Understanding Consumer Product Decisions When Shopping By Voice

Zhen Yang, Drexel University, USA
Yanliu Huang, Drexel University, USA
Barbara Kahn, University of Pennsylvania, USA
Jiangen He, Drexel University, USA

Using an online food ordering website, we show that consumers are more likely to choose recommended products and hedonic recommended products when shopping by voice than by other modalities (i.e., clicking or typing). Oral communication increases perceived socialness, which induces compliance behavior, and ultimately leads to recommendation acceptance.

[to cite]:

[url]:
http://www.acrwebsite.org/volumes/2551803/volumes/v47/NA-47

[copyright notice]:
This work is copyrighted by The Association for Consumer Research. For permission to copy or use this work in whole or in part, please contact the Copyright Clearance Center at http://www.copyright.com/.
The third and fourth papers examine the role of human-computer interaction in shaping brand perception and product choice. Chung and Jun investigate how AR-integration of digital products into a consumer’s physical environment affects consumer experience. They find that AR enhances self-brand connection for low-involved consumers. They further provide evidence that this improved self-brand connection is driven by a heightened consumer perception of the brand as communal rather than exploitative. Finally, Yang and colleagues examine how shopping by voice differs from the online shopping that is navigated by clicking or typing modalities. They find that that shopping by voice leads consumers to select more recommended products, compared to other modalities. They also identify the underlying mechanism for this discrepancy, which is voice shopping increases perceived socialness, which induces compliance behavior.

Our session contributes to the conference by documenting a range of unexplored expression modalities in research on consumer-technology interactions, a burgeoning area of research drawing on theory and methods from computer science, psychology, and design. We believe the session will spur the consumer researchers who are interested in innovations, consumer- technology interactions, and experiential consumption to seize the opportunity to further contribute foundational research in this growing and impactful research area.

**Horizontal Touch Interfaces and the Illusion of Discrete Content**

As computers have shifted from desktop and laptop computers to tablets and smartphone, interfaces have likewise shifted to direct-touch modalities where consumers directly interact with the content under their fingers. This has created a sea-change in the organization of e-commerce information (Wang, Malthouse, & Krishnamurthi 2015), as sites adapt to differing screen sizes and the lack of a mouse or other indirect pointer. Touch interfaces have been shown to affect choice (Shen, Zhang & Krishna 2016) and alter information salience (Brasel & Gips 2015) when compared to non-touch interfaces. The largely vertical orientation of smartphones and tablet screens in comparison to the horizontal layout of traditional computer screens create new design considerations for how information can be presented. Swiping between pages has become a ubiquitous interface gesture, but horizontal swiping has received little exploration in the interface literature (Kerckhove & Pandalærae 2018; Dou & Sundar 2016).

In this paper, we argue that moving left or right between ‘panes’ of information via swiping creates a different cognitive structure than vertical scrolling. This pane-switching creates a natural break-point in the information, both visually and in cognitive processing, suggesting that the information in separate panes is somehow discrete or separate from information in other panes. This is consistent with tabbed web-browsing suggesting separate source pages as users move horizontally across tabs (Warr & Chi 2013). This should increase the amount of perceived information contained within a site, even if the amount of information is held constant. In addition, if information in horizontal panes is considered more separate or discrete when compared to information presented vertically, reviews may be seen as more independent from the storefront, product specifications may

---

**SESSION OVERVIEW**

The consumer market is increasingly dominated by personal mobile devices, with the share of all e-commerce transactions on touch devices, such as tablets and smartphones, growing three times faster than US e-commerce overall. The increasing ubiquity of mobile devices in the market has been accompanied by a shift in the ways in which consumers relate to and interact with their devices. Beyond the usual visual stimulation encountered, consumers now also interact tactically through touch and vibrations, auditorily through voice commands issued via virtual assistants, and most recently via kinetics through augmented reality (AR) movements. Each element of this interactive mélange of visual, tactile, auditory, kinetic input has critical implications for consumer behavior, yet the latter three constituents are much less studied in academic research in marketing. This special session explores how these novel sensory-interface modalities of consumer interaction systematically impact consumer experience.

Specifically, this session addresses two central research streams in consumer-device interaction:

1. How haptic interfaces and vibrotactile feedback affect consumer perceptions and purchasing intent.

The first two papers focus on the first research stream. The opening paper by Brasel examines how horizontal (vs. vertical) touch interface interactions influence consumers’ information perception. Brasel shows that swiping between horizontal (vs. vertical) ‘panes’ of information creates perceptions of discrete breaks between information panels, making reviews appear more independent, and increasing perceived product information accuracy. In the second paper, Hampton and Hildebrand focus on another unique feature of haptic interfaces: vibrotactile feedback. They find that certain vibrations of intermediate length are perceived as rewarding, while longer vibrations are perceived as punishing. They apply this finding by pairing rewarding vibrational feedback (vs. non-vibrating) with online shopping behavior and show how vibrational feedback can stimulate purchase.
be viewed as more accurate, and consumers may feel they have processed more information and thus be more confident in their choices.

In Study 1, 104 laboratory participants used an Android Tablet to shop for a pair of noise cancelling headphones. In the Horizontal condition, the information for each pair of headphones was separated into four “panels” of information that could be swiped between left-to-right: Description, Specifications, Customer Reviews, and Ordering Information, each one tablet screen long in vertical length. In the Vertical condition, the four panels of information (with the same content) were arranged vertically and scrolled top-to-bottom; the information was kept one tablet screen wide in horizontal width. The participant could move between the four pairs of headphones either via a vertical menu bar running down the left of the screen in the Vertical condition, or a horizontal menu bar running across the top of the screen in the Horizontal condition. Results show that participants felt the Horizontal design site contained more information than the Vertical design site (5.7 vs 5, \(p < .05\)), and reporting evaluating both more products (3.4 vs 2.9 \(p < .05\)) and more of the information (85% vs 72% \(p < .01\)) when making their choice. Consistent with predictions, Horizontal participants also rated the product reviews as significantly more trustworthy and objective than the Vertical orientation participants \(p < .05\), and predicted the product specifications were marginally more accurate as well \(p < .10\). These results combined to increase overall choice confidence and predicted product satisfaction for participants in the Horizontal condition versus the Vertical condition.

In Study 2 (120 participants, similar stimuli to Study 1), for a cleaner test of the touch interaction and to explore whether an effect of orientation remained even if information in both orientations was similarly ‘paneled’, smooth scrolling was disabled so the only way to move between information was either horizontal or vertical swipes that would move a full panel of information at a time. While the Horizontal-vs-Vertical manipulation’s effect on perceptions of the amount of information available and perceived independence of reviews was reduced compared to study 1 (roughly halving in measures of effect size and power), the results remained significant (both \(p < .05\)), reinforcing that directionality plays as much of a role beyond just information compartmentalization via paneling. Respondents also rated the products as more varied when explored via the Horizontal versus Vertical menu bar (4.6 vs 3.8, \(p < .05\)). Participants in the Horizontal condition also recalled visiting significantly more panels of information in comparison to the Vertical orientation (similar to perceptions of information completeness in Study 1) when video evidence revealed there was actually no significant difference in the number of visited panels. This pattern of results is consistent with horizontal information paneling creating a higher salience of breaks between information, elevating perceptions of overall information and information independence.

In conclusion, touchscreen e-commerce sites that use horizontal panels of information that the user can swipe between (versus vertical information structures) create perceptions of discrete breaks between panels of information. This makes reviews appear potentially more independent, while increasing perceived accuracy of product specifications. It also increases the number of alternatives considered and artificially inflates perceptions of the amount of information considered. Future work could explore whether this effect may decrease with habituation, as horizontal swipes are a relatively recent addition to the e-commerce interface lexicon. It may also be useful to explore whether horizontal panel order leads to bias in product information processing, as other applications (such as dating apps) are creating a haptic language where right-swipes are positive and left-swipes are negative.

**Good Buzz, Bad Buzz: Using Vibrotactile Feedback to Shape Consumer Choice**

**EXTENDED ABSTRACT**

Nearly all consumers now carry a vibrating mobile device. Aside from the vibrating mobile phones and tablets many of us carry with us, an increasing number of consumer products have integrated vibrotactile feedback. For instance, the clip-on Lumo Sensor vibrates when its wearer begins to slouch (Peppet, 2014), the HAPIfork vibrates if the user is eating too quickly (Green, 2018), and the Fitbit wristband vibrates when a user reaches a fitness goal (Vanhemert, 2015). Despite the increasing ubiquity of devices producing vibrotactile feedback, rigorous research investigating how users process this feedback, and how it affects their decision making, is unclear.

Mechanical vibration, such as that emanating from a mobile phone, generates traveling waves of energy across the surface of the skin. Information from the skin, i.e. periphery, is carried to the spinal cord on individual nerve fibers, then to one or more second-order neurons which interface with the brain where the stimulus is processed. Tactile thresholds can be extraordinarily small with some people able to feel vibratory stimuli as minute as 250 Hz on the palm of the hand (Gescheider, Cappraro, Frisina, Hamer, & Verrillo, 1978). In parallel, consumer research has begun to underscore the importance of haptic properties of consumer goods (Peck & Childers, 2003) and communications (Peck & Wiggins, 2006), as well as their impact on purchase intention, donation behaviors, and feelings of ownership and endowment (Peck & Shu, 2009; Shu & Peck, 2011). Adding to this stream of research, Brasel and Gips focus on the touch-aspects of computer interfaces (Brasel & Gips, 2014a). Their findings suggest that the physical interaction with touch screens enhances consumers’ sense of psychological ownership and endowment.

While there is evidence that the length of a stimulus can have an effect on the perceived intensity (Gescheider, Berryhill, Verrillo, & Bolanowski, 1999), measures more relevant to consumer behavior such as pleasantness and reward as well as the downstream consequences on economically relevant outcomes (such as market basket size) have been missing in previous studies. Similarly, although one study using infants hints that vibration may be a potent reinforcer (Schaefer, 1960), it remains unclear whether and how mobile device vibrations modify behavior in adults. Finally, although no study has examined use of vibration as a reward in healthy adults, a study using clinical population suggests that vibration may have long-term effects in modifying behavior (Bailey & Meyerson, 1969). Inspired by this line of prior work, the current research is specifically designed to provide new insight into the perception and influence of vibrotactile feedback on consumer decision making.

In Study 1, one hundred and fifty participants were recruited via Amazon Mechanical Turk. Participants were asked to click buttons that produced a vibration of a particular duration, and then self-report their perceptions. We also assessed a broad range of potential control variables at the end of the study, including general mobile technology usage behaviors (such as the usage of vibration features), sensation seeking, need for touch, and experiential thinking.

We find that reward perceptions vary according to vibration duration \(F(3,147)=6.52, p<.001\). Specifically, vibrations of either 400 ms or 800 ms were perceived as significantly more rewarding than all other durations (all \(p<.01\)). The manner in which responses varied also appeared to be construct-dependent. For instance, we find that the relationship between vibration duration and perceived reward forms an inverted-u shape, with reward perception peaking around 400 ms. Notably, very short vibrations, e.g. 25 ms, or very
long vibrations, e.g. 3200 ms, were reported as punishing rather than rewarding. In examining the relationship between phone usage and other dependent variables, several theoretically and practically important relationships emerged. Higher mobile phone use positively correlated with both sensation seeking (r=.17, p<.05) and need for touch (r=.19, p=.05). Positive perception of vibration feedback also covaries as a function of participant traits: Participants who rated the vibrotactile feedback as more rewarding, also scored higher on measures of impulsivity (r=-.42, p<.001) and experiential thinking (r=.36, p=.001).

Our first study demonstrated that vibrations of certain durations are perceived as rewards. Given that rewards have been consistently shown to affect motivated behavior (for review, see Papes, Barsalou, & Press, 2015), it follows that vibration feedback could be used to modify behavior. Building on this prior work, Study 2 we presented one hundred and fifty participants with a “Box Task”. This task presented two boxes, with a button below each that adds an item to its respective box. These boxes resembled prototypical online shopping carts. Participants were instructed to press either button as many times as they liked to add items during a set 1.5 minute timeframe. Adding items to Box 1 produced a vibration (400 ms), while adding items to Box 2 did not. As predicted, we find that participants added more items to the vibration-producing box (M=14.69) than other box (M=10.71) that did not produce a vibration.

F(2,148)=4.08, p<.05. We also replicated the previously observed quadratic-shaped relationship between vibration duration and perceived reward. Together our data strongly support the notion that specific time-durations of vibrotactile feedback are perceived as a reward, and when paired with a purchasing behavior like adding an item to a cart, can induce increased purchase intent and greater market basket size. These findings have broad implications for the use of vibrational rewards to modify a wide ambit of other mobile consumer behaviors.

**Augmented Reality Helps Low-Involved Consumers Build Self-Brand Connection**

**EXTENDED ABSTRACT**

Augmented reality (AR) bridges physical and virtual worlds by overlaying an image of a product onto the physical world (Jaivork, 2016). For example, customers use their mobile devices to project how Ray-Ban sunglasses look on their face, or which CoverGirl lipstick looks good on them without physically trying them on. Since strong visual effects in product presentations help consumers imagine actual product experiences (Vessey and Galletta, 1991), we hypothesize that integrating AR into consumer’s shopping experience strengthens consumers’ self-brand connection, especially for low-involved consumers.

How does AR enhance low involved consumers’ connection with brands? We argue that AR-assisted shopping experiences help them to infer that the brands utilizing AR technology are communal, since AR-based assistance is intended to help fulfill consumers’ needs. Consumers build relationships with brands in much the same way as they form relationships with people (Fournier, 1998), and these can be either exchange-based or communal-based relationships (Clark and Mills, 1993). Exchange relationships are monetary transaction based, whereas communal relationships are like friendships, where the focus is on caring and satisfying the partner’s needs. For example, consumers who perceive a communal relationship with a brand are likely to perceive that the brand cares about their personal needs (Aggarwal, 2004).

Together, we predict that AR increases the perception that a brand is communal, which will then enhance self-brand connection. We expect this effect to ensue particularly among low-involved consumers who use peripheral route processing and rely on environmental cues like AR assistance during their choice of products (Petty and Cacioppo, 1986). By contrast, high-involved consumers are more likely to use central route processing and focus on actual product attributes and ad argument quality. Therefore, they are less likely to rely on AR-assisted shopping and the consequent brand image as communal to enhance their connection with the brand.

Study 1 shows that consumers who are low involved in shopping can build strong self-brand connection after using (vs. not using) AR. In a 2 (AR vs. control) × 2 (high vs. low involvement) between-subjects design, 169 participants reported their attitude toward Ray-Ban sunglasses twice – before and after negative brand information was given. This is a common measure of self-brand connection (Ahuwalia, Burnkrant, and Unnava, 2000). Our prediction was that participants who have a strong self-brand connection would be less disturbed by the negative information and continue to prefer the brand (i.e., small change in brand attitude due to negative information). We manipulated involvement (Litt and Tormala, 2010) by varying the importance of this shopping task (e.g., High involvement: “Your opinion counts!” vs. Low involvement: “Broad anonymous group’s preferences”). Participants shopped for Ray-Ban sunglasses, either by overlaying sunglasses onto their own photo (AR condition) or using a default model’s face (Control condition). Among less involved participants, those who used AR (vs. Non-AR) defended the brand by reporting small change in their attitudes after hearing negative information about the brand (M_AR = 0.42 vs. M_non-AR = 1.24; p = .018), whereas AR did not affect highly involved participants’ attitude change.

There were no differences in participants’ choice of sunglasses, their product liking, nor their psychological ownership over the chosen product (all n.s.).

Study 2 replicates the findings using Sephora brand. Female participants (N = 181) were randomly assigned to one of the 4 conditions in a 2 (AR: yes vs. no) × (Involvement: high vs. low) between-subjects design. We again manipulated involvement by varying the importance of the shopping task. Next, participants either used the AR feature or not to shop for lipsticks.

Participants in the AR condition used their webcam to take a picture of themselves, and applied lipsticks on their face using the AR feature. Participants in the non-AR condition used default model pictures provided from Sephora. Then, we measured self-brand connection using scale from prior literature (Escalas and Bettman, 2005). Low involved consumers increased self-brand connection after using (vs. not using) AR (M_AR = 5.46 vs. M_non-AR = 4.61; p = .005). There was no significant difference within the high involvement condition. People’s liking of the chosen lip products also did not differ across conditions. We also measured whether AR enhances mental simulation of the process and the outcome of the consumption experiences; however, we did not find significant differences across conditions (all n.s.).

Study 3 demonstrates that AR’s effect on self-brand connection among low involved participants are driven by their inference about brand’s communal characteristics. In a 2 (high vs low involvement) × 2 (AR vs. non-AR) design, we manipulated involvement by again varying the importance of the shopping task. Then, participants shopped for throw pillows from Amazon app using smartphone. Participants in the AR condition used the AR feature in the app, so they could virtually see an image of the throw pillow on an actual chair in the lab. That is, the app overlaid a digital image of a pillow on par-
participants’ camera view of the chair. Participants in the Non-AR condition shopped for a throw pillow using a typical search procedure on Amazon app. We measured perception on communal relationship with the brand on an 8-item scale (Aggarwal 2004).

Results revealed that among low involved participants, using AR increased their self-brand connection with Amazon (M_{AR} = 4.65 vs. M_{Non-AR} = 4.13; p = .044). There were no differences within the high-involvement condition (n.s.). Supporting our prediction on the process mechanism, we found a significant mediation effect of relative communal score (PROCESS Model 8: CI = [-.34, -.05]). Participants’ liking of the chosen products and positive affect arousal during the shopping task did not differ across conditions.

In summary, our results reveal that consumers who are less involved in shopping infer from AR-assisted shopping experience a communal brand image, which then increases their connection with the brand. In contrast, when consumers are already highly involved in shopping do not rely on such cues and heuristics for their inference, therefore using AR does not increase self-brand connection.

Understanding Consumer’s Product Decisions When Shopping by Voice

EXTENDED ABSTRACT

Despite the rapidly evolving of artificial intelligence and the technology infusion in retailing, sparse research exists on this topic in marketing literature (Grewal, Roggeveen and Nordfält 2017). Compared with traditional online shopping that navigated by point and click (i.e., click-based interaction), shopping with virtual assistant allows consumers to place an order either by talking (i.e., voice-based interaction) or texting message (i.e., text-based interaction). Based on previous research, which suggests that expression modality can influence consumer’s decision making (e.g., Kless, Levav and Goukens 2015; Shen and Sengupta 2018), we propose that voice-based interaction will make consumers more likely to choose the recommended product during their online shopping journey.

Prior research shows that people’s attitude towards an object can be determined by how they interact with it. For example, socially interact with a robot can increase the anthropomorphism of the robot (e.g., Peca et al. 2015). We propose that since voice-based interaction carries more social functions than text-based interaction or click-based interaction (e.g., Chafe 1982), engaging in voice-based interaction will incite people to feel that they’re interacting with a social actor rather than a computer. Consequently, this enhanced social perception will lead to compliance behavior (Cialdini 2016). As a result, consumers will end up purchasing more recommended products. Furthermore, we also argue the interaction effect may only exists in dialogical interaction, where people will get social responses from the computer reciprocally but will diminish in monologial interaction, where the computer’s social responses are absent. This is because social interaction is two-way in nature, while one-way communication will create social distance and undermine the social bonds between communicators (Andreoni and Rao 2011).

A fictitious online food ordering website was developed to examine our hypotheses. The website contains three product categories: pizza, sandwiches and chicken wings. Under each category, five to six products are provided, and some are labeled as recommended (i.e., regular recommended products). Before checkout, participants will be asked whether they would like to add a salad or fries to their orders (i.e., add-on products). The measurement of recommended products includes both regular recommended products and add-on products. We also created different website versions to manipulate both user communication modality (speak vs. click vs. type) and computer’s response modality (social response vs. non-social response). To simulate computer’s social response, we let the website communicate to participants in either synthesized voice or interactive text. Meanwhile, in the non-social condition, we used static text on the website and also made the content less interpersonal.

Five studies have been conducted with responses from more than 3,000 MTurk workers. In all the studies, participants were asked to complete an online shopping task on the given website. Study 1 demonstrates the main effect. It also examines the role of controllability (i.e., whether the communication can be interrupted). Participants were randomly assigned to either shopping by voice (incorporating synthesized speech), click-interruptible or click-non-interruptible (incorporating pop-up text) condition. Consistent with our hypotheses, when shopping by voice, participants purchased more quantity of recommended products (p < .01) and also more add-on products (p < .01). No differences were found between the two click conditions (ps > .06), suggesting controllability doesn’t influence product decisions.

Study 2 examines the effects of both user communication modality (speak vs. click) and computer’s response modality (synthesized speech vs. interactive text). We also included a plain website without any extra website information as baseline. The results revealed that speak modality leads to more recommended products (p < .01), and more add-on products purchased (p < .01). We also investigated the possible influences of computer response modality and controllability individually, however, the results didn’t reveal any significant differences on the major measurements (ps > .28).

Study 3 compares voice-based interaction with both click-based interaction and text-based interaction. Consistent with the last two studies, in the voice-based interaction, participants purchased more recommended products (p < .01) and more add-on recommended products (p < .01). Again, no significant differences were found between synthesized voice and interactive text conditions on the major dependent measurements (ps > .15).

Study 4 employs a 3 (user’s communication modality: speak vs. type vs. click) by 3 (computer’s response modality: synthesized voice vs. interactive text vs. non-social information) between subject design and investigates both user’s communication modality and the moderating role of computer’s response modality. We found that using speak modality leads to more recommended products purchased only when computer responds in synthesized voice (p < .01) or in interactive text (p < .01), but not in the monologial condition without computer’s social responses (p = .65). Similar pattern was also observed for the add-on product purchases.

Study 5 provides further support to the interaction effect between user’s communication modality and computer’s response modality. In addition, it also examines the proposed mediating role of perceived socialness (Wang et al. 2007). Specifically, using speak modality leads to more recommended products purchased only when computer respond in synthesized voice (p < .1) or in interactive text (p < .01), but not in the monologial condition where computer’s social responses are missing (p = .16). To test the mediation, with user’s interaction modality type as the independent variable, computer’s response modality as the moderator, perceived socialness as the mediator, and the number of recommended products purchased as the dependent variable, the 95% bias-corrected bootstrap confidence interval, obtained using 10,000 bootstrap samples, did not include zero, indicating a significant indirect effect of perceived socialness on choosing of recommended product.

In sum, this research shows that when shopping by voice, consumers are more likely to choose recommended products, and this effect is driven by perceived socialness. This work contributes to
existing communication and human-computer interaction literature and also provides important managerial implications to practitioners for designing a more effective virtual assistant system.

REFERENCES


Shen, H., Zhang, M., & Krishna, A. (2016) “Computer interfaces and the “Direct-Touch” effect: can iPads increase the choice of hedonic food?” *Journal of Marketing Research*


