International Journal of Research in Marketing 36 (2019) 117–136

Contents lists available at ScienceDirect

IJRM
International Journal of Research in Marketing

journal homepage: www.elsevier.com/locate/ijresmar

Full Length Article

Marketing analytics using anonymized and fragmented tracking data

Chinmay Kakatkar, Martin Spann *

Munich School of Management, Ludwig-Maximilians-Universität München, 80539 Munich, Germany

ARTICLE INFO

Article history:
First received on October 21, 2017 and was under review for 7 months
Available online 1 November 2018

Senior Editor: PK Kannan

Keywords:
Marketing analytics
Retailing
Data privacy
Data protection
GDPR
Network analysis
Big Data
Internet of Things

ABSTRACT

With the digitization of the retail industry, there is a growing abundance of event-based tracking data describing consumer behavior (e.g., online clickstreams and offline sensors tracking the movement of shoppers). However, stronger data privacy regulations and the growing privacy consciousness of consumers suggest that much of the data may increasingly only be available to retailers in an anonymized and fragmented form that does not identify individual consumers exactly. In response to the relative paucity of research on marketing analytics in retailing using anonymized and fragmented event-based (AFE) tracking data, this paper makes three interrelated contributions. First, we describe the relevance of AFE data in the future of retailing, contrasting it with other forms of aggregate and individual-level data. Second, we propose a methodology for analyzing AFE data, which allows us to approximately recover individual-level heterogeneity and derive meaningful variables from the raw data. Third, we validate the methodology using representative data collected by deploying sensor-enabled shelves in a field experiment within a store. We find that our approach to analyzing AFE data can help uncover interesting patterns of consumer behavior and could be applied across other online and offline retail settings in practice.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

The retail sector as we know it is going digital, and a number of retailers have begun to look at ways of exploiting the resulting trove of data to better understand consumer behavior. In particular, retailers increasingly have access to data tracking a consumer's actions as a stream of events over time, which can be a key enabler of marketing analytics (Wedel & Kannan, 2016). For example, online trackers can be used to gather clickstream data of an individual's web-browsing behavior, while offline trackers can take the form of sensor-enabled devices that can measure the movement and location of individuals via the mobile devices they carry around with them. However, stronger government regulations (e.g., the new European General Data Protection Regulation[1]) and greater consumer awareness of data privacy issues suggest that, in the coming years, a significant portion of the event-based tracking data available to retailers may be anonymized and fragmented across different data sources (Weber, 2015). The

* Corresponding author.
E-mail address: spann@spann.de. (M. Spann).

[1] The General Data Protection Regulation (GDPR) (Regulation (EU) 2016/679) came into effect on May 25, 2018 (see http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32016R0679).

https://doi.org/10.1016/j.ijresmar.2018.10.001
0167-8116/© 2018 Elsevier B.V. All rights reserved.

anonymization means that personally identifiable information (e.g., email addresses, geolocation) cannot be precisely uncovered from the data. Fragmentation means that there may be no consistent unique identifier (not even an anonymized one) that can help precisely tie data coming from one or more sources to a given person. Current trends suggest that such data will be on the rise (Storey, Reisman, Mayer, & Narayanan, 2017) — consumers will increasingly use tracker-blocking technology and be more selective about data sharing between devices, while retailers will have to adjust their data use policies to achieve regulatory compliance. Against this backdrop, we consider how retailers can make use of anonymized and fragmented event-based (AFE) tracking data[2] for marketing analytics.

Related literature in marketing has so far mainly considered aggregate and individual-level event-based data (Bradlow, Gangwar, Kopalle, & Voleti, 2017). Unlike aggregate data, individual-level data contains information about individual-level heterogeneity, which may be preferable for developing sophisticated models of consumer behavior. AFE data retains information about individual-level heterogeneity, making it different from aggregate data. Moreover, unlike individual-level data that researchers and practitioners have typically worked with in the past, the privacy-guarding nature of AFE data means that the individual-level heterogeneity is not directly accessible in the raw data.

AFE data is an emerging phenomenon that has received little attention in the extant literature. Thus, methodological work that addresses the challenges associated with the analysis of AFE data represents a fruitful area for marketing research. Such work would also be in alignment with recent calls for new research in marketing analytics. For example, Wedel and Kannan (2016) call for more work on diagnostic and predictive methods to support data-driven marketing decisions, especially in cases where the raw data may be problematic in some respects (such as tracking data that is anonymized and fragmented due to the imposition of legal or technical constraints). In response, this paper makes three interrelated contributions. First, we describe the relevance of AFE data in the future of retailing, contrasting it with conventional forms of aggregate and individual-level data. Second, we propose a methodology for analyzing AFE data, which allows us to approximately recover individual-level heterogeneity and derive meaningful variables from the raw data. Third, we collect representative data generated by deploying sensor-enabled shelves in a randomized in-store field experiment to validate the methodology. We find that our approach to analyzing AFE data can help uncover interesting patterns of consumer behavior and could be applied in online and offline retail settings.

2. Relevance of anonymized and fragmented event-based data in retailing

Event-based data in retailing typically takes the form of a sequence of observable actions (captured as "events") taken by one or more consumers in a given context over a period of time (Bradlow et al., 2017; Sismeiro & Bucklin, 2004). Each event is timestamped, making it possible to order events chronologically.[3] For instance, the web-browsing actions of a consumer in an online setting can be captured as events that track the opening of new webpages and links. Similarly, the physical location of a consumer, as well as her movements within and between stores, can be represented as a chain of events (e.g., shop visits, product interactions) in an offline setting.

However, the events that form AFE data are constrained by the fact that they do not directly track the identity of the individual being monitored, and that two such events cannot be tied to each other based on some unique identifier (e.g., a customer ID).

In the coming years, the privacy challenge of handling digital consumer data in a fair and secure manner will become increasingly important, and place a greater spotlight on AFE data. Government regulations are beginning to play a key role in forcing companies to adopt privacy-oriented data use policies. While the European Union currently appears to be leading the way, countries such as the USA, Canada, South Korea and Singapore are also actively taking steps to strengthen the privacy regulations in their respective jurisdictions (Maras, 2015; Weber, 2015). In general, the privacy challenge affects not only the retailer itself but also the technology providers that supply the software/hardware infrastructure for tracking consumer activity across online and offline channels and the application developers that build marketing analytics programs. The vast scope of the privacy challenge is thus a major reason for expecting a significant growth in AFE data in the coming years.

AFE data can generally arise in online and offline retail settings, resulting from flows of information related to consumers, products, channels through which consumers can access products, location and time (Bradlow et al., 2017; Burke, 2010). In this paper, we mainly consider the types of AFE data that may be generated by consumers in online and offline retail environments.

In the online setting, retailers may anonymize IP addresses or similar identity-revealing data, and prevent the sharing of unique identifiers across services to preempt data across different touchpoints from being tied to an individual; such measures could not only be taken to comply with government regulations, but also to improve consumer perception, unless the collection and use of such data is expressly permitted by the consumer (e.g., as part of a loyalty program). Meanwhile, consumers may choose to take privacy matters into their own hands by blocking online trackers from harvesting data in the background. Privacy-oriented web

[2] We henceforth refer to this data as simply "AFE data" for the sake of brevity.

[3] Additional technical characteristics of AFE data are as follows: (1) *atomicity*, i.e., an event cannot be broken down further into smaller events, (2) *immutability*, i.e., a once generated event is not overwritten at a later point in time but each new piece of information is codified as a separate, new event, and (3) *loose structure*, i.e., an event is described by a vector of variables recorded by the tracking technology (e.g., timestamp, location, duration), but the raw variables may only be loosely related to each other.

browsers, such as DuckDuckGo, are increasingly able to rival the search result quality of established giants like Google, while browser plugins such as AdBlocker and Ghostery give consumers a finer level of control over the specific online trackers they wish to share their browsing metadata with (Merzdovnik et al., 2017). The use of Virtual Private Networks (VPNs) can also provide strong protection against trackers by rerouting the online traffic of a user through a global network of servers, making it essentially impossible to trace the online activity back to the user's actual IP address. Extensive surveys carried out across several countries reveal the growing popularity of privacy-oriented online experiences, with as many as 26% of the respondents reporting the use of tracker-blocking tools in some instances (Storey et al., 2017).

In the offline setting, AFE data can broadly be constructed from the sequence of actions that a consumer takes in the physical environment (Hui, Fader, & Bradlow, 2009a). Faced with government regulations and a strong preference for privacy from consumers, retailers can deploy physical sensors to collect data on the movement patterns of consumers, while taking steps to protect the privacy of individuals (Inman & Nikolova, 2017). For instance, AFE data can be collected by sensors dispersed in a store or a shopping mall to unobtrusively track the movement of consumers within and between stores. The physical movement can be represented as a network in which the nodes are the data collection artifacts (e.g., sensor-enabled doors and shelves), and an edge between two nodes reflects the movement of a consumer between the corresponding artifacts. Note that, since we are concerned with AFE data, it would not be possible to precisely tie data collected at the same artifact (e.g., at different times in the day or week) or at different artifacts to an individual consumer.

Moreover, AFE data is different from aggregate data and individual-level data that tracks individuals. Aggregate data has played an important role in marketing research so far, mainly due to the wide availability of such data and its usefulness in deriving inferences about consumer segments (Ilfeld & Winer, 2002; Tellis, 2004). However, since aggregate data, by definition, does not preserve individual-level heterogeneity, it limits our ability to develop high-resolution, granular models of consumer behavior. To develop such models, marketing researchers have used individual-level data that explicitly tracks individuals via some unique identifier. Studies about online consumer behavior have relied on tracking IP addresses or customer IDs to map individuals to online activity (Bucklin & Sismeiro, 2009; Montgomery, Li, Srinivasan, & Liechty, 2004; Sherman & Deighton, 2001; Sismeiro & Bucklin, 2004). The literature looking at individual-level behavior in offline retailing has used video data and sensors tagged to individuals and/or their shopping trolleys (Hui, Fader, & Bradlow, 2009b; Hui, Huang, Suher, & Inman, 2013a; Hui, Inman, Huang, & Suher, 2013b; Larson, Bradlow, & Fader, 2005), or some form of mobile phone tracking using GPS information or near-field sensing (Luo, Andrews, Fang, & Phang, 2014; Molitor, Reichhart, Spann, & Ghose, 2016; Phua, Page, & Bogomolova, 2015), to effectively follow the individual within and between stores. However, the tightening data privacy regulations and the ability of consumers to block trackers may serve to limit retailers and researchers in their ability to track individuals explicitly. In this context, the use of AFE data represents a promising and increasingly relevant approach to overcoming the privacy challenge. Unlike aggregate data, AFE data does contain information about individual-level heterogeneity, making it possible to build more granular models of consumer behavior. However, since the individual-level information is hidden in the raw AFE data, the privacy of consumers is preserved.

The methodology we describe in the following section proposes one approach to recovering the individual-level information and deriving meaningful variables from AFE data to facilitate the analysis of consumer behavior in retailing.

3. Methodology

3.1. Notation and overview

First, we formalize the notion of events in AFE data and the artifacts in the environment that are responsible for triggering the events. Let E denote a finite set of event types $\{e_1, e_2, \ldots\}$. For example, a sensor-enabled door could trigger the events "person entered store" and "person left store", while a sensor-enabled shelf could also track whether the displayed product has been "picked up" or "put down". Being anonymized means that these events do not explicitly refer to any individual via a unique identifier, and any two events cannot be precisely tied to a particular individual. Arbitrarily long event streams consisting of the event types defined in E may be emitted by a physical sensor in a store (Guralnik & Srivastava, 1999; Inman & Nikolova, 2017), or an online analytics system that tracks the clickstreams of web users (Bucklin & Sismeiro, 2009). Next, let S denote a finite set of artifacts $\{s_1, s_2, \ldots\}$. Artifacts can be thought of as those entities in the environment that are able to emit events in response to the consumer's behavior. Sensor-enabled shelves in a store or shop doors mounted with visitor counters are examples of artifacts.

The above formalization allows us to further highlight the novelty of AFE data with respect to aggregate data and data that tracks the identity of individuals. Aggregate data would not let us see the actual event stream E, but rather the result of some aggregating function $f(\cdot)$ applied to one or more event streams. For example, the result of the aggregating function could be a single value (e.g., the total number of events in a given event stream or the average number of events across multiple event streams), or the result could be a set of related values (e.g., a vector describing the frequency distribution of the different events types $e_1, \ldots, e_n$ in a given event stream). Thus, aggregate data essentially ends up destroying information about individual-level heterogeneity (i.e., the information about the identities of the different consumers that may have been responsible for triggering the events in the event stream). Meanwhile, individual data that tracks individual identities alleviates the disadvantage of aggregate data by giving us not one but k event streams, one for each of the individuals being tracked. While the ability to access the information on individual-level heterogeneity for each of the k consumers is clearly attractive in terms of building high-resolution, individual-level behavioral models, it comes at the expense of the consumers' privacy. By contrast, AFE data protects the privacy of consumers by not directly encoding their identity in the data, but it leaves open the possibility of approximately recovering the individual-level heterogeneity via statistical and/or computational means.

In the following, we describe a three-step methodology for analyzing AFE data. The first step discusses different heuristic-based solution approaches to the problem of approximately recovering individual-level information from the raw event stream. Building on this, the second step adopts a network perspective of relationships between events and artifacts to elegantly derive contextually meaningful outcome and predictor variables, respectively. Finally, the third step links these outcome and predictor variables to facilitate the construction of high-resolution, individual-level models that describe consumer behavior in the retail setting. For the purpose of exposition, we initially consider the simple case of a single consumer interacting with a single artifact at any given point in time. We then discuss how our method can be extended to allow the analysis of consumer behavior in more complex cases involving multiple consumers and artifacts.

3.2. Identifying individuals from event streams

The first challenge is to approximately identify individuals in the event streams that constitute AFE data. Each observation in the dataset is timestamped, describing the event and the artifact that triggered it. In the simple case of a single consumer (regardless of how many artifacts there are), we can trivially identify the individual at all times. However, as soon as we allow for multiple consumers, it is no longer obvious from the anonymized and fragmented raw data whether two adjacent events $[\ldots, e_i, e_j, \ldots]$ were triggered by the same consumer or by two different ones. For example, an online clickstream might show an event describing someone searching for a product on an e-commerce website, and a following event that describes the same product being purchased. Similarly, a physical sensor may detect multiple instances of consumers visiting a particular product display in an offline store. In both of these examples, we cannot link the two events to the same consumer (despite them occurring next to each other in the anonymized event stream) without a unique identifier for the consumer. To facilitate individual-level analysis, we need to split up a given event sequence into several slices. Each slice should contain the events triggered by a unique individual, thus serving as an approximate identifier of consumers.

3.2.1. Slicing dimensions

The extant literature on information retrieval algorithms and data mining suggests that there are three basic heuristics for slicing an event stream in AFE data: temporal, spatial and combined/other (Kumar, Mahadevan, & Sivakumar, 2004; Nagarajan et al., 2009; Yang, Pierce, & Carbonell, 1998). From a theoretical perspective, each of these types of heuristics uses different assumptions about the relationship between a slicing dimension (e.g., time, space, or a combination thereof) and the hidden individual-level heterogeneity in the AFE data.

The temporal slicing heuristic assumes that events occurring in the same time interval belong to the same individual. The size of the time interval is an important slicing parameter, since a slice that is too large or too small may produce an inaccurate mapping of events to individuals. Slicing by time interval has also been used to segment online browsing sessions of users (Montgomery et al., 2004). Temporal slicing is theoretically justified in situations where consumer behavior is synchronous with respect to the retail environment, so that only one customer is able (or likely) to trigger events within the given timeframe. In practice, synchronous behavior is evident in queue-based systems (e.g., a check-out line in a store, speaking with a customer support representative) or systems with relatively low foot traffic, such that only one consumer is likely to be active in the retail setting at any given point in time. In spatial slicing, events that occur at artifacts located close together are grouped into a slice. Theoretically, spatial slicing is justified when the actions occurring at different locations are likely to be triggered by different consumers. Spatial slicing is therefore applicable in situations where there are several artifacts that are spaced far apart. For example, events generated by artifacts in different physical stores (e.g., sensor-enabled product displays), or at different aisles within a store, are unlikely to be triggered by the same consumers, and such events can thus be sliced along the spatial dimension. Finally, the slicing heuristic might combine temporal and spatial elements and leverage additional situational data (Nagarajan et al., 2009), which may involve augmenting the event data with information from related datasets.

3.2.2. Unguided or guided slicing

To slice the data along a given dimension, we can opt for either unguided or guided slicing (Kimball & Ross, 2009). For instance, let $[e_1, e_2, e_3, e_4, e_1, e_2, e_1, e_2, e_3, e_2, e_3, e_4]$ represent the event stream of a consumer's activity in a store, such as approaching a smart shelf (e_1), picking up (e_2) and putting down (e_3) the displayed product, and then leaving the shelf (e_4). Furthermore, suppose that we wish to slice the data along the temporal dimension.

The unguided approach to temporal slicing would rest on the assumption that the consumer's activity is linear in time, i.e., that the time between adjacent events is always the same. Given a stream of n events, we would slice the stream into k pieces with $\lfloor n/k \rfloor$ events per slice. For example, if $k = 3$, then the above event stream would be sliced into $[[e_1, e_2, e_3, e_4], [e_1, e_2, e_1, e_2], [e_3, e_2, e_3, e_4]]$. Since the length of each slice is fixed, unguided slicing is relatively easy to implement and the interpretation of statistics at the level of slices (e.g., the average occurrence of an event per slice) is also fairly straightforward. However, the linearity assumption underlying the fixed slice sizes in the unguided approach may not always be satisfied by the actual frequency of the event occurrences over time. From a theoretical perspective, the unguided approach is only justified in situations where each consumer is guaranteed — or at least likely — to trigger the same number of events. In practice, such situations may arise in a quota-based system that limits the number of actions per consumer. Applying unguided slicing in an arbitrary fashion may not yield an accurate mapping between the resulting event slices and the actual set of unobserved individuals we intend to recover.

By contrast, guided slicing does not assume fixed slice sizes and seeks instead to leverage outside information based on the retailer's past experience or preliminary analyses (qualitative and quantitative) of consumers in the store setting. For example, based on descriptive analyses of shopper behavior, a retailer may find that an average consumer spends approximately 1 minute at a product display. Using this valuable information as a guide, the event stream can be sliced into one-minute intervals. As such, guided slicing might split the above event stream into $[[e_1,e_2,e_3,e_4],[e_1,e_2],[e_1,e_2,e_3,e_2,e_3,e_4]]$, for instance. Notice that guided slicing may not necessarily yield evenly sized slices, which makes the interpretation of slice-level statistics more difficult. For example, with guided slicing, the event e_2 makes up 50% and 33% of the events in the second and third slices of our event stream, respectively, but occurs more often in the third slice. Similar statistics are easier to interpret with the result of unguided slicing in our example, since the fixed slice sizes provide a common denominator.

In general, to address the challenge of approximately recovering individuals in AFE data, guided slicing is arguably preferable where possible since it can potentially make use of deeper contextual information about the retail setting (e.g., characteristics of the retailer and the consumer base) to yield a more accurate recovery of individuals than the unguided approach would allow.

3.3. Deriving contextually meaningful variables

The second challenge is to derive contextually meaningful outcome variables and corresponding predictors from AFE data. We tackle this challenge by taking a network perspective. In particular, we propose the conceptualization of two types of networks — a network of events and a related network of artifacts. Examples of these two networks are depicted in Fig. 1(a) and (b).

The event network, as represented in Fig. 1(a), is often referred to as a "transition diagram" in engineering disciplines (Wang, 2013, pp. 4–5). In our methodology, a separate event network is constructed for each slice (guided or unguided) of the event stream. The set of nodes in the event network consists of all unique events seen in the slice; a directed edge between two nodes represents a permissible event transition. For example, the network in Fig. 1(a) might have been constructed from the slice $[e_1,e_2,e_3,e_3,e_2,e_1]$. Moreover, event streams in a retail setting must start with a plausible initial event (e.g., a consumer entering a store), and this corresponds to the entry point into the event network. In Fig. 1(a), the initial event happens to be e_1. An event transition matrix can now be computed, in which the element in the j-th row and the i-th column of the matrix captures the frequency (or proportion) of transitions from event e_i to event e_j in the slice. The event network then emerges naturally from the transition matrix — the nodes represent the event types and the edges denote event transitions. Note that while clickstream research has previously used event transition matrices in Markovian analyses (Montgomery, Li, Srinivasan, & Liechty, 2004), we use such matrices to construct contextually meaningful outcome variables at the individual level, as discussed later.

Whereas an event network is produced from the sliced data, the artifact network considers the event stream in its entirety. Two artifacts (the nodes) are linked in the artifact network if an event at one artifact is followed by an event at the other artifact in the aggregate event stream. For example, the artifact network in Fig. 1(b) depicts five artifacts, which may be different webpages of an e-commerce website or sensor-enabled shelves in an offline store. The edge between artifacts s_1 and s_2 suggests that an event at artifact s_1 was followed by an event at artifact s_2; in fact, the bidirectional arrows used for the edges in Fig. 1(b) indicate that an event in the other direction (from artifact s_2 and s_1) also occurred in the aggregate event stream. Notice that the existence of an artifact network with more than one node assumes a retail setting with more than one artifact but makes no claim about the number of consumers. The edges may correspond to the actions of a single consumer triggering events at each artifact

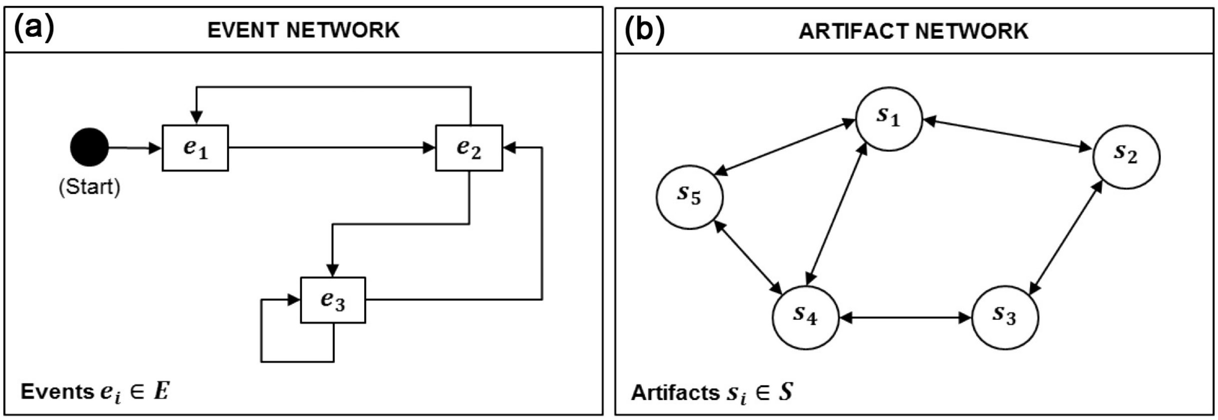

Fig. 1. Examples of network representations of AFE data.

sequentially, or many consumers triggering events at different artifacts simultaneously. Distinguishing between the cases of single and many consumers requires an understanding of the constraints imposed on the actions of individual consumers in the given retail context. For instance, it may be unlikely for events occurring in a very short timeframe at two ends of a store to be triggered by the same consumer.

Our decision to adopt a network view of AFE data is based on a number of reasons. First, the relational nature of event transitions and an individual's movement between artifacts lends itself naturally to a network-based conceptualization. Hui et al. (2009a) previously demonstrated the value of a network-based view in terms of its wide applicability to several instances of event-based data in marketing (e.g., data on online browsing behavior, sensor-based data of consumers' movements in a store, eye-tracking data). Second, casting the event data as a web of relationships between events/artifacts creates opportunities to transfer insights from the growing literature on applied network analysis to generate outcome variables and associated predictors. In particular, network thinking allows us to enrich our conceptualization of the retail setting by considering not just the entities themselves (e.g., events, artifacts, consumers) but also the relationships and interactions between them (Borgatti, Mehra, Brass, & Labianca, 2009). Third, networks in marketing science have so far typically been used to analyze relationships between fairly high-level representations of nodes, such as networks of consumers and firms (Borgatti et al., 2009; Newman, 2010; Wasserman & Faust, 1994). The application of network thinking that leverages a more granular set of nodes — as in the case of product networks (Dhar, Geva, Oestreicher-Singer, & Sundararajan, 2014) — is still a fairly nascent area in marketing research. In this context, adopting a network perspective allows us to test the value of network methods in the analysis of AFE data, thereby extending the literature on several related and emerging topics (e.g., clickstream analysis, path data analysis of shopper movement between and within stores, knowledge representation of consumer behavior in the Internet of Things). As we demonstrate in the following subsections, the event network can be especially useful for deriving individual-level outcome variables, while the artifact network can yield interesting situational predictors. Note that, although we focus on the network perspective in this paper, we also touch on interesting alternative representations of the event data in Section 5 in terms of avenues for future research.

3.3.1. Outcome variables

In general, we can derive two types of individual-level outcome variables from the information contained in the event network pertaining to AFE data. Essentially, the first type of outcome variables corresponds to the raw event itself, while the second type results from a combination of different events. Whether a raw event is meaningful as an outcome (e.g., in terms of its frequency of occurrence, duration or other properties) depends on various aspects of the data collection setup. For instance, consider an online tracker embedding in an e-commerce website that triggers a "purchase" event each time a consumer executes a purchase order, or a system in a physical store that triggers a "coupon used" event whenever a shopper uses a promotional coupon at the checkout counter. Such raw events may indeed be worth analyzing as an outcome variable in its own right, since the quantity being measured is directly relevant to contextually meaningful concepts such as purchase behavior and promotional effectiveness.

The need to combine raw events arises from the fact that most events that constitute AFE data tend to be too granular to be meaningful on their own. For example, a sensor-based in-store shopper tracking system might trigger events such as "shopper visits shelf", and "shopper picks up product". Considered in isolation, each of these atomic events provides an incomplete and fragmented view of consumer behavior; this is especially an issue in the offline setting, where several sensors may be collecting different pieces of information about the consumer (e.g., movement, interaction with products, purchases). To construct meaningful outcome variables, we can combine raw events by considering the transitions between these events in more detail. Building on the work by Montgomery et al. (2004), we can construct a matrix M that describes the conditional probability of each event transition in a given context. The element m_{ji} in M denotes $P(e_j|e_i)$, the probability of event e_j following e_i in the event stream. The conditional probability for each pair of adjacent events can be derived by taking a frequentist view of the event data. In general, an event stream of n raw events will yield $n-1$ adjacent event pairs.

For example, consider the event stream slice $[e_1,e_2,e_3,e_4,e_1,e_2,e_1,e_2,e_3,e_2,e_3,e_4]$, consisting of 12 raw events that can be grouped into the following 11 adjacent event pairs: (e_1,e_2), (e_2,e_3), (e_3,e_4), (e_4,e_1), (e_1,e_2), (e_2,e_1), (e_1,e_2), (e_2,e_3), (e_3,e_2), (e_2,e_3) and (e_3,e_4). Counting the number of occurrences of each observed transition (e_i,e_j) in the event stream yields the frequency matrix F of j rows and i columns,

$$F = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 3 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \end{pmatrix},$$

and the corresponding transition probability matrix,

$$M = \begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 3/4 & 0 & 1/4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

$P(e_j|e_i)$ in the transition matrix is the probability of seeing event e_j immediately after e_i in the event stream. For instance, the probability of seeing e_2 after e_3 in the event stream is $P(e_2|e_3) = 1/4$. If the set E of permissible events is sufficiently small and the event stream registers at least one occurrence of each $e_j \in E$ following some other event, then it is intuitive to see that the frequency matrix F will not be too sparse, so that each row of M will sum to 1.

Crucially, we can now combine the probabilities in M to derive composite conditions that may represent contextually relevant outcome variables. The type of composite conditions we consider in this paper are of the form $\sum_{j=1}^{k}\sum_{i=1}^{k}\omega_{ji}P(e_j|e_i)>\theta$, where the event set E consists of k unique event types, $\omega_{ji} \in [-1,0,1]$ is a weighting that essentially serves the purpose of including/excluding certain probabilities $P(e_j|e_i)$ from the condition (although we could let $\omega_{ji} \in \mathbb{R}$, in general), and θ represents a threshold value above which the condition is true. In a retail setting, a condition could be used to model a concept, such as the commercial value of facilitating product interactions in a store (e.g., by displaying a product prominently on a shelf). This can be operationalized by a simple outcome variable that tests whether the probability of a purchase (e_2) following a product interaction (e_1) is greater than zero, i.e., $P(e_2|e_1) > 0$. A more complex composite condition, such as $P(e_2|e_1) - P(e_2|e_3) > 0$, can be used to model a stricter outcome variable, which asserts that the transition from e_1 to e_2 has a higher probability of occurring than the transition from e_3 to e_2 (a purchase following, for example, an interaction with the cashier). While the simple condition models the commercial value of product interactions in isolation, the complex condition lets us compare the value of product interactions to the value of other events that may precede a purchase. In general, composite conditions offer a parsimonious yet flexible way of constructing outcome variables from event networks. To our knowledge, our demonstration in this paper also represents the first such application of composite conditions for analyzing AFE data in a contextually meaningful manner.

3.3.2. Predictors

Predictors that explain the variations in the outcome variables derived from AFE data may be artifact-independent or artifact-dependent. Predictors that exist independently of artifacts may include variables that are temporal (e.g., time of day, day of the week) or situational (e.g., the number and nature of sales assistants in a store, promotional offers and loyalty programs). Any number of these variables may potentially be good predictors of a given outcome variable and should thus be incorporated in the analysis where appropriate. Meanwhile, artifact-dependent predictors are those variables that are derived from the artifacts themselves. Such variables are fixed for a given artifact and may be time-invariant; therefore, they can be loosely thought of as artifact-specific effects (Wooldridge, 2013, pp. 484–485). For example, an in-store artifact may consist of variables such as the attributes of a smart shelf, the shelf's location on the shop floor, and the attributes of the product displayed on the shelf. Each of these variables can take on different values for each artifact, depending on the decisions made by the store manager (e.g., whether to change the displayed product or shelf location).

We can also leverage the relational view of artifacts, as shown in Fig. 1(b), to derive network measures that can serve as artifact-dependent predictors. The network itself can be described by a weighted matrix W, such that the element W_{ij} represents the strength of the edge between nodes i and j (e.g., frequency of co-occurrences in the case of event data), or an adjacency matrix A_{ij} that denotes a non-weighted version of W_{ij}, such that the element $A_{ij} = 1$ if nodes i and j are connected, and 0 otherwise. Table 1 provides an overview of key, artifact-level network measures that we will use in the empirical study in Section 4.

There are several benefits to using network measures as predictors of outcome variables in the context of AFE data. Fundamentally, in contrast to other predictors that are based on information about the environment or a particular artifact, network measures capture the heterogeneity that arises from relationships between artifacts. Since the nodes in the artifact network have no agency themselves, any relationships between the nodes are determined by the consumers that conceptually exist outside the artifact network (Dhar et al., 2014).

Measures of network centrality may reveal the popularity or influence of certain artifacts in a consumer's movement pattern. In particular, since the weighted degree measure only considers the relationships between the focal node and its first-degree neighbors, weighted degree centrality can serve as a "local" measure of the amount of consumer activity around a given artifact (e.g., a shelf in a store). Meanwhile, the recursive definition of eigenvector centrality effectively considers the focal node's relationship to all nodes in the network and can be a "global" measure of the importance of artifacts (e.g., whether a consumer

Table 1
Definitions of network measures.

Network measure of artifact i	Formal definitions
Weighed degree centrality $C_{Degree}{}^{W}$	$C_{Degree}{}^{W}(i) = \sum_j^{\lvert V\rvert} W_{ij}$, i.e., the degree of the node is weighted by the strengths of the ties with its immediate neighbors.
Eigenvector centrality $C_{Eigenvector}$	$C_{Eigenvector}(i) = \frac{1}{\lambda}\sum_j^{\lvert V\rvert} A_{ij} C_{Eigenvector}(j)$, where λ is the largest eigenvalue of A_{ij}. The eigenvector approach indirectly considers the importance of all nodes in the network.
Betweenness centrality $C_{Betweenness}$	$C_{Betweenness}(i) = \frac{g_{jk}(i)}{g_{jk}}$, where g_{jk} denotes the number of shortest paths between nodes j and k, and $g_{jk}(i)$ of these paths go via node i.
Modularity class c	The modularity of a network partition is given by $Q = \frac{1}{2m}\sum_{i,j}\left[W_{ij} - \frac{k_i k_j}{2m}\right]\delta(c_i, c_j)$, where $k_i = \sum_j W_{ij}$. c_i is the modularity class (or network community/cluster) to which node i is assigned. $\delta(c_i, c_j) = 1$ if $c_i = c_j$, and 0 otherwise. $m = \frac{1}{2}\sum_{ij} W_{ij}$.

typically arrives at a popular shelf in a store after having visited less popular shelves). The betweenness centrality measure complements the above insights by highlighting "bridge-builders" that connect different clusters of the network. For instance, an artifact with high betweenness may be an information board in a store that is consulted between visits to other product shelves.

Moreover, to gain an understating of the structure of the entire network, the modularity measure classifies the artifact network into clusters of nodes that are more densely connected together than with the rest of the network (Blondel, Guillaume, Lambiotte, & Lefebvre, 2008). Artifacts that fall in the same modularity class may be similar to each other in some respect, although this may not be obvious without the network data. For example, we may find that shelves with products that are part of a consumer's shopping list are part of the same modularity class.

Crucially, although AFE data prevents us from directly tracking an individual consumer or tying data from different artifacts to the same individual, the network-based data of the consumer's movement between artifacts can still reveal interesting insights about her behavior in a retail setting. Since the network measures also tend to be continuous variables, they allow for a high-resolution modeling of a consumer's interactions with the retail setting.

3.4. Building models of consumer behavior

Having constructed outcome variables and predictors, we are in a position to link these together. Notice that our approach of slicing the raw data lets us approximately track the interactions of individuals over the different artifacts in the retail environment. An elegant consequence of this data setup is that we can now reframe the information contained in the events as panel data in order to estimate event-based models of consumer behavior.

Suppose that our retail context consists of S artifacts, and we temporally slice the corresponding event stream into T time intervals, with each such interval approximately representing an individual consumer. Then, we have an $S \times T$ panel of artifacts observed over the set of all consumers that appear in the event data. Given an outcome variable y (e.g., whether a consumer interacts with a product), we can construct the linear model $y_{st} = \alpha + \beta X_s + \gamma Z_{st} + \varepsilon$ for $n = 1, \ldots, S$ and $t = 1, \ldots, T$, where y_{st} is the outcome variable observed for consumer t at artifact s, α is the average effect, X is a vector of artifact-dependent predictors (e.g., the location of an artifact, network measures), Z is the vector of artifact-independent predictors (e.g., time of day), and ε is the composite error term. Now, if the consumer-specific predictors are uncorrelated with the artifact-specific ones, then we can estimate the outcome variable using a random effects (RE) or a fixed effects (FE) model (Wooldridge, 2013, pp. 492–493).[4]

In general, by treating the consumer-artifact interactions as panel data and estimating the outcome using a panel model, we are able to control for unobserved heterogeneity. If the outcome variable is binary or categorical, we would treat it as count data using a suitable classification model (e.g., a Logit or Negative Binomial), whereas we would apply a regression model (e.g., ordered Probit or OLS) to an ordinal or continuous outcome variable. Other non-linear algorithms, such as decision trees and random forests, could also be applied (Varian, 2014). Moreover, assessed from the perspective of panel data analysis, it is clear that the slicing heuristic should afford sufficient variation in the consumer-artifact interaction. For instance, opting for temporal instead of spatial slicing may produce more slices, thus yielding a richer panel dataset that could approximate the between-consumer heterogeneity more accurately. Similarly, by finding the right balance in the quantity and quality of the artifacts derived from the data, the resulting panel interpretation can enable a nuanced analysis that accounts for within-consumer variation.

Finally, our methodology can also accommodate the presence of multiple individuals interacting with the same artifact at the same time. The simultaneous presence of more than one consumer at an artifact would be reflected in the event transition matrix M, such that transitions previously undefined in the event network would now receive non-zero probabilities of occurring. For example, suppose the artifact is a sensor-enabled shelf in a store that tracks customers arriving at — and leaving — the shelf. Now, if the shelf were to sense an individual arriving twice without leaving in-between, this would indicate the presence of two different individuals standing at the shelf together, and might conceptually reflect crowding or social interaction between consumers (Argo, Dahl, & Morales, 2008; Harrell, Hutt, & Anderson, 1980). An implication for our methodology is that allowing hitherto undefined event transitions can yield additional outcome variables related to social concepts, depending on the empirical context.

Fig. 2 provides a visual summary of the three steps of our proposed methodology, highlighting their flexible and modular nature.

4. Empirical study

4.1. Data

To validate our proposed methodology, we collected AFE data on consumer behavior in a store during a week-long, randomized field study. The data captures interactions of consumers with products displayed on interactive, sensor-enabled display shelves.

[4] The related specification test proposed by Hausman (1978) can be used to gauge the appropriateness of using the RE model as opposed to the fixed effects (FE) model. Unlike the FE model, the RE model would assume that the unobserved heterogeneity is uncorrelated with the consumer-specific predictors and, thus, includes it in the composite error term ε.

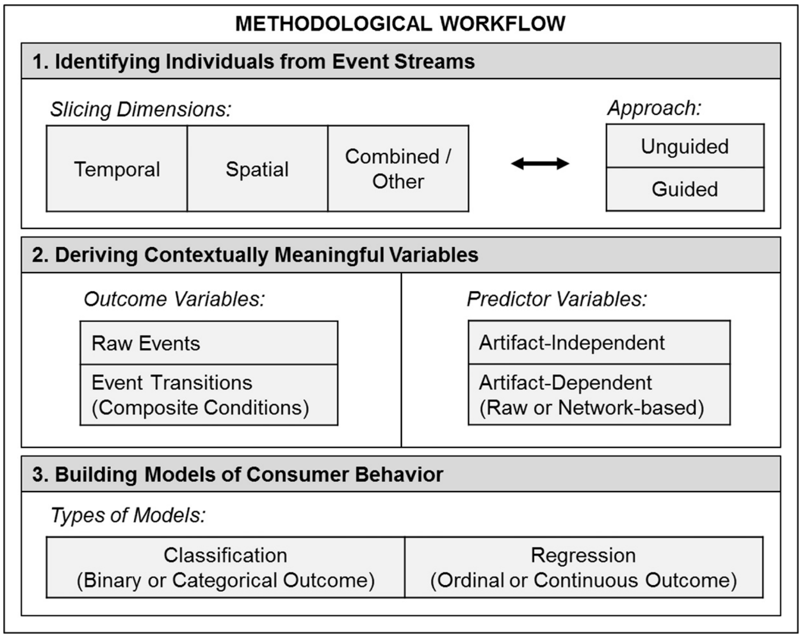

Fig. 2. Overview of methodological workflow.

The study took place in the merchandise store of a large European university at the beginning of the academic year. The store is open from Monday to Friday, and as expected, virtually all of the store's customers are affiliated with the university. The store sells a typical assortment of memorabilia (e.g., stationary, clothing, collectibles). We chose a university sweatshirt and a souvenir mug to represent the store's two main product categories, apparel and collectibles, respectively. According to sales data shared with us by the store manager, the chosen products were popular enough to ensure the generation of sufficient sensor-based data from consumer-product interactions during the study.

We partnered with a technology company to procure four identical prototypes of sensor-enabled display shelves that could be deployed in-store. The shelves were shaped as cubes of approximately 20 cm (8 in.) in length, height and width; this shelf size was large enough to give us flexibility in choosing different products to put on display, while being small enough to allow easy transport and handling within the store. Each sensor-enabled shelf was equipped with a distance and pressure sensor. The distance sensor would trigger an event whenever it detected the arrival or departure of a customer within approximately 50 cm (20 in.) of the shelf. The pressure sensor triggered an event whenever the product placed on the shelf was picked up or put down. The dataset obtained from each sensor-enabled shelf consisted of the timestamp, shelf number and event type per observation. Crucially, the event-based data generated by each sensor-enabled shelf was anonymized (no unique identifiers to link events to specific customers) and fragmented (no way of precisely tying data from different shelves to a single customer).

Upon detecting a person nearby, the distance sensor could be programmed to respond in varying degrees of complexity (e.g., lighting up in a certain color, or running a product-relevant ad on a nearby in-store TV). In our study, whenever the distance sensor was triggered, the front face of the shelf would light up in a mellow blue color. We chose this blue in an effort to minimize the confounding effect of color on the effect of the sensor-based nature of the stimulus. Past research on color and marketing suggests that the mellow blue would be noticeable enough to draw the attention of nearby shoppers, but not by itself induce any further consumer-product interaction (Singh, 2006); a qualitative pre-test appeared to confirm our expectation. The lighting functionality thus served as a simple stimulus to attract the attention of a customer standing near the shelf.

Two units of each product were placed on four sensor-enabled shelves (one unit per shelf) near the center of the store, as illustrated in Appendix A. The prominent location of the shelves ensured that they received high foot traffic during the study, which helped generate a large enough sample of product pick-ups to allow meaningful statistical analyses. To allow between-shelf identification, we randomized the product-shelf combinations every day before the store's opening hours in two ways. First, the lighting functionality of two randomly selected shelves was always kept disabled; this allowed us to separate the effects of the product type (sweatshirt or mug) and the display interactivity triggered by the sensor (light on or off). Second, the order of the shelves was randomly shuffled, allowing us to account for the effect of shelf location on the product-customer interaction. Reverse causality was not a problem in our study since the sensor-based stimulus of the shelves inherently preceded the customer's interaction with the product.

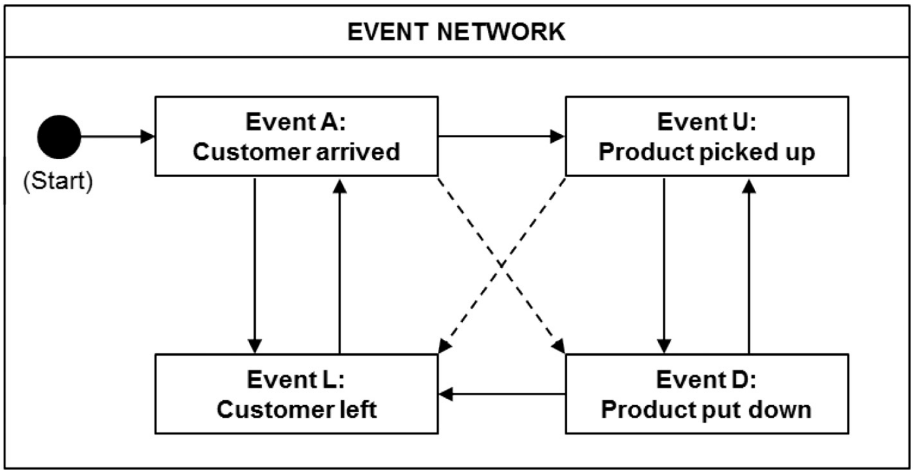

Fig. 3. Event network of data collected by sensor-enabled shelves.

In this context, we can define an artifact as a unique combination of shelf interactivity (light on/off), shelf location (four locations with respect to the checkout counter), and the type of displayed product (sweatshirt or mug). Thus, the in-store experiment yields $2 \times 4 \times 2 = 16$ possible artifacts, and 4 different events that can be triggered at each artifact: "customer arrived" (A), "product picked up" (U), "product put down" (D) and "customer left" (L). The corresponding event network for a single consumer is shown in Fig. 3. Notice that, for practical reasons, certain event transitions are deemed invalid (e.g., no product interaction before the consumer's arrival at the artifact is detected). The dashed arrows indicate that transitions that may be valid if a consumer is allowed to permanently remove (and presumably purchase) a product from a display shelf.

Tables 2 and 3 summarize the descriptive statistics of the raw event dataset. In total, 3016 event observations are captured. The mean values of event occurrences suggest that approximately 6% of the events involved touch-based interaction between consumers and products displayed on the sensor-enabled shelves. The means for the artifact-dependent variables reflect our randomization strategy; the observations are fairly evenly split between locations (approximately 25% per location), product types (approximately 50% each for the sweatshirt and mug), and the presence of shelf interactivity (approximately 50% each for shelves with the lighting function enabled and disabled). Note that the location variables are coded with respect to the shelf's distance from the checkout counter. The event distribution is split roughly evenly at approximately 12 pm (time of day), and between Monday–Tuesday and Wednesday–Friday (day of week). For simplicity, we aggregate the timestamp data into binary artifact-independent variables accordingly.

4.2. Slicing data and deriving contextually meaningful variables

Given the alternatives, we opt for temporal slicing in the following analysis for several theoretical and practical reasons. First, the consumers' interactions with the sensor-enabled shelves are likely to be synchronous, since the general foot traffic in the store

Table 2
Descriptive statistics of raw event data.

Variables	N	Mean	SD	Min	Max
Events					
A: Customer arrived	3016	0.471	0.499	0.000	1.000
U: Product picked up	3016	0.030	0.169	0.000	1.000
D: Product put down	3016	0.029	0.166	0.000	1.000
L: Customer left	3016	0.471	0.499	0.000	1.000
Artifact-dependent					
Shelf location "very far"	3016	0.281	0.449	0.000	1.000
Shelf location "far"	3016	0.248	0.432	0.000	1.000
Shelf location "near"	3016	0.225	0.418	0.000	1.000
Shelf location "very near"	3016	0.246	0.431	0.000	1.000
Product type (1 = sweatshirt, 0 = mug)	3016	0.508	0.500	0.000	1.000
Shelf interactivity (1 = on, 0 = off)	3016	0.491	0.500	0.000	1.000
Artifact-independent					
Time of day (1 = before 12 pm)	3016	0.560	0.496	0.000	1.000
Day of week (1 = before Wednesday)	3016	0.579	0.494	0.000	1.000

Note: Location variables are coded w.r.t. checkout counter.

Table 3
Distribution of raw events over time.

Time of day	Day of week					Total
	Mon	Tue	Wed	Thu	Fri	
<10 am	7	301	217	206	30	761
10 am–12 pm	329	233	92	218	56	928
12 pm–2 pm	206	320	92	160	48	826
>2 pm	256	94	100	31	20	501
Total	798	948	501	615	154	3016

tends to be low and only one consumer is likely to be at a given shelf at a given point in time. Second, slicing by time offers more variation than slicing by spatial location, allowing us to build richer behavioral models. Unlike artifacts — whose number is essentially fixed (e.g., the set of shelves in a store) — we have more freedom in choosing the size of temporal slices. Our choice of the slice size may depend on considerations such as the rate at which events are typically emitted in the given context and what that implies for the recovery of individuals from the anonymized data. Third, consumer behavior in a temporally sliced event stream can be reframed as panel data, in which the time slicing yields the time variable and the artifact becomes the panel variable in the model. Finally, temporal slicing can leverage the fact that event streams are typically already ordered by time, so sorting the data again is not required.

Based on our qualitative observation of shopper behavior and anecdotal evidence from the store assistant, individual consumers tended to stay at a given shelf for approximately 1 minute, on average. Using the guided approach, we can thus break up the full dataset of 3016 events into minute-long slices, which yields approximately 90 slices (approximating individual consumers) with 16 artifacts each (one for every artifact). If all 90 of these approximately identified consumers were to visit all 16 artifacts, then we would have $90 \times 16 = 1440$ observations. However, since many of the individuals only visit a subset of the artifacts, we have 745 observations in the panel.

Due to the slicing, we can now construct individual-level outcome variables, and we do so using composite conditions. While there are a number of outcomes that we could consider, in this paper we focus on three conceptual outcome variables in particular: *touch*, *re-evaluation*, and *purchase*. Touch refers to the propensity of consumers to physically interact with the product, which in our case means picking up the product from the shelf. A growing stream of literature argues that touch — which is inherently an often underused sensory modality — should receive more attention from retailers to better understand consumer behavior with respect to the product offerings (Grohmann, Spangenberg, & Sprott, 2007; Krishna, 2012; Peck & Childers, 2008). The works of these authors also suggest that re-evaluating a product by touching it more than once may be a useful outcome variable to measure — repeated touch interactions may signify a greater interest in the product and an improved ability to evaluate the product. Finally, although we cannot directly observe product purchases in the event data, a contextually meaningful proxy may be to track the probability of a product being removed from a shelf for an unusually long period of time. Based on the store manager's anecdotal evidence, we can plausibly assume that a product is typically only removed from a shelf as a result of a purchase.

Table 4 summarizes the composite conditions for each of the three outcome variables, distinguishing between weak and strong conditions. Weak conditions should be easier to satisfy than strong conditions. For example, the conceptual outcome of "touch" is captured by event transitions in which a touch-related event occurs between a customer's arrival and departure from the artifact. Checking whether these event transitions occur at least once, as captured by the condition $(P(U|A) + P(D|A) > 0) \vee (P(L|U) +$

Table 4
Deriving outcome variables using composite conditions.

Outcome variables	Composite conditions
Touch	Weak condition: $(P(U|A) + P(D|A) > 0) \vee (P(L|U) + P(L|D) > 0)$ *"A product interaction occurred at least once."* Strong condition: $(P(U|A) + P(D|A) > P(L|A)) \vee (P(L|U) + P(L|D) > P(L|A))$ *"A product interaction occurred during most visits to the shelf."*
Re-evaluation	Weak condition: $P(U|D) > 0$ *"A product was re-evaluated (picked up again after being put down) at least once."* Strong condition: $(P(U|D) > P(L|D)) \vee (P(U|D) > P(U|A))$ *"A product was re-evaluated during most visits to the shelf."*
Purchase	Weak condition: $(P(D|A) > 0) \vee (P(L|U) > 0)$ *"A product was purchased (removed from the shelf) at least once."* Strong condition: $(P(D|A) > P(L|A)) \vee (P(L|U) > P(L|A))$ *"A product was purchased during most visits to the shelf."*

Notes: A: "customer arrived", U: "product picked up", D: "product put down", L: "customer left".
As an example, $P(U|A)$ reflects the probability of a customer picking up the product upon arriving at the shelf.

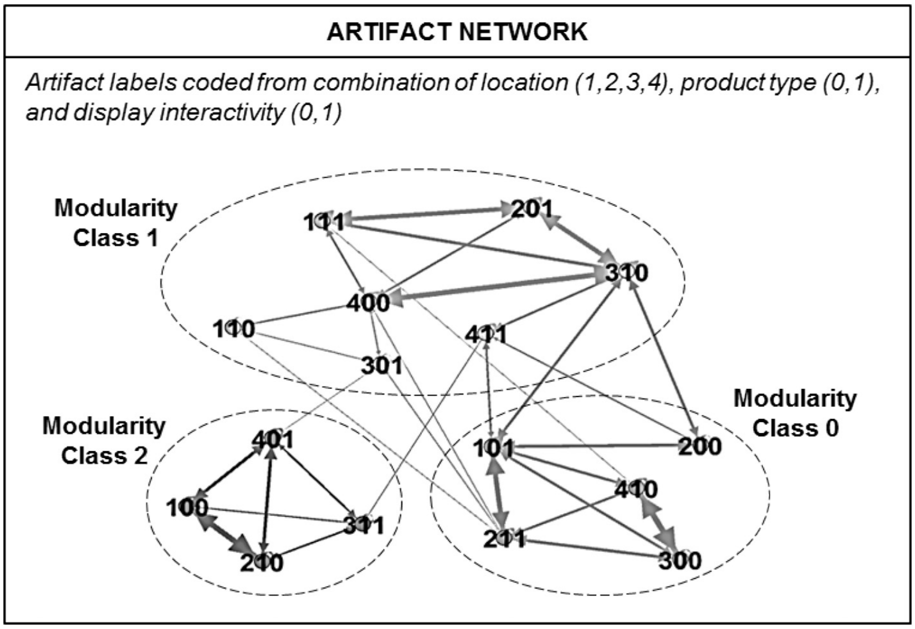

Fig. 4. Artifact network of in-store field experiment.

$P(L|D) > 0)$,[5] amounts to a weak condition. By contrast, a strong condition could require the probability of these touch-based event transitions to be greater than $P(L|A)$, which reflects a customer visit that does not involve touching the product. Similarly, the "re-evaluation" conditions capture the notion of repeated touch. The "purchase" conditions approximate the sequence of a customer's moves in which a product is picked up but not put back down by the same customer.

Each composite condition can now be analyzed as a binary outcome on its own, or the weak/strong condition pairs can be combined to form ordinal outcome variables. For example, in our dataset, the weak and strong conditions for "touch" are each satisfied 64 and 52 times, respectively; the weak condition alone is satisfied 12 times. Thus, we can derive an ordinal, three-level variable for "touch", coded 0 (if weak and strong conditions are both unsatisfied), 1 (if the weak condition alone is satisfied), and 2 (if the strong condition is also satisfied). The same can be done for the variables "re-evaluation" and "purchase." Note that we never have a situation where the strong condition is satisfied but the weak condition is not; this ensures the internal consistency of the ordinal outcome variable.

We can now derive predictors from the data. The artifact-independent predictors in our data are time-related and we expect these to be contextually relevant. A consumer's in-store behavior may be constrained by out-of-store factors (e.g., shopping goals, demographics) that are captured to some extent by the time of day and/or the day of the week of the store visit (Chandon, Hutchinson, Bradlow, & Young, 2009). Meanwhile, the raw artifact-dependent predictors are as follows:

- *Shelf location*: Previous research suggests that shelf location may affect the customer-product interaction (Russell & Urban, 2010). Some studies have also found that shoppers may avoid hovering sales assistants, implying a preference for unobserved buying (Kukar-Kinney, Ridgway, & Monroe, 2009). To analyze this in our experimental setting, we code the locations of the four sensor-enabled shelves based on their distance from the checkout counter where the store assistant tends to be (e.g., "very far," "far," "near" or "very near").
- *Product type*: Tactile product attributes, such as texture and thickness, are known to affect a consumer's desire to touch the product (Grohmann et al., 2007; McCabe & Nowlis, 2003). Previous studies show that consumers react differently to simple and complex-looking products (Blijlevens, Creusen, & Schoormans, 2009). In our case, the mug would be considered fairly simple (i.e., easy to size up), whereas the sweatshirt may require additional information processing by the consumer. We may expect the sweatshirt to register more touch-related events than the mug, as consumers try to gain a better sense of the product.
- *Shelf interactivity*: Based on attentional accounts of in-store consumer behavior, the sensor-based stimulus (the shelf lighting up) should attract the consumer's attention and draw her to the product display (Baker, Parasuraman, Grewal, & Voss, 2002; Kaltcheva & Weitz, 2006; Shankar, Inman, Mantrala, Kelley, & Rizley, 2011). Upon visiting the shelf, the consumer's attention may translate to a touch-based interaction with the product, involving re-evaluation and even purchase (Chandon et al., 2009).

[5] Notations: A: "customer arrived", U: "product picked up", D: "product put down", L: "customer left". As an example, $P(U|A)$ reflects the probability of a customer picking up the product upon arriving at the shelf.

As described in Section 3, we can also derive artifact-dependent predictors by constructing an artifact network. Fig. 4 shows the artifact network of our aggregate in-store AFE data, generated using the programming language Python and the network analysis software Gephi (Bastian, Heymann, & Jacomy, 2009). Each node represents a unique artifact, and each artifact is coded based on its specific combination of shelf location, product type and shelf interactivity. For instance, artifact "201" is located "far" from the checkout counter (2), displays a mug (0), and the shelf interactivity (lighting) has been enabled (1). The edges between the artifacts reflect the aggregation of the in-store paths of several shoppers over the course of the field experiment; thicker edges denote paths that were observed more frequently. For example, based on the edge thicknesses in Fig. 4, we see that several consumers moved between artifacts 100 and 210, 300 and 410, and 101 and 211. The most popular paths tended to occur between shelves placed directly next each other and shelves that have the same shelf interactivity setting. Moreover, clustering the artifact network using the modularity algorithm by Blondel et al. (2008) yields three different modularity classes, which can be used for descriptive analyses. Modularity class 1, for instance, encompasses the most artifacts and has the highest proportion of light-enabled shelves. Meanwhile, the location variable alone offers little differentiation between the modularity classes, since locations 1–4 are present in the artifacts of all three classes, although artifacts in modularity class 0 tend to be located farther away from the checkout counter.

4.3. Results

Table 5 shows the descriptive statistics for our dataset, temporally sliced into 1-minute intervals using a guided approach. As described in Section 3 (Table 1), we derive the following network-based predictors from the artifact network: weighted degree centrality, eigenvector centrality, betweenness centrality, and modularity class.

By interpreting the data as a panel of individual consumers interacting with artifacts in the retail environment (i.e., shelves in our study), we can estimate a negative binomial random effects model. Table 6 presents the results of estimating a full model for "touch" that includes the network-based measures and a partial model that does not. The Hausman (1978) test supports our use of the random effects model over the fixed effects alternative. Note that we do not model the remaining two outcome variables, re-evaluation and purchase, due to the sparsity of the data — the re-evaluation outcome was only observed 7 times, and the purchase outcome was observed only 16 times.

The dummy variables for location are partly significant with respect to the base shelf location "very near" in the regression models with and without the network measures, suggesting that more consumer activity took place at the shelves farthest away from the checkout counter. Past studies suggest that consumers would be less likely to touch products displayed on the two central shelves due to a perception of those products being "contaminated" by being generally touched more often (Argo, Dahl, & Morales, 2006). However, we do not find evidence of a contamination effect since the (non-central) shelf closest to the checkout counter accounts for an even lower level of touch activity than the central shelves. Given that the artifact features (location, product type, shelf interactivity) are randomized, our findings suggest a causal link between touch activity and shelf location, which might imply a tendency towards unobserved buying (Kukar-Kinney, Ridgway, & Monroe, 2009). Moreover, the regressions also suggest that the presence of a sweatshirt and shelf interactivity can increase the likelihood of a touch interaction taking place. The time of day is also a significant predictor, such that consumers appear to interact with the artifacts more in the morning. Note that, at no

Table 5
Descriptive statistics after temporal slicing.

Variables	*N*	Mean	SD	Min	Max
Outcomes (2 = strong, 1 = weak, 0 = none)					
Touch	745	0.156	0.521	0.000	2.000
Re-evaluation	745	0.015	0.159	0.000	2.000
Possible purchase	745	0.034	0.238	0.000	2.000
Predictors					
Artifact-dependent (non-network)					
Shelf location "very far"	745	0.281	0.450	0.000	1.000
Shelf location "far"	745	0.239	0.427	0.000	1.000
Shelf location "near"	745	0.226	0.418	0.000	1.000
Shelf location "very near"	745	0.255	0.436	0.000	1.000
Product type (1 = sweatshirt, 0 = mug)	745	0.468	0.499	0.000	1.000
Shelf interactivity (1 = on, 0 = off)	745	0.466	0.499	0.000	1.000
Artifact-dependent (network-based)					
Weighted degree centrality	745	458.977	174.967	76.000	782.000
Eigenvector centrality	745	0.679	0.278	0.236	1.000
Betweenness centrality	745	23.408	20.546	0.000	44.000
Modularity class	745	0.875	0.740	0.000	2.000
Artifact-independent					
Time of day (1 = before 12 pm)	745	0.383	0.486	0.000	1.000
Day of week (1 = before Wednesday)	745	0.848	0.359	0.000	1.000

Note: Location variables are coded w.r.t. checkout counter.

Table 6
Random effects models of the "touch" outcome.

Predictors	Outcome variable: touch	
	Partial model (without network measures)	Full model
Artifact-dependent (non-network)		
Shelf location "very far"[L]	−0.192	0.347
	[0.334]	[0.518]
Shelf location "far"[L]	−0.665*	0.067
	[0.369]	[0.526]
Shelf location "near"[L]	−0.240	−0.688*
	[0.358]	[0.386]
Product type (1 = sweatshirt, 0 = mug)	0.983***	1.082***
	[0.269]	[0.302]
Shelf interactivity (1 = on, 0 = off)	0.768***	−0.248
	[0.272]	[0.556]
Artifact-dependent (network-based)		
Weighted degree centrality		0.005**
		[0.002]
Eigenvector centrality		−10.995**
		[4.718]
Betweenness centrality		0.072**
		[0.033]
Modularity class 1[M]		−1.196
		[0.789]
Modularity class 2[M]		−3.992*
		[2.227]
Artifact-independent		
Time of day (1 = before 12 pm)	1.194***	1.165***
	[0.293]	[0.299]
Day of week (1 = before Wednesday)	−0.049	−0.371
	[0.521]	[0.565]
Constant	−3.569***	1.728
	[0.564]	[2.505]
Wald X^2	45.99***	63.92***
Observations	745	745
Number of groups (artifacts)	16	16

Negative binomial regressions with standard errors in brackets.
[L] Base shelf location is "very near"; location variables are coded w.r.t. checkout counter.
[M] Base modularity class is 0.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

point were the consumers primed or otherwise made aware of their participation in the field experiment. Care was therefore taken to make the experiment and its results as representative of a real-world retail setting as possible.

The regression analysis highlights the value of including network measures in modeling the consumers' touch-related behavior. All centrality measures are statistically significant, suggesting that the heterogeneity contained in the artifact network — which in turn can be traced back to the in-store movement patterns of consumers in the store — may explain some of the variance in the consumers' touch behavior.

It is interesting to note that, although eigenvector centrality is negatively correlated with touch, degree centrality and betweenness centrality are positively correlated with touch. The difference in the direction of correlation is attributable to the subtle difference in the computations of these centrality measures (Newman, 2010, pp. 168–192). The recursive eigenvector computations effectively consider all nodes in the network, thereby producing a global measure of nodal centrality. The eigenvector centrality of a node will be high if it is connected to relatively "important" (or central) nodes. Our finding that touch activity is negatively correlated with the eigenvector centrality of an artifact in the store is intriguing, since it suggests a possible cannibalizing effect whereby a central artifact ends up ceding some touch activity to its highly central neighboring artifacts. By contrast, weighted degree centrality, which is computed by taking the edge-weighted sum of a node's immediate neighbors regardless of their respective degree measures, is a more local measure. Through this lens we find that, as intuitively expected, a central artifact is likely to yield more touch activity in general. Meanwhile, betweenness centrality computes the extent to which a node lies on the shortest path between other node pairs. As seen in Fig. 4, the shelves with interactive lighting tend to be the artifacts with the highest betweenness in our study and appear to serve as "bridge builders" between otherwise disconnected clusters in the artifact network. Our regression results go further, by suggesting that such bridging artifacts also encourage touch activity.

The modularity class (i.e., node cluster) of an artifact also seems to affect the level of touch interaction. In particular, modularity class 2 is correlated with a lower level of touch activity than the base modularity class 0. The possibility of interpreting the modularity class as a latent measure of the effect of the observed artifact-dependent predictors (i.e., shelf location, product type, shelf interactivity) underscores the value of the network perspective (Blondel et al., 2008). For example, while the direct correlations between touch activity and the observed predictors are not always statistically significant in our regressions, the large effect of modularity class 2 suggests the presence of a latent property that makes some artifacts more amenable to touch than others.

To check the robustness of the above findings, we perform two types of additional tests. First, we repeat the guided approach that we have demonstrated so far with different temporal slice sizes (e.g., in the range of 1–3 minutes). Although the specific coefficient values change slightly, the results are similar to those shown in Table 6 in terms of the insights they produce. Second, we conduct the analysis using an unguided approach to temporal slicing. Since the guided approach yields approximately 90 slices, we let the unguided approach also slice the event stream into a similar number slices to allow for comparison between the two approaches. Appendix B shows the regression results of the unguided approach to highlight some interesting differences from Table 6. By artificially stipulating a fixed number of evenly sized slices, the unguided approach produces a larger panel dataset (1440 observations across the 16 artifacts). The effect sizes and statistical significance are seemingly larger — and are potentially exaggerated due to the artificialness of fixing the slices — in the unguided approach.

Moreover, we investigate the predictive validity of the full model of touch activity. As touch activity does not occur in the majority of the observations, the data is unbalanced. Assessing predictive validity by fitting the model on unbalanced data can lead to low model performance in terms of precision and recall, since the model will tend to err on the side of predicting no touch activity for a given observation in the holdout sample even when touch activity actually did occur. One approach to address the problem of an unbalanced ordinal outcome variable is to oversample the levels of the variable that are in the minority (i.e., weak and strong touch in our case). We employ the commonly used synthetic minority over-sampling technique (SMOTE) to produce a balanced sample for predictive purposes (Chawla, Bowyer, Hall, & Kegelmeyer, 2002). For the case of guided slicing, SMOTE generates a balanced dataset consisting of 681 observations for each of the three levels of the touch variable (no touch, weak and strong touch).[6] For unguided slicing, the balanced dataset contained 1332 observations for each level of touch. We fit the full model for guided and unguided slicing on a randomly selected training sample of 70% of the respective data, and predictions are carried out on the remaining holdout sample. The results of the holdout testing are documented in Appendix C. Note that, randomly guessing the value of a three-level variable should yield an expected predictive accuracy, precision and recall of about 33%. With both guided and unguided slicing, the full models of touch tend to achieve accuracy, precision and recall statistics between 60 and 80%, and are thus substantially better than the random expectation in the training and holdout samples. The results are indicative of strong predictive validity.

5. Discussion and conclusion

Advances in sensor-based and mobile technology have allowed retailers to track consumers at the individual level in online and offline settings using event-based data (Shankar et al., 2011). In essence, the data captures a sequence of events that reflect actions taken by the consumer in various scenarios (e.g., browsing online, traveling to stores and movement inside a store). Recent literature has led to some important insights by looking at event-based data that tracks individuals or sensor-enabled devices on their person (Hui et al., 2009a; Larson et al., 2005). However, growing concerns related to the data privacy of consumers has led various government regulators across Europe, America and Asia to begin tightening the laws governing the collection and use of consumer data by retailers (Maras, 2015; Weber, 2015). Meanwhile, a growing awareness of the data privacy risks among consumers has seen an uptick in the usage of tools to block trackers that collect data on consumer behavior in online and offline environments. The upshot of the privacy challenge — especially as we enter the age of ubiquitous digitization and the Internet of Things — is that anonymized and fragmented event-based (AFE) data will increasingly account for a large portion of the customer tracking data available to retailers. Unlike aggregate data and the types of individual-level data that have been typically used by researchers and retailers in the past, AFE data represents a novel form of data that balances the preservation of individual-level heterogeneity with an enhanced protection of consumer privacy. In particular, AFE data ensures a high level of privacy that is in line with regulatory prescriptions (e.g., the new European General Data Protection Regulation), while a combination of computational techniques (as presented in our methodology, for instance) allows us to approximately recover the individual-level heterogeneity captured in the data.

In this context, the work presented in this paper makes several contributions to marketing research as follows:

- We aim to fill the gap in the existing literature by developing a theoretically grounded and practically usable methodology to enable the analysis of consumer behavior using AFE data. Our research is thus well-aligned with the recent calls for research that emphasize the need for new methods to analyze such data in retail (Grewal, Roggeveen, & Nordfält, 2016; Wedel & Kannan, 2016).
- Using heuristics that are common in research on event-based data (Nagarajan et al., 2009), we propose that the raw data can be sliced along one or more dimensions (e.g., temporal, spatial, or a combination thereof) to approximately recover the information about individual-level heterogeneity. The slicing heuristic can also be guided by contextual information that the retailer may know about the consumers and the retail setting to help the slicing achieve a more accurate representation of reality.
- We address the difficulty of deriving contextually meaningful outcome variables and explanatory variables from the sliced data to facilitate descriptive and predictive analyses. We build on the theory of event transition probabilities to develop an elegant

[6] There were 681 "no touch" observations in the unbalanced dataset. By default, SMOTE oversampled the minority variable levels ("weak touch" and "strong touch") to achieve 681 observations for all variable levels.

formulation of so-called composite conditions, which can enable the flexible combination of several event transition probabilities into a conceptually meaningful outcome variable.

- We suggest the use of network methods to derive a rich set of explanatory variables that capture the heterogeneity embedded in the relationships between different artifacts in the retail environment; this heterogeneity arises from the paths that different consumers take and the sequence in which the artifacts are visited along the way. We argue that the network-based measures can complement other explanatory variables related to an artifact's location, characteristics of the displayed product, and the time of day/week.
- We validate the methodology by applying it to a carefully designed, randomized field experiment in an in-store setting. By deploying sensor-based shelves to display the store's merchandise, we are able to generate AFE data that reflects the movement patterns of consumers within the store. The empirical data allows us to show the value of guided slicing and using network measures to improve the fitness of the resulting models of consumer behavior.

Our work has some interesting implications for marketing research. First and foremost, we draw attention to the opportunities and challenges associated with AFE data, juxtaposing it with other forms of aggregate and individual data that have been the mainstay of marketing research so far (Ghose & Yang, 2009). Rather than ignoring the hidden, individual-level information contained in the anonymized data by treating it as aggregate data, we suggest that it may be worthwhile to recover such information. Second, our methodological workflow for analyzing AFE data — going from raw, unstructured event streams to insights on shopper behavior — can be seen as an extension of the general framework described by Hui et al. (2009a) for analyzing "path data" in marketing. Path data broadly refers to the sequence of an individual's movement in a given environment (e.g., online clickstreams, eye-tracking data, and offline movement patterns in a store) and is thus closely related to event-based data. While it is likely that Hui et al.'s framework was primarily developed for path data that explicitly tracks individuals, we show that their framework can also be extended and operationalized to facilitate the analysis of AFE data. Third, our comprehensive methodology highlights the value of transferring knowledge (e.g., concepts and techniques) from other fields to make use of event-based "Big Data" in marketing science. As discussed previously by Einav and Levin (2014), Big Data is fundamentally different to the structured, "rectangular" data that marketing scholars typically work with; Big Data tends to be unstructured, fragmented across data sources and overwhelming in terms of volume, velocity and variety. In the coming years, Big Data is likely to account for a larger portion of the type of AFE data we study in this paper.

The methodological and empirical work presented in this paper also has important managerial implications. The above findings point to some interesting implications for the specific store in our study. If the store manager wants to increase the touch activity at a certain group of chosen products (e.g., newly launched products or even flagship products), she could try placing the products farther away from the checkout counter (and thus away from the sales assistants) to facilitate unobserved buying, and on interactive shelves to attract attention. Based on the network perspective, the manager could increase touch activity for the chosen products by increasing the degree and betweenness centrality of these products; this could be achieved by placing them in places of high visibility and foot traffic within the store. However, since eigenvector centrality reduces touch activity, the manager should ensure that the chosen products are not directly surrounded by other highly central products.

Regarding implications on a broader scale, we highlight the relevance of the privacy challenge for retailers in the coming years. Stricter government regulations on data privacy and increasing consumer awareness of data-related risks are likely to place limits on the ability of retailers to collect sensitive, individual-level customer data. At the same time, ubiquitous digitization is creating a wealth of data that managers can potentially use to better serve their customers. The related conundrum that retailers face has been called the "personalization privacy paradox" in past literature (Aguirre, Mahr, Grewal, de Ruyter, & Wetzels, 2015; Awad & Krishnan, 2006); the idea is that privacy-conscious consumers are less likely to part with their data to allow personalization despite its benefits. Our work allows retailers to address the privacy challenge by shifting the focus to mining AFE data, which does not directly expose specific consumers. AFE data — and methods to analyze it — will arguably also gain in importance in the post-GDPR world, as businesses endeavor to achieve regulatory compliance with stricter data use policies while using marketing analytics to better serve customers. Furthermore, our methodology underscores an important point made in the extant literature (Bradlow et al., 2017) — despite the allure created by its sheer scale and granularity, more data may not translate directly into commercially relevant insights. While the concept of "data cleaning" will not be new for data analysts, it becomes even more important in the context of AFE data; as our methodology and empirical study demonstrates, AFE data in its raw form is likely to require significant preprocessing in terms of reorganizing and structuring to yield meaningful variables for further analysis. Finally, the practical applicability of our methodology is an attractive feature of our work for retail managers. Through the empirical study, we have shown that using our methodology can help extract interesting insights from AFE data in a real-world environment. The methodology is also sufficiently modular and flexible for managers to customize it to their needs, both in an online and offline setting.

Nonetheless, the methodology presented in this paper has two main limitations. First, while the use of a slicing heuristic can help approximately separate the events attributable to different individuals, it does not let us detect repeat visits by the same individual. For example, the sensor-based data in our empirical study does not keep track of whether a certain consumer visited the store — and interacted with the products — on multiple days. Each slice of the event stream is implicitly treated as a different and independent individual. Discerning this particular case of interaction between consumers and artifacts from AFE data alone may be generally intractable and would arguably require some outside insight in practice (e.g., past experience or qualitative observation). Second, there may be — at least in a managerial sense — a conceptual limit to the interpretability of variables generated from composite conditions. Conditions composed of more than four or five transition probabilities may become too difficult to interpret in practice. Nevertheless, our methodology could serve as a useful starting point for deriving valuable insights from AFE data.

Future work can address the limitations of the methodology and its validation. One possible avenue is to explore more sophisticated methods of slicing that combine the spatial and temporal dimensions to recover individual-level heterogeneity more accurately. The accuracy of the slicing could be validated using a combined dataset that is collected both as AFE data and as individual-level data using tracking equipment like the PathTracker belt worn by shoppers in Hui et al. (2009a,b) — the better the slicing heuristic, the closer it would be to the individual-level data. Similarly, alternatives to the network representation of AFE data can be explored. For example, Heins and Stern (2014) describe an approach to modeling event data sequences that may be used to detect possible patterns in aggregate or individual-level data; the applicability of such approaches to AFE data can be investigated. Finally, we would also encourage researchers to validate our approach to analyzing AFE data in other retail scenarios. For example, it may be possible to partner with an online retailer to procure anonymized clickstream data (e.g., which obscures part of a user's IP address, location, or other sensitive details); the artifacts in an online store may be webpages or hyperlinks that users can interact with to produce the clickstream.

Acknowledgements

The authors gratefully acknowledge the provision of smart display shelves by SAP Hybris and logistical support by Uni Merch Merchandising GmbH during the field experiment. The authors also thank Dominik Molitor for providing useful feedback on an earlier revision of the article.

Appendix A. Store layout in field study

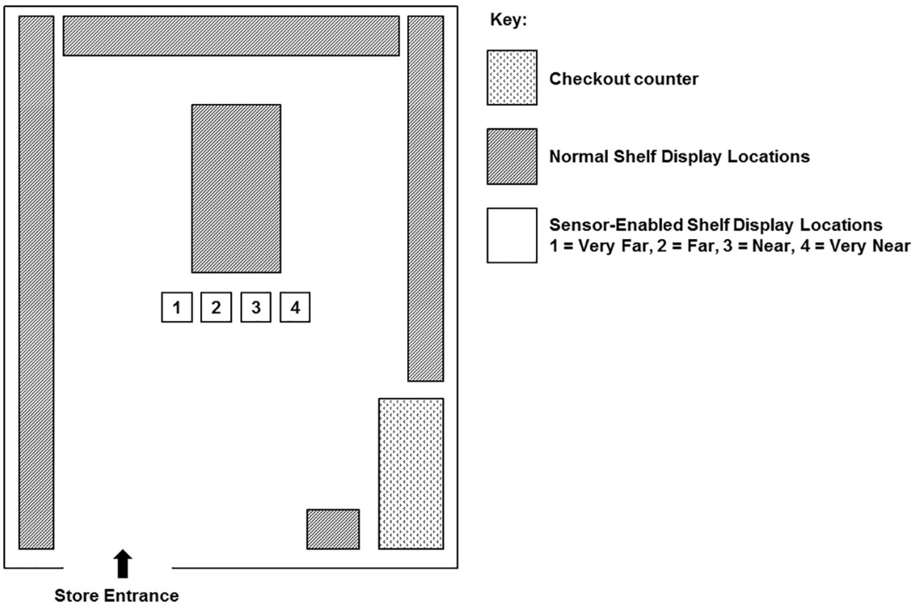

Fig. A1. Illustration of store layout in field study.

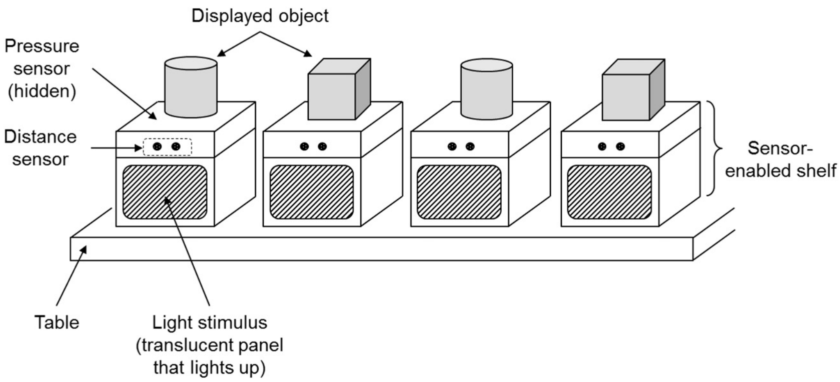

Fig. A2. Illustration of sensor-enabled shelves in a row.

Appendix B. Robustness of results to unguided approach

Table A1
Random effects models of "touch" outcome with unguided approach.

Predictors	Outcome variable: touch	
	Partial model (without network measures)	Full model
Artifact-dependent (non-network)		
Shelf location "very far"[L]	−0.122	0.496
	[0.473]	[0.415]
Shelf location "far"[L]	−0.804*	0.670*
	[0.486]	[0.403]
Shelf location "near"[L]	−1.107*	−2.321***
	[0.621]	[0.453]
Product type (1 = sweatshirt, 0 = mug)	1.312***	0.926***
	[0.380]	[0.249]
Shelf interactivity (1 = on, 0 = off)	−0.057	−1.518***
	[0.396]	[0.444]
Artifact-dependent (network-based)		
Weighted degree centrality		0.010***
		[0.002]
Eigenvector centrality		−16.042***
		[3.492]
Betweenness centrality		0.117***
		[0.024]
Modularity class 1[M]		−1.440**
		[0.573]
Modularity class 2[M]		−6.318***
		[1.643]
Artifact-independent		
Time of day (1 = before 12 pm)	0.584***	0.586***
	[0.214]	[0.213]
Day of week (1 = before Wednesday)	0.833***	0.822***
	[0.253]	[0.252]
Constant	−3.133***	2.585
	[0.534]	[1.775]
Wald X^2	46.83***	109.07***
Observations	1440	1440
Number of groups (artifacts)	16	16

Negative binomial regressions with standard errors in brackets.
Event stream is sliced into 90 evenly sized slices.
[L] Base shelf location is "very near"; location variables are coded w.r.t. checkout counter.
[M] Base modularity class is 0.
* $p < 0.1$.
** $p < 0.05$.
*** $p < 0.01$.

Appendix C. Predictive validity of full models for touch outcome

Table A2
Holdout testing results.

Overall predictive accuracy	Guided slicing		Unguided slicing	
	Training sample ($N = 1436$)	Holdout sample ($N = 607$)	Training sample ($N = 2774$)	Holdout sample ($N = 1222$)
	60.8%	61.8%	77.4%	77.2%
Touch = 0 (no touch)				
Precision	59.5%	63.5%	71.2%	72.4%
Recall	58.7%	59.7%	67.4%	67.5%
Touch = 1 (weak touch)				
Precision	47.2%	48.0%	76.6%	75.2%
Recall	41.9%	41.2%	64.0%	65.8%
Touch = 2 (strong touch)				
Precision	72.4%	70.6%	82.5%	82.7%
Recall	81.2%	86.1%	100.0%	100.0%

Overall predictive accuracy = (True Positives for Touch = 0, 1 and 2) / N.
Precision = True Positives / (True Positives + False Positives).
Recall = True Positives / (True Positives + False Negatives).
Synthetic minority over-sampling technique (SMOTE) is used to balance the three levels of Touch in the data used for testing predictive validity.

References

Aguirre, E., Mahr, D., Grewal, D., de Ruyter, K., & Wetzels, M. (2015). Unraveling the personalization paradox: The effect of information collection and trust-building strategies on online advertisement effectiveness. *Journal of Retailing*, *91*(1), 34–49.

Argo, J. J., Dahl, D. W., & Morales, A. C. (2006). Consumer contamination: How consumers react to products touched by others. *Journal of Marketing*, *70*(2), 81–94.

Argo, J. J., Dahl, D. W., & Morales, A. C. (2008). Positive consumer contagion: Responses to attractive others in a retail context. *Journal of Marketing Research*, *45*(6), 690–701.

Awad, N. F., & Krishnan, M. S. (2006). The personalization privacy paradox: An empirical evaluation of information transparency and the willingness to be profiled online for personalization. *MIS Quarterly*, 13–28.

Baker, J., Parasuraman, A., Grewal, D., & Voss, G. B. (2002). The influence of multiple store environment cues on perceived merchandise value and patronage intentions. *Journal of Marketing*, *66*(2), 120–141.

Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. *ICWSM*. *8*. (pp. 361–362).

Blijlevens, J., Creusen, M., & Schoormans, J. (2009). How consumers perceive product appearance: The identification of three product appearance attributes. *International Journal of Design*, *3*(3), 27–35.

Blondel, V. D., Guillaume, J. -L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008.

Borgatti, S. P., Mehra, A., Brass, D. J., & Labianca, G. (2009). Network analysis in the social sciences. *Science*, *323*(5916), 892–895.

Bradlow, E. T., Gangwar, M., Kopalle, P., & Voleti, S. (2017). The role of big data and predictive analytics in retailing. *Journal of Retailing*, *93*(1), 79–95.

Bucklin, R. E., & Sismeiro, C. (2009). Click here for internet insight: Advances in clickstream data analysis in marketing. *Journal of Interactive Marketing*, *23*(1), 35–48.

Burke, R. R. (2010). The third wave of marketing intelligence. *Retailing in the 21st century* (pp. 159–171). Springer.

Chandon, P., Hutchinson, J. W., Bradlow, E. T., & Young, S. H. (2009). Does in-store marketing work? Effects of the number and position of shelf facings on brand attention and evaluation at the point of purchase. *Journal of Marketing*, *73*(6), 1–17.

Chawla, N. V., Bowyer, K. W., Hall, L. O., & Kegelmeyer, W. P. (2002). SMOTE: Synthetic minority over-sampling technique. *Journal of Artificial Intelligence Research*, *16*, 321–357.

Dhar, V., Geva, T., Oestreicher-Singer, G., & Sundararajan, A. (2014). Prediction in economic networks. *Information Systems Research*, *25*(2), 264–284.

Einav, L., & Levin, J. (2014). Economics in the age of big data. *Science*, *346*(6210), 1243089.

Ghose, A., & Yang, S. (2009). An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Science*, *55*(10), 1605–1622.

Grewal, D., Roggeveen, A., & Nordfält, J. (2016). The future of retailing. *Journal of Retailing*, *93*(1), 1–6.

Grohmann, B., Spangenberg, E. R., & Sprott, D. E. (2007). The influence of tactile input on the evaluation of retail product offerings. *Journal of Retailing*, *83*(2), 237–245.

Guralnik, V., & Srivastava, J. (1999). Event detection from time series data. *Proceedings of the fifth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 33–42).

Harrell, G. D., Hutt, M. D., & Anderson, J. C. (1980). Path analysis of buyer behavior under conditions of crowding. *Journal of Marketing Research*, *17*, 45–51.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica: Journal of the Econometric Society*, *46*(6), 1251–1271.

Heins, K., & Stern, H. (2014). A statistical model for event sequence data. *Proceedings of the seventeenth international conference on artificial intelligence and statistics* (pp. 338–346).

Hui, S. K., Fader, P. S., & Bradlow, E. T. (2009a). Path data in marketing: An integrative framework and prospectus for model building. *Marketing Science*, *28*(2), 320–335.

Hui, S. K., Fader, P. S., & Bradlow, E. T. (2009b). The traveling salesman goes shopping: The systematic deviations of grocery paths from TSP optimality. *Marketing Science*, *28*(3), 566–572.

Hui, S. K., Huang, Y., Suher, J., & Inman, J. J. (2013a). Deconstructing the "first moment of truth": Understanding unplanned consideration and purchase conversion using in-store video tracking. *Journal of Marketing Research*, *50*(4), 445–462.

Hui, S. K., Inman, J. J., Huang, Y., & Suher, J. (2013b). The effect of in-store travel distance on unplanned spending: Applications to mobile promotion strategies. *Journal of Marketing*, *77*(2), 1–16.

Ilfeld, J. S., & Winer, R. S. (2002). Generating website traffic. *Journal of Advertising Research*, *42*(5), 49–61.

Inman, J., & Nikolova, H. (2017). Shopper-facing retail technology: A retailer adoption decision framework incorporating shopper attitudes and privacy concerns. *Journal of Retailing*, *93*(1), 7–28.

Kaltcheva, V. D., & Weitz, B. A. (2006). When should a retailer create an exciting store environment? *Journal of Marketing*, *70*(1), 107–118.

Kimball, R., & Ross, M. (2009). *The data warehouse toolkit: The complete guide to dimensional modeling* (2nd ed.). New York: Wiley.

Krishna, A. (2012). An integrative review of sensory marketing: Engaging the senses to affect perception, judgment and behavior. *Journal of Consumer Psychology*, *22*(3), 332–351.

Kukar-Kinney, M., Ridgway, N. M., & Monroe, K. B. (2009). The relationship between consumers' tendencies to buy compulsively and their motivations to shop and buy on the internet. *Journal of Retailing*, *85*(3), 298–307.

Kumar, R., Mahadevan, U., & Sivakumar, D. (2004). A graph-theoretic approach to extract storylines from search results. *Proceedings of the tenth ACM SIGKDD international conference on knowledge discovery and data mining* (pp. 216–225).

Larson, J. S., Bradlow, E. T., & Fader, P. S. (2005). An exploratory look at supermarket shopping paths. *International Journal of Research in Marketing*, *22*(4), 395–414.

Luo, X., Andrews, M., Fang, Z., & Phang, C. W. (2014). Mobile targeting. *Management Science*, *60*(7), 1738–1756.

Maras, M. -H. (2015). Internet of things: Security and privacy implications. *International Data Privacy Law*, *5*(2), 99.

McCabe, D. B., & Nowlis, S. M. (2003). The effect of examining actual products or product descriptions on consumer preference. *Journal of Consumer Psychology*, *13*(4), 431–439.

Merzdovnik, G., Huber, M., Buhov, D., Nikiforakis, N., Neuner, S., Schmiedecker, M., & Weippl, E. (2017). Block me if you can: A large-scale study of tracker-blocking tools. *IEEE European symposium on security and privacy (EuroS&P)* (pp. 319–333).

Molitor, D., Reichhart, P., Spann, M., & Ghose, A. (2016). *Measuring the Effectiveness of Location-Based Advertising: A Randomized Field Experiment.* Available at https://ssrn.com/abstract=2645281.

Montgomery, A. L., Li, S., Srinivasan, K., & Liechty, J. C. (2004). Modeling online browsing and path analysis using clickstream data. *Marketing Science*, *23*(4), 579–595.

Nagarajan, M., Gomadam, K., Sheth, A. P., Ranabahu, A., Mutharaju, R., & Jadhav, A. (2009). Spatio-temporal-thematic analysis of citizen sensor data: Challenges and experiences. *International conference on web information systems engineering* (pp. 539–553).

Newman, M. (2010). *Networks: An introduction* (9th ed.). Oxford: Oxford University Press.

Peck, J., & Childers, T. L. (2008). Sensory factors and consumer behavior. *Handbook of consumer psychology* (pp. 193–219).

Phua, P., Page, B., & Bogomolova, S. (2015). Validating Bluetooth logging as metric for shopper behaviour studies. *Journal of Retailing and Consumer Services*, *22*, 158–163.

Russell, R. A., & Urban, T. L. (2010). The location and allocation of products and product families on retail shelves. *Annals of Operations Research*, *179*(1), 131–147.

Shankar, V., Inman, J. J., Mantrala, M., Kelley, E., & Rizley, R. (2011). Innovations in shopper marketing: Current insights and future research issues. *Journal of Retailing*, *87*(S1), S29–S42.

Sherman, L., & Deighton, J. (2001). Banner advertising: Measuring effectiveness and optimizing placement. *Journal of Interactive Marketing*, *15*(2), 60–64.

Singh, S. (2006). Impact of color on marketing. *Management Decision*, *44*(6), 783–789.

Sismeiro, C., & Bucklin, R. E. (2004). Modeling purchase behavior at an e-commerce web site: A task-completion approach. *Journal of Marketing Research*, *41*(3), 306–323.

Storey, G., Reisman, D., Mayer, J., & Narayanan, A. (2017). *The Future of Ad Blocking: An Analytical Framework and New Techniques.* Available at https://arxiv.org/abs/1705.08568.

Tellis, G. (2004). *Effective advertising.* Thousand Oaks, CA: Sage.
Varian, H. R. (2014). Big data: New tricks for econometrics. *The Journal of Economic Perspectives*, *28*(2), 3–27.
Wang, J. (2013). *Handbook of finite state based models and applications. Discrete mathematics and its applications.* Boca Raton: CRC Press.
Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications. Structural analysis in the social sciences. Vol. 8.*Cambridge: Cambridge University Press.
Weber, R. H. (2015). The digital future—A challenge for privacy? *Computer Law and Security Review*, *31*(2), 234–242.
Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing*, *80*(6), 97–121.
Wooldridge, J. (2013). *Introductory econometrics: A modern approach* (5th ed.). Ohio: South-Western.
Yang, Y., Pierce, T., & Carbonell, J. (1998). A study of retrospective and on-line event detection. *Proceedings of the 21st annual international ACM SIGIR conference on research and development in information retrieval* (pp. 28–36).