In today’s emerging and competitive live streaming e-commerce, streamers must continually work to improve their sales performance. Past research has highlighted the performance impact of live interaction, treating its effects as enduring based on general or average levels. Live interaction in some moments is more salient than that in others, however. Building on gestalt characteristics and event system theories, this study takes a short-term momentary perspective to examine the intensity and duration of peak interaction during product display and non-product display phases of a live streaming video. Using moving window and computer vision techniques, these variables are extracted from videos data obtained on the Taobao Live platform. Their effects are then empirically quantified with econometric model estimations.

Keywords: Live streaming e-commerce, Peak interaction, Duration, Phases, Sales performance

1. Introduction
Live streaming, an emerging type of live information diffusion and communication technology, has been widely adopted by online sellers to boost their sales performance. The growth of live streaming e-commerce (LSE) has been accompanied by increasingly intense competition among streamers. In China, for example, about 20,000 streamers sold products via Taobao Live in 2019, an increase of 233.3% from 2018. In 2020, Taobao Live attracted 6 billion USD on the Double 11, China’s largest annual online marketing event. Considering these huge numbers and growing trends, streamers face major challenges in maximizing sales performance in the LSE context, and are keen to understand how to achieve this.

LSE enables the synchronous conversation between streamers and consumers. Such live interaction (LI) is significantly different from live chat tools in traditional e-commerce (Ou et al. 2014; Tan et al. 2019). Compared with live chat that only allows for one-to-one interactions in text or image formats, LSE helps streamers interact with multiple buyers in a real-time video (Sun et al. 2019; Xue et al. 2020). Specifically, a consumer can type text comments on the bullet screen, which is shown to peer consumers and streamers. Peer consumers can then make responses after reviewing the comments sent by the consumer. Streamers can also review and react to these comments through visual and verbal demonstration synchronously, allowing the interactivity of content to be manipulated in a broader and more effective way (Sun et al. 2019).
Understanding the performance impact of LI has been of great interest to LSE researchers in recent years (e.g., Xue et al. 2020). However, the momentary or episodic features of LI during a video remain understudied. As live interaction intensity changes on a momentary basis, where some moments are more salient than others (Kang et al. 2021), this feature of LSE warrants further examination. For example, as consumers enter the live streaming video, streamers greet consumers and call for their feedback via visual and (verbal) sound responses. When the group of consumers is large enough, streamers begin to introduce a sequence of products and talk about the price incentives and selling policies of a certain product at a particular moment before moving on to the next item (Wongkitrungrueng et al. 2020). In other words, a live streaming video featuring a particular product is not a single timestamp, but the dynamic unfolding of communication, where the level of interactivity intensity can be expected to fluctuate. This suggests that consumers may be more influenced by these interactions at some moments than at others, where information on products may also be retained at different levels over time. This study therefore develops a short-term episodic view of interaction and its effects during a live streaming video. Treating LSE interaction as having momentary rather than enduring effects, our research question is posed: How do momentary interactions influence sales performance in LSE?

We propose a novel and temporal feature of LI based on gestalt characteristics and event system theories: peak interaction, the moment when interactivity intensity is at its highest level. Gestalt characteristics theory suggests that the most intense (peak) moments of individuals’ experienced event significantly shape their overall summary evaluations of this event (Ariely and Carmon 2000; Geng et al. 2013; Jiang et al. 2019). We argue that the intensity of the peak moment of LI can affect the sales performance of a video. Breaking from previous studies of investigating the influence of LI on an overall or general level (Hu and Chaudhry 2020; Xue et al. 2020), we highlight a novel perspective of the influence of LI. Additionally, event system theory—which proposes that peak events become salient when they are novel, disruptive, and critical—defines peak interaction in terms of not only its intensity, but also its duration and phases (Morgeson et al. 2015). Given that the influence of an event varies from moment to moment, researchers should further investigate the temporal features of events. Besides considering peak interaction’s intensity, we explore how long it should last to achieve satisfying sales outcomes. In considering peak interaction in different temporal phases, we focus on the key phases of a live streaming video, that of product display phases. In LSE, streamers adopt various visual skills (e.g., trial, zoom-in, and rotation) to display each product (Wongkitrungrueng et al. 2020). When a consumer watches the product displayed on the screen, that person is more likely to evaluate details and immerse themselves in their contexts (Sun et al. 2019). By considering both the intensity and temporal features of peak interaction, we aim to reveal the role of peak interaction over the short retention of product messaging in LSE.

2. Improving Sales Performance Through Live Interaction

LSE research has investigated how LI drives consumers to purchase (e.g., Kang et al. 2021; Xu et al. 2020; Xue et al. 2020). LI refers to the degree to which communication parties can act on each other, and to what extent the process is synchronized (Ou et al. 2014; Xue et al. 2020). Extant literature has shown the various benefits of LI, such as offering more product information, immersing consumers in visual and verbal demonstrations, increasing emotional connectedness, and facilitating the formation of associate relationships among streamer, viewer, and product/brand (e.g., Kang et al. 2021; Sun et al. 2019). As the current LSE research has treated
LI is enduring by investigating its effects at a general or average level (e.g., Xu et al. 2020; Xue et al. 2020), the lack of attention on its temporal aspects limits our understanding of the phenomenon. Because the influence of product display and promotion varies from moment to moment in a live streaming video (Kang et al. 2021; Wongkitrungrueng et al. 2020), it is therefore necessary to focus on LI from a short-term momentary perspective. We thus study the performance impact of peak interaction and its total time length based on gestalt characteristics and event system theories. Moreover, we study the intensity and duration of the peak moments of LI in the product display phases of a live streaming video.

3. The Intensity of Peak Interaction
According to gestalt characteristics theory and related peak-end rule (Geng et al. 2013), when individuals summarize their experienced event, they often focus on a few defining features (gestalt characteristics) of this event, such as the most intense moment (peak), the final moment (end), and the overall linear trend of the experience (slope), among which peaks have attracted increasing attention (Ariely and Carmon 2000). Many events exist such peak effect, such as when experiencing offensive or pleasant films, encountering annoying sounds or cold water, and receiving painful medical treatments. Responses to these experiences or stimuli are generally dominated by their extreme moments, causing them to be remembered more vividly and more accurately (Geng et al. 2013). Following gestalt characteristics theory, this study focuses on peak interaction, the moment when the intensity of interactivity is the highest.

Peak interaction is key to influencing the consumers’ response to a live streaming video. First, peak interaction is especially noticeable and salient to consumers. Gestalt characteristics theory suggests that the most dominant stimuli for individuals who experience an event (peak moments) cause better (more accurate) performance on their overall assessment of this event than general stimuli (Ariely and Carmon 2000; Jiang et al. 2019). Geng et al. (2013) also found that the peak moments can improve the “hedonic value” a customer receives based on the subjective experience of fun, playfulness, and enjoyment. In LSE, consumers often tune out unnecessary details and pay most attention to the most vivid and enjoyable stimuli, including the most frequent LI (Xu et al. 2018; Xue et al. 2020). Consumers thus notice the peak moments of LI above and beyond that in other periods. Second, peak interaction is most likely to guide consumers to develop positive inferences about streamers. Frequent LI contributes to consumers’ perception that streamers are friendly and warm, enhances their confidence that streamers are reliable and competent, and improves their positive emotions (Hu and Chaudhry 2020; Xue et al. 2020). When observing these peak interactions, consumers are more likely to make positive judgements of streamers, which can influence purchase decisions. We therefore hypothesize:

H1: The intensity of peak interaction during a video is positively related to sales performance.

4. The Duration of Peak Interaction
Drawing from event system theory, we concur that events related to peak interaction are bounded in time (Morgeson et al. 2015). Although gestalt characteristics theory also addresses the duration of an event, it suggests that individuals often use gestalt characteristics (notably, peak moments) to summarize the entire event and neglect how long it lasts (Ariely and Carmon 2000; Jiang et al. 2019). Integrating gestalt characteristics and event system theory, we argue that the duration of peak interaction in a live streaming video may also influence sales performance.
Event system theory proposes that the impact of an event may be increased with its duration (Morgeson et al. 2015). Because consumers give greater attention to peak interactions with a longer time length in LSE, longer LI among all parties at the peak level makes peak interaction more influential. This means that when consumers focus more on peak interaction, they are more likely to get caught up in the pleasure of LSE, and make favorable inferences about the streamers. In this way, consumers will make purchase decisions. However, beyond a certain duration, the longer peak interaction may hurt sales performance. First, prolonged peak interaction may distract consumers’ attention from the most important details. Although a brief moment of peak interaction helps consumers get quick responses from streamers, some viewers may be less likely to obtain useful information delivered from streamers when streamers focus too much on responding to comments from multiple consumers. Second, overly long peak interaction requires consumers to expend more cognitive effort to process diverse information, which can lead to negative reactions. The decision-making process of consumers entails their cognitive resources (Li and Karahanna 2015). If more participants begin to interact with each other simultaneously in a video, consumers will need to access the visual and auditory information from products, streamers, and peer consumers; as such, they will need to engage more cognitive resources when making their purchase decision. Since humans are characterized as cognitive “miser” who tend to simplify complex decisions (Fiske 1984), prolonged peak interaction may make consumers feel uncomfortable and resistant. We therefore hypothesize an inverted U-shaped relationship between the duration of peak interaction and sales performance.

H2: The duration of peak interaction during a video has an inverted U-shaped relationship with sales performance.

5. Product Display Phases for Peak Interaction

According to event system theory, when stimuli happen in different temporal phases of individuals’ cognition or experience, they may still have distinct evaluation even though the intensity of stimuli is at the same level (Morgeson et al. 2015). The literature on temporal phases mainly focuses on the beginning and ending phases of an event (Geng et al. 2013). In our context, however, the product display phases play a more salient role, because consumers come and leave at any moment, neither showing up at the beginning nor staying until the ending. When streamers bring together multiple products in a video and display them one-by-one on the screen, they often show the various physical angles of products and tactfully interact with a product (e.g., using touch, smell, or taste) (Wongkitrungruang et al. 2020). In turn, consumers can acquire more product knowledge, thus reducing perceived uncertainty while giving them a sense of immersion in the product experience (Sun et al. 2019). This evidence suggests that peak interaction and its duration at the product display phases of a video is more likely to shape consumers’ evaluation and impression about the products compared to non-product display phases. We therefore suggest that the proposed effects of the intensity and duration of peak interaction will be especially dominant during the product display phases of a video:

H3: (a) The positive performance impact of the intensity of peak interaction at product display phases is more effective than that at non-product phases; (b) The duration of peak interaction during product display phases has a more salient inverted U-shaped relationship with sales performance than that during non-product phases.
6. Proposed Research Methodology, Future Plan, and Intended Contribution

To examine the proposed hypotheses, data at the video level are gathered from Taobao Live, one of the largest LSE platforms in China. We collect preliminary data from the platform by selecting 508 third-party streamers who sell products of various online sellers, from October 2020 to January 2021. The collected data offer video files, sales performance data, and statistical information about streamers, videos, and products. We extract the variables according to the following operationalization. In subsequent work, we plan to propose and estimate an econometric model to quantify the peak interaction-performance relationships.

The operationalization of data are as follows. Peak interaction refers to the moment when the intensity of interactivity is the highest during a video, manifested as a timestamp when the number of comments reaches the maximum. It is measured by using moving window technology. We treat each second of a video as the beginning of a window, with each window lasting ten seconds, and calculate the number of comments for each window. This allows us to obtain a number for the interactivity intensity level for each window. For each video, we identify the maximum intensity moment within all windows of the video as the peak interaction and its number of comments as peak intensity. Peak interaction duration here refers to the total time length (in seconds) when the interactivity reaches a range of peak intensity in a video. We measure it by counting the number of windows where the number of comments is more than half the peak intensity. The analyses in a robustness check are then rerun, with the total length of time as the number of windows, where the number of comments is more than 80 percent of the peak intensity. The product display phases of peak interaction are measured in three steps. For each video, we first identify the moments in which products are displayed by using moving object detection techniques from the field of computer vision (Li et al. 2019). The video is then partitioned into two parts: product display and non-product display phases. Third, the peak intensity of LI and its duration are identified in each of the two phases. Sales performance is measured using the transaction revenue generated during a video. Streamer demographics (e.g., gender, total updates, and total followers), video attributes (e.g., views, the total number of comments, duration, average product price, total discounts) are incorporated as control variables.

This study intends to contribute to the e-commerce literature by examining the performance impact of the temporal aspects of LI during a live streaming video. Previous LSE studies have mostly regarded LI as an enduring construct by examining its overall influence. This study goes beyond that approach by exploring the impact of temporal features (i.e., intensity, duration, and phase) of peak interaction on sales performance. Specifically, we build on gestalt characteristics and event system theory to study the performance impact of the peak intensity of LI. Compared with the average or general level of individuals’ experience, the peak moment of that is found to be more influential in shaping their overall evaluation, and subsequent decision-making. Second, we add to the understanding of a short-term moment perspective of LI. According to event system theory that posits events with different duration and timing determine the different reactions of individuals, we suggest that temporal duration and product display phases of peak interaction also matter in improving sales performance. Third, we quantify these specific temporal dimensions using moving window and computer vision techniques, which has drawn very limited attention in the LSE literature. Fourth, from a practical view, streamers can draw inspiration from the dimensions of peak interaction and better compete by manipulating the interactivity of content toward improving sales performance.
7. References