The Neural and Psychological Basis of Herding in Purchasing Books Online: An Event-Related Potential Study

Mingliang Chen, Ph.D., Qingguo Ma, B.A., Minle Li, B.A., Shenyi Dai, Ph.D., Xiaoyi Wang, Ph.D., and Liangchao Shu, Ph.D.

Abstract

In this study, event-related brain potentials (ERPs) were used to investigate the neural and psychological bases of consumer herding decision in purchasing books online. Sixteen participants were asked to decide as quickly as possible whether to buy a book on the basis of its title keywords and the numbers of positive and negative reviews in stimulus. The given title keywords were very similar, and participants did not have special preference for any particular one. Hence, they were forced to adopt the strategy of herding decision: choosing to buy the book when there were consistent positive reviews, choosing not to buy when there were consistent negative reviews, randomly choosing to buy or not to buy when there were no consistent reviews. The herding decision triggers a categorical processing of the consistency level of customer reviews. Remarkable late positive potential (LPP), a component of ERP sensitive to categorization processes, was elicited. The LPP amplitudes varied as a function of review consistency. The LPP amplitudes for three categories of review consistency were significantly different, and their order is such that absolute consistent review was greater than relative consistent review, which was greater than inconsistent review. In addition, behavioral data revealed that the higher the consistency of the customer reviews, the higher the herd rate. It is possible that customer reviews with higher consistency let participants make herding decisions more resolutely. The present results suggest that the LPP may be regarded as an endogenous neural signal of the herding mechanism in a sense and that the LPP amplitude is potentially a measure of consumers’ herd tendency in purchase decisions.

Introduction

Herding behavior describes various social situations in which individuals are strongly influenced by the decisions of others and everyone does what everyone else is doing. Herding effect is widely used by marketing practitioners to induce consumer purchase intentions. When consumers use the product evaluations of others to indicate product quality and make their purchasing decision based on the opinions of most others on the Internet, herd behavior in online shopping occurs. Herd behaviors in online shopping occur widely for many product categories, such as books, electronics, games, videos, music, beverages, and wine. There are two reasons. First, the Internet and information technology provide a new opportunity for consumers to share their product evaluations online. In online environments, consumers can search and get the product evaluations of others conveniently via the sellers’ Web sites, message boards, Internet forums, chat rooms, and so on. Second, these products are essentially low-involved products and are relatively low priced. Compared with high-involved and high-priced products, consumers spend less time and cognitive efforts in considering low-involved and low-priced product choices. Such products are more likely to elicit online herd behavior.

In order to induce consumer purchase intentions, it is important to understand the occurrence mechanism of online herd behavior. Previous studies have examined herding on the Internet, such as in digital auctions, software downloading, and purchasing books online. These studies focused on the determinants of online herding or its impacts on consumer decisions by means of surveys and behavior tests. Scholars and scientists have not started to investigate the neural basis of herd behavior. This study tries to investigate the neural basis and psychological process of herding decision that leads to online herd behavior.

Books are experience products, which are low involved and low priced. Consumers’ herd behaviors are evoked easily in purchasing books online, so we chose book purchases as the background to investigate the neural and psychological basis of herding decision. There are two additional reasons to choose this particular research context. First,
many people have experienced herding in purchasing books online, and we were able to recruit suitable participants conveniently. Second, the cues for evoking herd behavior in real book-purchasing environments can be simulated easily in laboratory. Previous studies have demonstrated that customer review and sales volume are the main cues for eliciting herd behavior in purchasing books online.\(^4\) Star ratings (average number of stars the reviewers assigned on a scale of 1 to 5 stars, with 5 stars being the best) are used as an indicator of product quality by consumers.\(^4,10\) When consumers make their buying decisions based on the valence of all customer reviews, online herd behavior occurs. The information about sales volume, such as best-seller lists, is shown to drive consumer reviews, online herd behavior occurs. In order to simplify experiment conditions, this study used only the customer review (consumer book review) as cues to evoke herd behavior in purchasing books online.

People tend to believe what most others believe, even though these beliefs may not be true.\(^12\) This tendency results in herd behavior. The essence of consumer herd behavior is that consumers make their buying decisions based on the product evaluations (reviews) of other consumers. The following is a possible mechanism of decision making: when the reviews of most customers are positive, the consumer decides to buy it; when the reviews of most customers are negative, the consumer decides not to buy it. This is an example of herding decision leading to herd behavior. But before making herding decisions, consumers must judge whether the product reviews of most customers are positive or negative. Five situations may occur in customer reviews: all customer reviews are negative; most customer reviews are negative; customer reviews are not consistent; most customer reviews are positive; and all customer reviews are positive. These five situations are the five categories of consistency level of customer reviews: absolute negative review, relative negative review, inconsistent review, relative positive review, and absolute positive review. Obviously, herding decision includes a categorization process of review consistency. The categorization of others’ opinions consistency in herding decision is the focus of our studies on herd behavior.

Given that herding decision is a categorization processing, it is possible that event-related brain potential (ERP) opens the neural basis of herding decision. Numerous studies have explored the neural mechanisms underlying categorization based on ERP.\(^13-15\) Several components of ERP can be used to examine categorization processing. Simple categorizations (e.g., the discrimination between animal and nonanimal photos) evoke some early components, such as N1.\(^15,16\) The main components of ERP elicited by complex categorizations are P300 and late positive potential (LPP).\(^14,15,17\) For example, the classification of male-or-female name\(^16\) and semantic human face\(^15\) evoked P300; the implicit and explicit categorization of stimuli along evaluative (positive and negative) and nonevaluative (people and no-people) dimensions evoked LPP;\(^14\) the visual categorization evoked LPP.\(^17\)

The categorization of customer review consistency is a complex process, so this study pays attention only to P300 and LPP. The P300 component is measured by assessing its amplitude and latency. Amplitude (\(\mu V\)) is defined as the voltage difference between a prestimulus baseline and the largest positive-going peak of the ERP waveform within a latency window (e.g., 250–400 ms). Latency (measured in milliseconds [ms]) is defined as the time from stimulus onset to the point of maximum positive amplitude within the latency window.\(^19,20\) The P300 component reflects neural activity related to basic aspects of cognition. However, because the P300 is derived from neural activity, it is necessarily affected by the physical state of its underlying physiology.\(^19\) The LPP shares many signature characteristics of the P300, such as in scalp distribution and eliciting conditions.\(^21\) But their peak latencies are different. The peak latency of P300 is at approximately 300 ms poststimulus presentation,\(^22\) and LPP appears after P300, so the LPP is interpreted as a long latency P300\(^21\) or sustained P300-like component.\(^23\) As the two categories of late positive component of ERP, P300 and LPP provide a great deal of information about the neural activity of fundamental cognitive operations, especially the activity of categorization processing.\(^14,24,25\) Their amplitudes reflect the allocations of attentional resources. The more the attentional resources captured by participants are, the larger the amplitude of P300\(^23\) or LPP.\(^26\) Their peak latencies reflect the stimulus classification time.\(^18,27\) The more complex the categorization is, the longer the peak latency. When a categorization is complex enough, it needs process time that is long enough, and it is more likely to evoke LPP.

We designed an ERP experiment to study the neural basis of consumer’s herding decisions in purchasing books online. In this study, participants were asked to decide whether to buy a book on the basis of the title keyword and of positive and negative review numbers shown in the center of a computer monitor. This is a simulation environment for purchasing books online. Participants were told that the positive review number equals the sum of the numbers of 5-star and 4-star reviews, and the negative review number equals the sum of the numbers of 1-star and 2-star reviews, and 3-star reviews are neglected. In addition, in order to avoid the difference of the attraction of book title, only title keyword rather than concrete title was provided in stimulus. The given title keywords were similar, and the participants had no special preference for any particular keyword. Under the circumstances, title keyword was neglected, and the numbers of positive and negative reviews were almost all information that can be used to make decisions by the participants. So herding strategy based on the valence of all customer reviews is the unique rational mechanism for participants to decide whether to buy the book. In other words, participants choose to buy the book in the context of consistent positive reviews, choose not to buy in the context of inconsistent negative reviews, and randomly choose to buy or not to buy in the context of inconsistent reviews. It is apparent that the process of herding decision includes a categorization process of customer review consistency. In this experiment, we speculated that P300 or LPP, two categories of the ERP component sensitive to complex categorization process, would be recorded.

Materials and Methods

Participants

Sixteen undergraduates from Zhejiang University in China (9 men, 7 women; age 21–25 years, \(M = 22.5\) years) gave written informed consent and participated in the experiment as paid volunteers. All participants were right-handed, had
normal or corrected-to-normal vision, and had no history of neurological or mental diseases. In addition, all of them were enrolled in customer relationship management (CRM) courses and needed to buy reference books about CRM.

**Materials**

A stimulus was composed of three items: a book title keyword, the number of positive reviews, and the number of negative reviews about the book. Forty-five books from the field of CRM were offered as choices to participants. The title keyword of each book was limited to four Chinese characters. The total number of reviews presented in each trial for a book was randomly generated from the interval of 400 to 500, which was determined by the mean number of reviews on the books in the best-seller lists at the Web sites of Amazon and Dangdang (the largest e-bookstore in China). The percentages of the number of positive reviews in the total number of reviews were fixed at 0, 25, 50, 75, and 100, corresponding to absolute negative review, relative negative review, inconsistent review, relative positive review, and absolute positive review respectively. So the stimuli consisted of 225 pictures (i.e., 45 books×5 categories of review consistency).

**Procedure**

Participants sat on a comfortable sofa located in a shielded room and were instructed to avoid blinking or moving their eyes and to keep their eyes fixed on the monitor located 1 m away from his or her eyes. Those stimuli (white on a black background) were presented both sequentially and randomly in the center of a computer screen, with a visual angle of 2.58°×2.4°. Each trial began with a fixation point (+), which appeared for 500 ms. Then a stimulus was presented for 1400 ms. Based on title keyword, positive review number, and negative review number in stimulus, participants were asked to decide as quickly as possible whether to buy a book. They were instructed to press the left key if deciding to buy and the right key if deciding not to buy. The next trial started after the participant responded or after 1400 ms had elapsed since the onset of the stimulus. In order to reduce carryover effects, the previous behaviors of participants would influence their later behaviors, 225 trials were divided into four blocks, and the sequence of trials in each block and the sequence of blocks were randomly assigned. There was a 5-minute interval between the two blocks for the participants to rest. The trials in which participants did not press either the left or right key were considered invalid.

**Electroencephalogram recording and analysis**

Electroencephalogram (EEG) was continuously recorded (band pass 0.05–100 Hz, sampling rate 500 Hz) with Neuroscan Synamp2 Amplifier (Scan 4.3.1, Neurosoft Labs, Inc., Sterling, VA), using an electrode cap with 64 Ag/AgCl electrodes mounted according to the extended international 10–20 system and referenced to linked mastoids. Vertical and horizontal electrooculograms were recorded with two pairs of electrodes, one pair placed above and below the right eye, and the other 10 mm from the lateral canthi. Electrode impedance was maintained below 10 kΩ throughout the experiment. EEG recordings were segmented for the epoch from 100 ms before onset of each picture appearing on the video monitor to 800 ms after this onset with the first 100 ms prestimulus as a baseline. Electrooculogram artifacts were corrected using the method proposed by Semlitsch et al. Trails contaminated by amplifier clipping, bursts of electromyographic activity, or peak-to-peak deflection exceeding ±80 μV were excluded from averaging. The remaining trials were baseline corrected. The EEG epochs were averaged separately for absolute consistent review, relative consistent review, and inconsistent review, and the averaged ERPs were digitally filtered with a low-pass filter at 30 Hz (24 dB/octave). To investigate the neurophysiologic correlates of the processing of different review consistency categories, we compared the amplitudes of the three ERPs using a within-participants repeated measure ANOVA. The ANOVA factors were review consistency category (absolute consistent review, relative consistent review, inconsistent review) and electrode site (P1, PZ, P2, CP1, CPZ, CP2, C1, CZ, C2).

**Results**

**Behavioral data**

Behavioral data are given in Table 1, which shows the mean response times (RTs) and herd rate (or buy proportion) in three different categories of customer review consistency. Table 1 reveals that when customer reviews are consistent, the herd rate is more than 90%; and when customer reviews are inconsistent, the buy proportion is about 50%. This result supports our speculation that herding decision based on the valence of all customer reviews is a main mechanism for participants to make purchasing decision in the given experiment conditions. The repeated-measure ANOVAs for RTs in three conditions (absolute consistent, relative consistent and inconsistent reviews) and for herd rate in two conditions (absolute consistent and relative consistent reviews) were conducted respectively, and indicated that the category of review consistency had a significant main effect on RTs, F(2, 15) = 54.909, p = 0.000, and herd rate, F(1, 15) = 5.319, p = 0.036. Post hoc tests revealed that the higher the consistency of customer reviews, the shorter the RTs and the higher the herd rate. These results suggest that the higher the consistency

<table>
<thead>
<tr>
<th>Condition</th>
<th>RTs</th>
<th>Herd rate (or buy proportion)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Absolute consistent review</td>
<td>625.0452</td>
<td>97.23% (2.63%)</td>
</tr>
<tr>
<td>Relative consistent review</td>
<td>721.2417</td>
<td>90.90% (11.38%)</td>
</tr>
<tr>
<td>Inconsistent review</td>
<td>827.3736</td>
<td>47.92% (27.90%)</td>
</tr>
</tbody>
</table>

Note: Absolute consistent review is a combination of absolute negative review and absolute positive review; and relative consistent review is a combination of relative negative review and relative positive review. Herd rate is a percentage of herding decision number to total decision number. When participants decide to buy in the context of consistent positive reviews or not to buy in the context of consistent negative reviews, herding decision occurs. Whether participants decide to buy or not to buy in the context of inconsistent reviews, herding decision does not occur. Buy proportion is the percentage of the number of purchases of a given book among the total number of books offered.
of customer reviews, the easier its categorization processing and the more resolute the participants’ herding decision.

**Event-related potential data**

Following the onset of the stimulus, remarkable LPP was recorded from widespread scalp areas in all three conditions of review consistency. The latency of the LPP was approximately 320 to 650 ms. Figure 1 shows the grand-averaged ERP waveforms in each condition at sites P1, PZ, P2, CP1, CPZ, CP2, C1, CZ, and C2.

Mean LPP amplitudes in the 350-ms to 650-ms time windows were analyzed. To examine the main effect of the LPP evoked by review consistency categorization processing, the 350-ms to 650-ms time window of LPP is divided into three equal sections: 350–450 ms, 450–550 ms, 550–650 ms. Then we performed three 3×9 (category of review consistency: absolute consistent review, relative consistent review, inconsistent review by scalp site: P1, PZ, P2, CP1, CPZ, CP2, C1, CZ, C2) within-participants repeated measure ANOVA on their amplitudes. For the time window of 350 to 450 ms, main effects of review consistency category, F(2, 30) = 7.805, p = 0.002, were found. Post hoc tests revealed that the LPP amplitudes evoked by relative consistent reviews and inconsistent reviews were not significantly different and that the LPP amplitude evoked by absolute consistent review was significantly larger than the LPP amplitude evoked by relative consistent reviews or inconsistent reviews. For this time window section, main effects of scalp site, F(8, 120) = 7.894, p = 0.000, was found, and significant interaction between review consistency category and scalp site, F(16, 240) = 1.670, p = 0.144, was not found. For the time windows of 450 to 550 ms and 550 to 650 ms, main effects of review consistency category, F(2, 30) = 12.813, p = 0.000; F(2, 30) = 5.566, p = 0.009, were obtained. Post hoc tests revealed that the higher the review consistency, the larger the amplitude of LPP. The order of LPP amplitude is as follows: absolute consistent review was greater than relative consistent review, which was greater than inconsistent review. For the latter two time window sections, main effects of scalp site, F(8, 120) = 3.973, p = 0.000; F(8, 120) = 2.466, p = 0.016, were obtained, and no significant interaction between review consistency category and scalp site, F(16, 240) = 1.283, p = 0.279; F(16, 240) = 1.127, p = 0.353, were obtained.

**FIG. 1.** Grand average event-related potentials at P1, Pz, P2, CP1, CPz, CP2, C1, Cz, C2 for absolute consistent reviews (heavy lines), relative consistent reviews (solid lines) and inconsistent reviews (dotted lines) conditions.
Figure 2 displays 64-channel topographical voltage maps across the scalp for maximal amplitudes of the LPP (480 ms). The visual inspection of the topographical maps shows that the peak potential of LPP was distributed over the central-parietal scalp regions in three categories of review consistency. The LPP for absolute consistent review was more remarkable than that for relative consistent review and inconsistent review, and it is consistent with the outcomes from ANOVAs.

Discussion

Previous studies found that LPP could be evoked by some complex categorizations. Our study reveals that (a) remarkable LPP were elicited in all three categories of customer review consistency; (b) the amplitudes of LPP for three categories of review consistency were significantly different (their order is absolute consistent review > relative consistent review > inconsistent review).

These results suggest the following three inferences, at least. In order to explore the psychological process of herding decision, we explain these inferences mainly from a psychological viewpoint.

First, the results suggest that herding decision in purchasing books online includes an explicit categorization processing of customer review consistency, as we speculated. In our experiment, a stimulus consists of a title keyword, the number of positive reviews, and the number of negative reviews about a book. According to the given information, the participants must decide within 1400 ms after the onset of stimulus whether to buy a book. As mentioned previously, title keywords are very similar, and the participants have no special preference for any particular one. The participants are forced to adopt herding strategy to make a purchase decision: when most reviews are positive, they decide to buy the book; when most reviews are negative, they decide not to buy. Apparently, the core process of herding decision is to judge what the opinion of most reviews is. The judgment triggers a categorical processing of the consistency level of customer reviews. There are multiple methods to classify review consistency. The five-category method was used in this study: it classifies review consistency into absolute negative review, relative negative review, inconsistency review, relative positive review, and absolute positive review.

We speculated that the participants first classify the consistency level of current customer reviews as one of the five categories before they make a purchase decision. The findings of previous studies strongly support interpreting the LPP as reflecting neurophysiologic mechanisms of categorical processing. The present experiment reveals remarkable LPP, which was exactly elicited by the categorization processing of customer review consistency. Thus, the result supports our speculation that the participants’ herding-decision process includes a categorization processing of customer review consistency.

Second, the results suggest that the categorization of customer review consistency is a very complex categorization, and it could elicit LPP. The time course of the categorization of review consistency involves four stages: (1) selective attention: attending only to the numbers of positive and negative reviews in stimulus, neglecting book title because the attractions of various title keywords were not significantly different; (2) information extraction: extracting the numbers of positive and negative reviews; (3) valence computation: computing the valence of customer reviews and its size; (4) category judgment: judging the category of review consistency. Figure 3 illustrates the four stages.

The categorization processing of review consistency involves the extraction of review information and the computation of proportion of positive review number to negative review number. Compared with other categorizations in previous studies investigating P300 or LPP (e.g., classification of male and female name, semantic human face, positive picture and negative picture, picture with people and without people, vegetable food and nonvegetable food), the categorization of review consistency is much more difficult, so it is reasonable that remarkable LPP was evoked by the categorization of review consistency in the present experiment.

Third, the present results suggest that the consistency level of customer reviews determines the amplitude of LPP, and
the higher the consistency level of customer reviews, the larger the amplitude of LPP. The differences in LPP amplitude among three categories of review consistency can be interpreted in three ways. A first explanation is based on category similarity. Similarity is a central factor in classification, and most psychological models of categorization are based on similarity. Past studies found that category similarity would affect the amplitude of P300. The higher the similarity, the larger the amplitude of the P300. The LPP is a kind of late positive component of ERP very similar to the P300 and is also sensitive to the categorization process. The categorizations evoking P300 and LPP should involve the similar cognitive and neural processes. An obvious inference is that the amplitude of LPP is also sensitive to category similarity. Based on the inference, we can explain why higher review consistency elicits the larger amplitude of the LPP in our study. The absolute consistent review is a criterion category of review consistency in all possible categories. It is possible that the discrimination of all other consistency categories is based on their similarity to the criterion category. For example, absolute consistent review has complete similarity to the criterion category, inconsistent review has no similarity to the criterion category, and relative consistent review has relatively high similarity to the criterion category. They would produce the largest, the smallest, and the intermediate LPPs respectively. The results of our experiment are consistent with this deduction. A second explanation is based on the difference in decision difficulty and/or confidence. LPP or P300 amplitude is inversely related to decision difficulty and/or is positively related to decision confidence. In our study, it is obvious that the higher the consistency of customer review, the easier the purchase decision and the higher the decision confidence. So we obtained the result that the LPP amplitude is positively related to the consistency level of customer review. A third explanation is based on the difference in the allocation of attentional resources. The results suggest that the stimuli with higher consistent reviews tended to attract greater processing resources. Past studies have shown that attentional resources affect the amplitude of P300 and LPP and the stimuli with higher similarity capture attention more easily and lead to larger amplitude. In our experiment, it is possible that the customer reviews with higher consistency level excites the participants and attracts more attentional resources, because this information lets them make herding decisions more resolutely and confidently.

The results of this study suggest that LPP can be elicited by the categorization processing of customer review consistency embedded in herding decision in purchasing books online. The higher the review consistency, the larger the LPP amplitude. In addition, behavioral data revealed that the higher the review consistency, the shorter the RTs and the higher the herd rate. These conclusions have an important methodological implication: the techniques of neuroscience (such as ERP used in this study) can improve the understanding of herding phenomena. First, the ERP results support the conclusion of behavioral data. In this study, behavioral data and ERP data were recorded simultaneously. The former revealed that higher review consistency led to higher herd rate and shorter RTs. The latter revealed that higher review consistency elicited larger LPP amplitude. Larger LPP amplitude means that customer reviews with higher consistency are more likely to lead participants to make herding decisions more resolutely. Obviously, both data reached virtually the same conclusions. Second, ERP results deliver addition information. In this study, ERP data revealed that herding decisions evoked remarkable LPP, which reflects neurophysiologic mechanisms of categorical processing. The result means that the participants’ herding-decision process includes a categorization processing of review consistency. This is a deduction that behavioral data cannot deliver. Third, ERP can help us explain herding decision from the level of neural mechanism. In this study, it is possible that the LPP indirectly reflects consumers’ herding mechanism in purchase decisions. Taken in this sense, the LPP should be regarded as an endogenous neural signal of herding mechanism. In the researches and practices of word-of-mouth (WOM)–based herding marketing, companies could potentially use the LPP as an endogenous neural indicator for measuring consumers’ herd tendency in purchase decisions. Consider, for example, a herding marketing program based on WOM recommendation agent (WOM-RA). WOM-RAs are “virtual consultants” from online shops that provide consumers with statistical information on customer WOMs about a given product, such as the total number of WOMs, the numbers of positive and negative WOMs, and the proportion of positive WOMs and negative WOMs. The aim of this marketing program is to evoke consumers’ positive herd tendency through advantageous statistical information of WOMs and induce consumers to buy recommended products. How can we evaluate whether this herding marketing program will evoke a strong enough herding effect? In addition to measuring herd rate using behavioral experiment, we can adopt ERP technique to measure LPP elicited by participants’ herding decisions in simulated purchase environments. Both herd rate and LPP can be used to reflect herding effect. Compared to herd rate,
LPP is a more objective indicator, and they are mutually complementary.

Notes

a. The 1400-ms time window was chosen for two reasons. First, the pre-experiment showed that participants were able to finish their purchase decisions in 1400 ms, although feeling a little time pressure. Second, appropriate time pressure helps to promote participants to adopt a herding mechanism, which is what we expected to happen. When extracting information useful to purchase decisions, participants neglect the title keywords, especially under time pressure. Under the circumstances, herding becomes a rational decision strategy.

b. In the experimental instruction, the participants were told that (1) the price of every book is assumed to be the same—for example, 30 Yuan (Chinese unit of currency), (2) the purchasing decision on each book is an independent one.

c. Ito and Cacioppo define an explicit process as requiring intentional and conscious control.14 According to their definition, when a categorization becomes the direct or indirect task of participations, it is an explicit categorization. In our experiment, participations’ direct task is to make a book-purchasing decision, but they must first classify the consistency of customer reviews. Categorization is participations’ indirect task, so it is explicit.

Acknowledgments

This work was supported by grant No. 70372020, No. 70671092, No. 70971116, and No. 90924304 from the National Natural Science Foundation of China.

Disclosure Statement

No competing financial interests exist.

References


33. Cutmore TRH, Muckert TD. Event-related potentials can reveal differences between two decision-making groups. Biological Psychology 1998; 47:159–79.


Address correspondence to:
Qingguo Ma
School of Management
Zhejiang University
38# Zheda Road
Hangzhou 310027
People’s Republic of China

E-mail: neuromanagementlab@gmail.com