_{\text{age}} = 2.30\)).

The study employed a 2-level (decision context: paper vs. digital device) between-subjects design and followed a procedure similar to that of Study 1a, with research assistants randomly approaching participants to ask them to complete “A Short Paper and Pen [Tablet] Survey.” In this study, to further extend our investigation to other types of device interactions, we also gave a stylus to participants in the tablet (iPad) condition. We note that the stylus in this study acted more like an extended finger than like the digital pen in Study 4. Specifically, participants could not handwrite with the stylus, but rather could click on their selected responses. After indicating their gender and age, participants completed a “preference survey,” in which they made choices from nine neutral pairs of options (e.g., water vs. tea; see Appendix B-IX). We then assessed their perceptions of the self-diagnosticity of these choices, using four items (\(\alpha = .82\); e.g., “Right now, what I do says a lot about who I am;” see Appendix B-IX) adapted from Touré-Tillery and Light (2018). An attention check instructing participants to answer “3” on a 7-point scale (see Oppenheimer et al. 2009) followed these questions. On the next page/screen, we measured participants’ perceptions of the realness of these choices (three items, \(\alpha = .87\); e.g., “To what extent do your choices seem real right now?” see Appendix B-IX).

\(^{16}\) When we retain these participants in the analyses, the significance of the effect on perceived self-diagnosticity in unchanged, but the effect on realness becomes non-significant (\(t(169) = 1.48, p = .140, d = .22\)).
In addition to the differences in digital fluency we explored in Study 1b, people might believe information will last longer on a digital device than on paper (i.e., perceived digital longevity) or vice versa. These beliefs might in turn interact with the decision context to influence judgments and choices. To test these possibilities, before thanking and debriefing participants, we assessed digital fluency (using the same three items as in Study 1b; $\alpha = .79$) and perceived digital longevity (one item: “Do you believe information can be kept longer on paper or on electronic devices?” (1 = definitely longer on paper, 7 = definitely longer on electronic device; see also Tully, Hershfield, and Meyvis 2015). We found no evidence that the good-on-paper effect depends on beliefs about digital longevity or digital fluency for the judgments measured in this study (see Appendix C-III), for the prosocial decision Study 1b, or for task performance in Supplemental Study 2.

### 8.1.2. Results

Analyses showed participants perceived their choices as more real on paper ($M = 5.50, SD = 1.14$) than on the tablet ($M = 5.11, SD = 1.21$; $t(146)^{17} = 2.00, p = .047; d = .33$), and as more self-diagnostic on paper ($M = 5.35, SD = 1.09$) than on the tablet ($M = 4.90, SD = .98$; $t(147) = 2.61, p = .010; d = .43$).

Moreover, perceived realness mediated the effect of the decision context ($0 = \text{paper}, 1 = \text{digital device}$) on perceived self-diagnosticity: a bootstrap test of the indirect effect (Hayes 2017) using 5,000 replications revealed a significant mean indirect effect ($a \times b = -.11 (.07)$, 95% CI $[-.26, -.002]$; see Figure 3).

**Figure 3.** Perceived Realness Mediates the Effect of Decision Context on Perceived Self-Diagnosticity (Study 6a)

---

17 One participant (paper condition) did not respond to the realness questions.
8.2. Study 6b – Perceived Self-Diagnosticity Mediates the Effect of Perceived Realness on Virtuous Behavior

8.2.1. Methods. We recruited 207 students (gender: 164 female, 43 male; age: $M = 22.06, SD = 3.80$) from several Chinese universities in a large city to complete this study online in exchange for monetary compensation. We excluded seven participants who did not complete the manipulation of realness (i.e., they did not write the essay as instructed; three in the high realness condition and four in the low realness condition).

The study employed a 2-level (perceived realness: low vs. high) between-subjects design and started with an experimental manipulation of perceived realness. Participants first read: “Human beings have an incredible ability to imagine and create alternative realities. Sometimes the lines between fantasy and reality can be blurry: a simple matter of perspective. In this section, we are interested in how this perspective can shift from one moment to the next.” Then, in the high [low] perceived realness condition, participants were instructed to: “Please take a minute to consider everything that might make the present moment feel real [imaginary] as opposed to imaginary [real], and tangible [intangible] as opposed to intangible [tangible]. In the space provided below, please describe how your actions and the things around you at this

Note. ***$p < .001$, *$p < .05$, †$p < .10$
moment could be real or tangible.” This experimental manipulation was meant to bring forth a mindset that things are (vs. are not always) real or tangible, which would subsequently influence behavior. The task was followed by a (filler) preference survey similar to the one used in Study 1a (five choices, e.g., water vs. tea).

Next, participants read a charitable appeal soliciting volunteering time for a local orphanage (see Appendix B-X). To measure virtuous behavior, we asked participants to indicate their interest in volunteering for the organization by leaving their cellphone numbers so that the organization could contact them directly (i.e., sign-up rate). Then, we assessed their perceptions of the self-diagnosticity of their choices in the survey \( r = .70, p < .001 \), and a manipulation check of perceived realness \( r = .63, p < .001 \), using the first two items from Study 6a for each measure (see Appendix B-IX). The study ended with a basic demographic questionnaire.

**8.2.2. Results.** In response to the experimental manipulation of realness, participants wrote about a variety of topics. For example, in the high realness condition, one participant wrote “(…) I am going to have breakfast now, which is specific, and breakfast can be touched.” Another wrote: “I am going to the examination room to take the third exam. This is the behavior that makes me feel real and specific at this moment.” By contrast, in the low realness condition, one participant wrote: “Imagination is intangible and abstract, and sometimes it can be felt, and sometimes it feels non-existent. There is always a feeling of being invisible and invisible to people (…)” An analysis of the manipulation check showed participants in the high realness condition perceived their actions and choices in the context of the survey as more real/tangible \( (M = 5.47, SD = 1.15) \) than participants in the low realness condition \( (M = 5.02, SD = 1.25) \); \( t(198) = -2.66, p = .009; d = -.37 \)—indicating the manipulation produced the intended effect.
Additionally, two independent coders blind to our hypothesis rated the extent to which participants’ responses to the experimental-manipulation question related to each of the four dimensions of psychological distance (see Appendix C-IV for coding instructions). All participants wrote about (un)realness, and some also mentioned other dimensions of psychological distance: 49.00% wrote about spatial distance, 46.50% about temporal distance, and 20.50% about social distance. The coders’ ratings of realness (ICC = .82) showed that participants in the high-realness condition wrote about realer, more tangible things ($M = 1.48$, $SD = .83$) than participants in the low-realness condition ($M = 3.38$, $SD = 1.29$; $t(198) = 12.60$, $p < .001$; $d = 1.75$). On average participants in the high (vs. low) realness condition also wrote about things that were closer spatially, temporally, and socially (see Appendix C-IV for summary of results)—although, as noted above, less than 50% of participants mentioned these three other dimensions. These results are consistent with the notion that the dimensions of psychological distance are interconnected but distinct (Trope and Liberman 2010).

We then examined the effect of perceived realness ($0 = low$, 1 = high) on sign-up rate (0 = did not provide cellphone number, 1 = provided cellphone number) using a logistic regression. We found more participants signed up in the high realness condition (22.73%) than in the low realness condition (10.00%; $b = .97(.42)$, $z = 2.33$, $p = .020$; odds ratio = 2.65). Furthermore, participants in the high realness condition ($M = 5.21$, $SD = 1.33$) perceived their choices as more self-diagnostic than those in the low realness condition ($M = 4.77$, $SD = 1.51$; $t(198) = -2.18$, $p = .031$; $d = -.31$).

Finally, a bootstrap test of the indirect effect (5,000 replications) revealed a significant mean indirect effect ($a \times b = .25 (.16)$, 95% CI [.03, .62]; see Figure 4), indicating perceived self-diagnosticity mediated the relationship between perceived realness and sign-up rate.
Figure 4. Perceived Self-Diagnosticity Mediates the Effect of Perceived Realness on Sign-up Rate (Study 6b)

Perceived Realness (Low vs. High) \( \rightarrow \) Perceived Self-diagnosticity \( \rightarrow \) Sign-up Rate

\[ a = .44^* \]
\[ b = .57^{***} \]
\[ c = .97^* \]
\[ c' = .84^† \]

Note. \( *** p < .001, ^* p < .05, ^† p < .10 \)

8.3. Discussion

We found perceived realness mediated the effect of the decision context on perceived self-diagnosticity (Study 6a), which in turn mediated the effect of perceived realness on virtuous behavior (Study 6b). Through this experimental causal-chain design, we provided additional evidence for the sequential mediating roles of perceptions of realness and self-diagnosticity on the good-on-paper effect (H2). In Supplemental Study 3 (Appendix A-III), we provide an additional demonstration of the effect of the decision context on perceptions of self-diagnosticity by showing that feedback about actions performed on paper (vs. a digital device) have a greater influence on people’s self-concept. The next study continued to explore the underlying role of self-diagnosticity through moderation.

9. Study 7: Moderation by Chronic Self-Diagnosticity

This study tested the moderating role of perceptions of self-diagnosticity on the good-on-paper effect (H3). Participants completed a survey on paper or on a digital device (tablet), in which they read a petition for an environmental protection campaign asking them to express their support by signing their names and providing their contact information. We measured perceptions of self-diagnosticity as an individual difference (see Touré-Tillery and Light 2018).
We predicted that chronic self-diagnosticity would moderate the effect of decision context on virtuous behavior such that it would attenuate among people who are higher in chronic self-diagnosticity. We preregistered the design and analysis plan for this study at: http://aspredicted.org/blind.php?x=u3cj39.

9.1. Methods

9.1.1. Participants. We recruited 315 students (gender: 132 female, 181 male, 2 undisclosed; age: $M = 21.29$, $SD = 2.75$) on the campus of a large Chinese university and gave them candy bars in appreciation for their time. We did not exclude any participant.

9.1.2. Design and Procedure. The study employed a 2 (decision context: paper vs. digital device) $\times$ self-diagnosticity (continuous) between-subjects design, with decision context manipulated and self-diagnosticity measured. Research assistants approached passersby, and those who agreed to participate received a survey on paper or tablet (iPad). The survey started with demographic questions (gender, age) and the same filler survey about preferences as in Study 6b.

Participants then read a message about an environmental protection campaign (see Appendix B-XI) and indicated their support for the campaign by signing their names and leaving their cellphone numbers. Our measure of virtuous behavior consisted of the number of participants who provided both pieces of information (i.e., support rate). On the next page of the survey, we assessed participants’ dispositional levels of self-diagnosticity using Touré-Tillery and Light’s (2018) seven-item self-diagnosticity scale ($\alpha = .88$). The scale captures individual differences in people’s propensities to see their own actions as representative of who they are (e.g., “What I do is a reflection of who I am” 1 = strongly disagree, 7 = strongly agree) by computing a self-diagnosticity score (SDS) corresponding to the average of the seven items.
9.2. Results

First, a logistic regression of support rate (0 = did not support the cause; 1 = supported to the cause) on decision context (0 = paper, 1 = digital device) replicated the good-on-paper effect: participants in the paper condition exhibited a higher support rate (51.90%) than participants in the tablet condition (36.13%; b = -.65(.23), z = -2.80, p = .005; odds ratio = .52).

Next, we examined the moderating role of self-diagnosticity. A logistic regression of support rate on decision context (0 = paper, 1 = digital device), SDS (mean-centered), and their interaction showed a significant effect of decision context (b = -.58(.24), z = -2.41, p = .016; odds ratio = .036), but no effect of SDS (b = .12(.19), z = .66, p = .511; odds ratio = 1.13).

Importantly, the predicted interaction of decision context × SDS emerged (b = .52(.26), z = 1.98, p = .048; odds ratio = 1.67): At and below +.21 standard deviations of the mean of SDS (i.e., a value of 5.54 on the 7-point scale; 57.74% of the sample), support rate was higher on paper (vs. digital device; all ps ≤ .05). However, above this level of SDS (42.26% of the sample), there was no effect of decision context on support rate (all ps > .05; see Figure 5). We note that our sample included mostly people with moderate and high levels of self-diagnosticity. Indeed, only 9% of our sample reported chronic SDS below the midpoint of the 7-point scale (i.e., 4). Furthermore, an examination of the quantiles of the SDS measure showed that the median (5.29) and the first quantile (4.71) were all much higher than the mid-point of the scale. These features of the SDS distribution—i.e., the fact that few participants in our sample were truly low in SDS—can explain why we did not observe an attenuation of the effect at the lower levels of SDS in this study. Nonetheless, our results are consistent with our hypothesis, indicating that the good-on-paper effect occurs at moderate levels of self-diagnosticity—where perceptions of diagnosticity might be more malleable—but attenuates at the observed higher levels of self-diagnosticity.
Finally, we examined the effect of self-diagnosticity on environmental support in each experimental condition. In the tablet condition, we found that participants’ support for the cause increased as their levels of SDS increased ($b = .64(.18)$, $z = 3.50$, $p < .001$; odds ratio = 1.89). However, there was no effect of SDS in the paper condition ($b = .12(.19)$, $z = .66$, $p = .511$; odds ratio = 1.13): participants moderate in SDS were as supportive of the cause as those high in SDS. This last set of results further highlighted the important underlying role of self-diagnosticity in the good-on-paper effect.

**Figure 5.** Chronic Self-Diagnosticity Moderates the Effect of Decision Context on Support for Environmental Protection (Study 7)

9.3. Discussion

This study replicated the good-on-paper effect (H1) and demonstrated the moderating role of chronic perceptions of self-diagnosticity on this effect (H3), showing the effect attenuates at higher levels of self-diagnosticity. Because this study did not allow us to capture the attenuation of the good-on-paper effect at low levels of self-diagnosticity, in the next study, we conducted another test of the moderating role of perceptions of self-diagnosticity on the good-
on-paper effect with a focus on showing that this effect attenuates at low levels of self-diacasticity. This final study also sought to extend the good-on-paper effect to the context of (healthy vs. indulgent) food choices.

10. Study 8: Moderation by Target of Choice

Study 8 examined the underlying role of perceptions of self-diacasticity on the good-on-paper effect by testing the moderating role of whether a decision is being made for the self or for another person (H4). Participants chose between healthy and indulgent food options on paper or on a digital device (tablet), either for themselves or for another person. Because choices made for another person should be less representative of the choosers’ own preferences and personal characteristics (i.e., low in self-diacasticity) than choices made for the self, we expected the good-on-paper effect to replicate when participants chose for themselves, but not when they chose for another person.

10.1. Methods

10.1.1 Participants. Three hundred forty-six (346) adults (gender: 176 female, 167 males, 3 undisclosed; age: \( M = 42.35, SD = 16.52 \)) recruited in the downtown area of a large city in the United States completed this study and received gum packets in appreciation for their time. We excluded three participants who did not make the critical food choice and two participants who chose both a healthy and an indulgent entrée, leaving 341 participants (174 female, 164 male, 3 undisclosed; age: \( M = 42.51, SD = 16.53 \)) for subsequent analyses.

10.1.2. Design and Procedure. The study employed a 2 (decision context: paper vs. digital device) \( \times \) 2 (target of choice: self vs. other) between-subjects design and followed a procedure similar to that of Study 1a. Research assistants approached passersby to ask them to take part in a short study and handed them either a paper survey (with a pen) or a tablet
displaying a digital survey. The survey started with the same short set of questions as in Study 1a. Next, participants in the choice-for-self [choice-for-other] condition read: “In this survey, we are interested in people’s food preferences [food orders for others]. Please take a moment to review the menu below and circle the entrée you would be most likely to order for yourself for your next meal [select the entrée you would be most likely to choose for a friend for his or her next meal].” The menu, adopted from Fishbach and Zhang (2008), appeared below these instructions and featured 10 entrées in two separate columns, following Fishbach and Zhang’s (2008) logic to highlight the difference between vices and virtues. The five entrées displayed in the left column were healthy, whereas the five entrées displayed in the right column were indulgent (see Appendix B-XII). Participants’ choice of entrée constituted our measure of virtuous behavior. A final question asked participants whether the target of the choice had any dietary restrictions or allergies (No/Yes, please specify).

10.2. Results

A logistic regression of food choice (0 = indulgent, 1 = healthy) on decision medium (0 = paper, 1 = digital device), the target of the choice (0 = other, 1 = self) and their interaction showed a main effect of the target of the choice, such that participants were more likely to choose a healthy entrée for themselves (55.56%) than for another person (46.47%; b = .88(.31), z = 2.82, p = .005; odds ratio = 2.41). There was no effect of the decision context (b = -.45(.31), z = -1.45, p = .147; odds ratio = .64). Importantly, the predicted interaction of the decision context × the target of the choice emerged (b = -1.08(.45) z = -2.37, p = .018; odds ratio = .34).

Specifically, when participants chose for themselves, a greater proportion made a healthy food choice on paper (72.04%) than on the tablet (35.90%, b = -1.52(.33), z = - 4.62, p < .001; odds ratio = .22). However, when participants chose for another person, there was no significant
difference in food choices between the paper (51.65%) and tablet conditions (40.51%, \( b = -0.45(0.31), z = -1.45, p = .147; \) odds ratio = .64; see Figure 6).

**Figure 6.** Moderation by Target of Choice (Study 8)

![Bar chart showing the percentage of healthy entrée choices for self and other in paper and tablet conditions.](image)

*Note:* Error bars represent the standard error.

10.3. Discussion

Study 8 extended our findings to the context of food choices and provided further evidence for the role of perceptions of self-diagnosticity on the good-on-paper effect by demonstrating the moderating role of choosing for oneself rather than for another person (an act low in self-diagnosticity; \( H4 \)).

11. General Discussion

The present research explored the effect of the decision context (paper vs. digital device) on perceptions of realness and self-diagnosticity, and hence on virtuous behavior. We first showed consumers are more likely to behave virtuously when they make choices on paper (vs. a digital device) across various contexts, including prosocial choices (Studies 1, 2, 3, 5, 6b, and 7), book choices (Study 4) and food choices (Study 8). We labelled this pattern of behavior the good-on-paper effect and showed that it occurs whether decisions are objectively real (e.g., Studies 1, 2, 3, 5, 6b, and 7) or hypothetical (e.g., Studies 4 and 8). We then showed that
consumers perceive decisions made on paper (vs. a digital device) as more real and hence more self-diagnostic, which in turn increases virtuous behavior (sequential mediation: Study 5, and Studies 6a and 6b). As further evidence for the role of perceptions of self-diagnosticity on the good-on-paper effect, we showed chronic perceptions of self-diagnosticity (Study 7) and the target of the choice (self vs. other; Study 8) moderate this effect, such that the effect attenuates at high levels of chronic self-diagnosticity and for choices made for another person (i.e., choices low in self-diagnosticity). These findings have important theoretical and practical implications.

11.1. Theoretical Implications and Future Research

Recognizing the importance of the medium through which actions are performed, prior research has extensively explored the effects of using computers compared to using paper on performances related to reading and text processing (e.g., Mangen, Walgermo, and Brønnick 2014), learning, and test-taking (e.g., Clariana and Wallace 2002; DeAngelis 2000; Mazzeo et al. 1991; Watson 2001). More recently, neuromarketing studies have examined the relative effectiveness of advertising via print and digital media in terms of consumers’ neurophysiological responses (Dimoka et al. 2015; Millward Brown Study 2009; United States Postal Service Report 2015). The present research extends these lines of inquiries by investigating the differential effects on virtuous behavior of using paper versus digital devices (e.g., tablet with and without pen, computer). Whereas much of the prior work focused on the neurophysiological processing of information viewed on these different media (Magen, Walgermo, and Brønnick 2013), our research explores how people’s behaviors change when they make decisions on paper versus on digital devices, due to differential perceptions of realness and self-diagnosticity. In addition to broadening the scope of knowledge on the effects of different decision contexts/media, our findings contribute to the literatures on virtuous behavior,
motivation, and choice by uncovering a factor that influences judgments and behaviors through self-concept management.

Furthermore, this article is the first to document the link between the decision context (paper vs. digital device) and perceptions of realness and self-diagnosticity. Notably, the link between perceived realness and self-diagnosticity suggests a potential connection between other dimensions of psychological distance (i.e., social, spatial, and temporal; Trope and Liberman 2010) and perceiving choices are representative of the self. The implication is that people might perceive actions they will perform in the distant future (or actions they have performed in the distant past) as less self-diagnostic than temporally closer actions. Similarly, people might perceive choices related to physically distant (vs. close) places to be less self-diagnostic. We find that increasing realness (i.e., reducing this dimension of psychological distance) increases prosocial behavior through differential perceptions of self diagnosticity. As noted in the introduction, a construal level theory (CLT) perspective might have made a different prediction. According to CLT, morals and values are more likely to guide decisions and intentions for psychologically distant (vs. proximal) situations (Eyal, Liberman, and Trope 2008). Thus, virtuous behaviors should increase as psychological distance increases. For example, previous studies suggest that increasing temporal distance increases virtuous intentions, such that people are more willing to commit to donate blood in the distant (vs. near) future (Choi, Park, and Oh 2012) and make healthier food choices for future (vs. immediate) consumption (Read and Van Leeuwen 1998). However, this prediction does not hold for all dimensions of psychological distance: several studies show that increasing physical distance (Touré-Tillery and Fishbach 2017) or social distance decrease prosocial intentions (Krebs 1975; Small 2011). For example, Levine et al. (2002) find that bystanders are less likely to help victims who are described as out-
group (higher social distance) as opposed to in-group members (lower social distance). These seemingly contradictory findings indicate that the four dimensions of psychological distance are distinct and thus influence virtuous behavior in different ways. Furthermore, the effect of psychological distance on virtuous behavior often depends on other factors, such as victim identification (Kogut et al. 2018), choice target (self vs. other; Mehta, Zhu and Meyers-Levy 2014), and goals (Xu, Rodas, and Torelli 2020). Future research is needed to further resolve discrepancies in the effects of various dimensions of psychological distance and explore their influence on perceptions of self-diagnosticity and virtuous behavior.

In most of our experiments, we operationalized digital devices through tablets (i.e., touch interface; iPads) for several reasons. First, the use of tablets allowed us to conduct our experiments in the most naturalistic manner possible. Indeed, tablets are now common devices, which consumers encounter in a variety of settings (doctor’s office, restaurants, charity auction), and hence fit naturally in the contexts we examine. Second, the use of tablets allowed us to minimize the differences in participants’ experiences while completing our surveys. Unlike personal computers or laptops, the shape and portability of a tablet closely mirrors that of a sheet of paper, such that we could put participants in similar situations (e.g., completing a survey while standing on the sidewalk) whether they were in the paper or tablet condition.

Notably, in three studies, we show that the good-on-paper effect is robust across different types of digital devices and device interactions: tablet with a digital pen (Study 4) or stylus (Study 6a), and desktop computer (Supplemental Studies 2 and 3 in Appendix A-II and -III). These findings suggest that the good-on-paper effect and its antecedents (perceived realness and self-diagnosticity) do not stem from differences related to touch interfaces, or the ability to handwrite/inscribe responses. Instead, we posit that the differential perceptions of realness (and
resulting self-diagnosticity and virtuous behavior) on paper (vs. digital devices) occurs because paper (vs. digital) contexts feels more tangible, belonging to the physical world, rather than the virtual world. We note other factors might be at play. For example, the differential perceptions of realness for paper (vs. digital devices) might, in part, stem from the knowledge that paper existed long before digital devices, such that over time, as digital devices become even more entrenched in society, people might come to perceive them to be just as real as paper.

Finally, the notion that people perceive actions on digital device (vs. paper) as less real could have far-reaching implications in other contexts and for other forms of judgments and behaviors—beyond virtuous behaviors. For example, would people signing legal documents on a tablet (e.g., via docusign) versus on paper perceive their actions as less legally binding with implications for how carefully they review these documents and subsequently abide to their terms? Would consumers making financial decisions take more risks on digital devices (vs. paper), with implications for investment and retirement portfolios? More research is needed to delve deeper into the antecedents and consequences of perceptions of realness.

11.2. Practical Implications

The present research as important practical implications for marketers, policymakers, and anyone seeking to encourage prosocial acts (e.g., charitable, political, or social-justice organizations) or other forms of virtuous behaviors (e.g., healthy eating, learning). For example, restaurants might consider opting for paper menus rather than digital ones to encourage healthier food choices, whereas parents and educators might provide students with paper (vs. digital) book order forms to encourage the selection of more educational reading materials. Similarly, to increase pledges of money and time, charitable organizations might consider sticking to paper pledge forms and sign-up forms. Of course, given the significant negative influence of paper
production on the environment (see Ince, Cetecioglu, and Ince 2011), the use of recycled paper, as well as the re-use and proper disposal of paper products should become standard practice. Finally, and importantly, our theory and findings suggest that beyond soliciting virtuous behaviors on paper (vs. digital devices), which carries obvious environmental consequences, interventions that heighten perceptions of realness or self-diagnosticity could also promote virtuous behavior in both paper and digital contexts.


Millward Brown Study (2009) Using neuroscience to understand the role of direct mail. Accessed July 11, 2021, 
https://static1.squarespace.com/static/58ee4bac414fb53d228c3532/t/5d30cff8e172f9000021e612/1563480057602/MillwardBrown_CaseStudy_Neuroscience.pdf.


Statista Research Department (2020) *Forecast number of tablet users worldwide 2013-2021*.


